DEPARTAMENTO DE INFORMACIÓN Y COMUNICACIÓN

PROGRAMA OFICIAL DE POSGRADO EN CIENCIAS SOCIALES

CLASSIFYING AND VISUALIZING THE DISCIPLINARY FOCUS OF UNIVERSITIES

THE INVISIBLE FACTOR OF UNIVERSITY RANKINGS

NICOLÁS ROBINSON-GARCÍA

TESIS DOCTORAL
GRANADA, 2014

Editor: Editorial de la Universidad de Granada
Autor: Nicolás Robinson-García
D.L.: GR 2014-2014
ISBN: 978-84-9083-205-9

El doctorando **Nicolás Robinson García** y los directores de la tesis **Evaristo Jiménez Contreras** y **Emilio Delgado López-Cózar** garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, a de de 2014

Director/es de la Tesis Doctorando

Fdo.: Fdo.:

Acknowledgments

The present thesis is the result of a four-year journey during which I have been able to work with and learn from many people who have played different roles at different stages of this trip. This thesis is intended to reflect my initiation in the research profession. Formally, this has been possible thanks to an FPU Grant from the Ministerio de Economía y Competitividad, as well as other grants which have allowed me to enjoy a three-month research stay at Leiden University in The Netherlands and participate in different congresses and seminars. However, there are many people who have been implicated personally in my career development who I wish to thank.

First of all I would like to thank Evaristo and Emilio, my supervisors, who have guided and supported me on this journey. Thank you Evaristo for your constant support and encouragement. Thank you Emilio for our endless conversations and discussions on research evaluation. You have certainly stimulated my longing to learn and undertake research. Both of you have played important, complementary roles in my research development for which I am grateful.

I would especially like to thank Daniel for everything he has taught me. He has been a mentor and a friend. It has been a delight to work side by side with you. I consider myself privileged to have been able to collaborate with you. Moreover, being able to argue about bibliometrics in a bar over a couple of beers is always helpful. You even like Pink Floyd, which is undoubtedly a sign of good taste and knowledge!

Thanks to my colleagues in the EC3 Research Group, especially to Álvaro, with whom I first worked when I joined as a research assistant, and to Enrique, with whom I studied the Licenciatura and Master and who got on board in the EC3 research group at the same time I did. And of course to the rest of my colleagues: Rafa, a.k.a. 'The Godfather', Repi, Diego, Alberto, Juanma and Orduña.

Thanks to Clara and Rodrigo who I met and with whom I worked during my research stay in Leiden. They have become friends as well as colleagues and have supported me constantly since then. After all, it is our research isn't it?

Thanks to all the people with whom I have been able to collaborate during these four years. Thanks to the team from the I-UGR Rankings: Paco, Jose, Isaac. Thanks to Jose Antonio, Rosa, Joaquín, Campanario... It has been a great pleasure to work with you.

My parents and my sister have played an important role on my professional development. They are my unconditional fans and supporters. Lydia, always enthusiastic and proud of anything I do sometimes even makes me feel embarrassed as I don't think I deserve it. Bryan, you've been there and done that, and know how difficult it gets sometimes. Thank you for being my first reviewer and supporting me no matter what. Mom, thanks for listening to my whining and moaning and for encouraging me and worrying about me ALL THE TIME!!! Despite the incomprehensibility of what we bibliometricians do, they have always been there for me. Listening to me when I was excited about some weird indicator (let's say the IFQ^{2}A Index, for example) as well as when I was low-spirited because things didn't work out. I am more than conscious of how stubborn I can get and how difficult it can often be to deal with me. Thank you.

And last but not least, Patricia. What can I say? You know a great deal about dealing with me, don't you? You don't know how grateful I am. You were next to me when I started and you are still here with me. I must have done a pretty good job! Thank you very much for your love and support. Without you I would have never made it. It is time now to enjoy the benefits of your efforts!

Resumen

Introducción

Los rankings de universidades internacionales se han convertido en una herramienta de gran importancia para los gestores universitarios y de política científica. Han cobrado tal fuerza que forman parte del discurso de políticos, perfilan la imagen de las universidades e influyen de manera determinante en la elección de alumnos e investigadores a la hora de elegir un centro académico al que vincularse. Al convertirse en parte de nuestro código social (Hakanen, 2002), tienen un atractivo en la sociedad que hace que sea imposible evadirse de ellos. Esta atracción se debe a su capacidad para reducir una realidad tan compleja como es la 'calidad' o el rendimiento científico de una universidad. Lejos de ofrecer descripciones precisas y sencillas del objeto analizado, se centran en indicar dónde se sitúa dicho objeto con respecto a otros. De esta manera, el lector observa dónde está su institución con respecto a las demás y no en qué características definen a esta institución y en qué aspectos es destacable.

Además, los rankings no sólo transforman nuestra forma de evaluar y 'elegir', sino que también se internalizan rápidamente en nuestra forma de percibir la realidad. En un estudio sobre el efecto de los rankings entre decanos de facultades de Derecho en los Estados Unidos, Sauder y Espeland (2009) descubrieron que incluso en los casos en los que los decanos se mostraron más reacios a incorporar estas herramientas a la hora de tomar decisiones, al final se vieron forzados a tomarlas en cuenta debido a lo influyentes que eran en su entorno. Desde la irrupción del Ranking de Shanghai, el primer ranking internacional de universidades, a principios del presente siglo, la influencia de estas herramientas no ha hecho más que crecer y actualmente se han convertido en la verdadera vara de medir de las universidades a nivel mundial. Sus metodologías no sólo son tenidas en cuenta, sino que se incluyen en los propios planes estratégicos de las universidades para así mejorar su posicionamiento en futuras ediciones (Hazelkorn, 2011).

En este sentido, la mayor parte de los rankings se centran en el análisis del rendimiento científico de las universidades. Esto se debe principalmente, al amplio desarrollo de indicadores de actividad científica que existe proveniente del campo de la bibliometría. Así, entran en escenas las dos principales empresas que proveen datos bibliométricos sobre las publicaciones científicas:

Thomson Reuters, a través de la base de datos Web of Science, y Elsevier, a través de Scopus. Estas dos bases de datos nutren a los principales rankings internacionales de donde extraen la información necesaria para medir el rendimiento científico de las universidades que finalmente será determinante para establecer el ranking.

Sin embargo, son muchas las limitaciones que presentan los rankings desde el punto de vista bibliométrico, lo que hace que desde la comunidad científica, sean vistos con recelo (van Raan, 2005). Las principales limitaciones tienen que ver con cuestiones técnicas en la recogida de datos (Waltman, et al., 2012), sesgos idiomáticos (Altbach, Reisberg & Rumbley, 2010) y sesgos disciplinares (Ishikawa, 2009).

Objetivos de la tesis

En el presente trabajo, presentado mediante un compendio de cinco artículos de revista, se analizan dos aspectos relativos con el uso de los rankings internacionales para la evaluación científica. En primer lugar, la necesidad de construir rankings nacionales que complementen la información que proveen los internacionales, donde países como España están infrarrepresentadas. En segundo lugar, la necesidad de desarrollar herramientas que permitan la estratificación horizontal de las universidades, partiendo de la premisa de que se trata de instituciones muy heterogéneas entre sí que no son percibidas como tales en los rankings. Mientras que éstos hacen peligrar la diversificación institucional mediante una estratificación vertical de las universidades (Marginson, 2008), la estratificación horizontal pretende establecer distintos perfiles institucionales que nos permitan realizar comparaciones entre universidades similares entre sí.

Específicamente, se plantean las siguientes preguntas de investigación:

1. *¿Son los rankings nacionales de universidades necesarios en un contexto internacional?* Los rankings internacionales de universidades no se están empleando únicamente para establecer políticas institucionales, sino también para evaluar y reformar sistemas universitarios nacionales. En este sentido, la poca cobertura que estos rankings a sistemas como el español, hacen que resulte peligroso considerar a las pocas universidades presentes en los mismos como una muestra fiable del sistema. En este sentido, se plantea la creación de rankings nacionales de universidades que complementen la información ofrecida por los otros.

2. *¿Cómo podemos desarrollar herramientas que permitan a los gestores de política científica identificar distintos perfiles disciplinares de universidad?* Los rankings por campos y disciplinas científicas han empezado a proliferar en los años. Sin embargo, plantean problemas de interpretación derivados de la construcción de los campos científicos. Para dar solución a esta pregunta, se analizan distintas metodologías de visualización de la ciencia, debido a la importancia que están cobrando los mapas científicos de cara a la toma de decisiones (Noyons 2005).

Justificación y estructura de la tesis

Esta tesis surge como resultado de mi participación dentro del equipo que desarrolla los Rankings I-UGR de Universidades Españolas por Campos y Disciplinas Científicas (http://rankingsuniversidades.es). Aunque de carácter nacional, los Rankings I-UGR nacen precisamente como respuesta a la existencia de rankings internacionales. Las universidades españolas tienen poca autonomía (Estermann, Nokkola & Steinel, 2011) y una alta dependencia en la financiación pública que se ha repartido históricamente en función del número de estudiantes, lo que hace que se trate de un sistema nacional donde hay poca competencia entre universidades. Esta falta de competitividad hace que el país no esté preparado a hacer frente al impacto de los rankings internacionales en concreto, y en general, al impacto de un nuevo escenario de educación universitaria ampliamente globalizado.

Todo esto deriva en la falta de información que tienen los gestores de política científica españoles sobre el posicionamiento del sistema español en referencia a los de otros países. Los rankings I-UGR pretenden precisamente servir de herramienta para ofrecer dicha información (Torres-Salinas et al., 2011). Además, plantean la necesaria importancia que tiene ofrecer listados de rankings por campos y disciplinas en lugar de una tabla general, respetando así, los distintos perfiles de especialización disciplinar que caracterizan a cada universidad. Esta tesis nace y bebe del trabajo anterior desarrollado por el equipo de los Rankings I-UGR, y pretende ser mi aportación a dicho proyecto que continúa vigente.

La presente tesis doctoral sigue la modalidad de compendio de publicaciones. Dichas publicaciones van encaminadas a plantear soluciones y análisis diversos que respondan a las preguntas de investigación planteadas. Las publicaciones aportadas son las siguientes:

- Aportación 1. Robinson-Garcia, N.; Torres-Salinas, D.; Delgado López-Cózar, E.; Herrera, F. An insight on the importance of national university rankings in an international context: The case of the I-UGR Rankings of Spanish universities. *Scientometrics*. Online first. doi:10.1007/s11192-014-1263-1

- Aportación 2. Robinson-Garcia, N.; Calero-Medina, C. What do rankings by fields rank? Exploring discrepancies between the organizational structure of universities and bibliometric classifications. *Scientometrics, 98*(3), 1955-1970. doi: 10.1007/s11192-013-1157-7.

- Aportación 3. Torres-Salinas, D.; Robinson-Garcia, N.; Jiménez-Contreras, E.; Herrera, F.; Delgado López-Cózar, E. (2013). On the use of biplot analysis for multivariate bibliometric and scientific indicators. *Journal of the American Society for Information Science and Technology. 64*(7): 1468-1479. doi:10.1002/asi.22837

- Aportación 4. Garcia, J.; Rodriguez-Sanchez, R.; Fdez-Valdivia, J.; Robinson-Garcia, N.; Torres-Salinas, D. (2012). Mapping Academic Institutions According to Their Journal Publication Profile: Spanish Universities as a Case Study. *Journal of the American Society for Information Science and Technology, 63*(11): 2328-2340. doi: 10.1002/asi.22735

- Aportación 5. Robinson-Garcia, N.; Rodriguez-Sánchez, R.; García, J.A.; Torres-Salinas, D.; Fdez-Valdivia, J. Análisis de redes de las universidades españolas de acuerdo a su perfil de publicación en revistas por áreas científicas. *Revista Española de Documentación Científica, 36*(4), e027. doi:10.3989/redc.2013.4.1042.

Revisión bibliográfica

El campo de la bibliometría ha cobrado especial importancia en las últimas décadas debido a su capacidad para proporcionar indicadores que asesoren sobre la evaluación de la actividad investigadora de instituciones, personas y países. Con menos de 100 años de historia, las metodologías, técnicas e indicadores bibliométricos son cada vez más empleados de cara a la distribución de financiación para proyectos científicos o la concesión de plazas de trabajo de investigadores. Nacida a principios de los años veinte, fue en los años cincuenta cuando Derek de Solla Price, Eugene Garfield y Robert K. Merton desarrollaron las bases metodológicas y teóricas que permitirían su posterior desarrollo.

En este sentido, resulta esencial la creación del *Science Citation Index* en 1963 (Garfield, 1964), base de datos que será el principal suministro de datos bibliométricos para la evaluación de la actividad investigadora a nivel internacional. Al contar con un amplio corpus de datos bibliométricos, la comunidad científica tiene la oportunidad de analizar el uso de la cita como principal indicador para medir la influencia científica de las contribuciones. Así, se desarrollan indicadores bibliométricos de tanta trascendencia como el Factor de Impacto, se analiza la distribución que siguen las citas y cómo éstas tienden a concentrarse en un grupo reducido de trabajos. Se profundiza en cuestiones como la colaboración científica, la distribución de la producción de los investigadores y demás aspectos que, aunque originalmente pretendían profundizar en el estudio de la ciencia y cómo se general el conocimiento científico, finalmente tendrá una importante repercusión en la evaluación científica.

Dicha repercusión se pone en evidencia tras el fin de la Guerra Fría. La grave crisis económica de los años ochenta, unida a la necesidad de encontrar una utilidad civil a toda la investigación militar desarrollada anteriormente y así traducirla en competitividad económica, hacen que la universidad se convierta en un elemento central para la generación de nuevo conocimiento científico (Geisler, 1995). Así pues, los gobiernos de los distintos países empiezan a introducir ejercicios de evaluación universitaria a nivel nacional. En este sentido, cabe destacar la importancia del Reino Unido como pionero en la introducción de este tipo de evaluaciones (Moed, 2008). Estos sistemas de evaluación se caracterizan por (Hicks, 2012):

- Estar centradas principalmente en la evaluación de la actividad investigadora, ignorando otros aspectos como la calidad docente o la transferencia de conocimiento.

- El uso de una evaluación *ex post* basada en la premisa de que no hay mejor indicador del rendimiento científico futuro que analizando el pasado.

- La exclusión de indicadores de entrada, centrándose únicamente en los indicadores de salida como son las publicaciones científicas.

- Tener como objetivo la distribución de financiación y por tanto, afectan directamente al presupuesto final de las universidades.

- Considerar las características nacionales específicas del sistema sobre el que se implantan.

- Ser sistemas dinámicos que deben reformarse y modificarse para así estar siempre en consonancia con las transformaciones que sufra el sistema universitario.

- Analizar y evaluar el rendimiento científico de las universidades, ignorando otros tipos de instituciones de investigación.

Como consecuencia de la internacionalización de las universidades y la globalización del modelo de universidad centrada en la investigación, en 2003 nace el Ranking de Shanghai, que posiciona las 500 universidades más productivas del mundo de acuerdo a una serie de indicadores de rendimiento académico y científico, donde prevalecen éstos últimos (Docampo, 2013). Rápidamente empiezan a surgir nuevos rankings internacionales que emplean distintas metodologías a la hora de ordenar las universidades pero que tienen en común su respaldo metodológico en datos e indicadores bibliométricos.

Aunque surgen en su mayoría de iniciativas privadas y no tienen pretensiones evaluativas, su impacto es prácticamente inmediato, y pronto empiezan a surgir estudios que analizan el rendimiento de distintas nacionales basándose en la información que éstos ofrecen (Aghion et al., 2010; Docampo, 2011). Así pues, surgen diversas iniciativas en Europa para paliar el efecto de estos rankings, especialmente por las numerosas críticas metodológicas que reciben y por el peligro que entraña hacer política científica basándose en ellos ya que ponen en peligro la diversidad institucional (Marginson, 2008). Los países europeos, al contrario de lo que sucede en Estados Unidos, se caracterizan por tener sistemas universitarios menos preparados para enfrentarse a un entorno competitivo, tal y como promueven los rankings de universidades. Ante este nuevo escenario, la Union Europea plantea rankings y sistemas de evaluación institucional alternativos como el proyecto U-Map (http://www.u-map.eu) o el proyecto U-Multirank (http://www.u-multirank.eu).

Estos proyectos europeos se caracterizan por enfatizar la necesidad de establecer tipologías institucionales que permitan hacer comparaciones entre instituciones similares y no comparar instituciones que resultan muy heterogéneas entre sí. En este sentido, son muchos los estudios que resaltan la necesidad de elaborar herramientas complementarias a los rankings que reflejen el perfil disciplinar de las universidades (Calero et al., 2008; Bornmann et al., 2014). Esta tesis se centra en este aspecto mediante el desarrollo de metodologías para la creación de herramientas de visualización de la ciencia.

Los mapas científicos han demostrado ser herramientas muy útiles para asesorar a gestores de política científica (Noyons, 2005). Su desarrollo se enmarca dentro del campo de la bibliometría, empleando generalmente la teoría de redes para facilitar su interpretación. Estos mapas se elaboran estableciendo redes entre los objetos de análisis. Estas redes se definen mediante relaciones que pueden definirse a través de la coautoría de trabajos científicos, la referencia simultánea de dos trabajos científicos a un tercero, etc. Otras metodologías emplean el análisis de componentes principales para analizar datos multivariantes y representar las relaciones entre los objetos de estudio mediante una reducción en dos dimensiones de las variables analizadas.

Aportación 1. An insight on the importance of national university rankings in an international context: The case of the I-UGR Rankings of Spanish universities

La gran importancia que tienen actualmente los rankings internacionales en política científica hace que debamos ser conscientes de las muchas limitaciones que presentan estas herramientas. Entre otras, destacamos su incapacidad para representar de manera precisa sistemas universitarios nacionales, ya que su objetivo es exclusivamente incluir a las principales universidades a nivel mundial. Otra de sus limitaciones es su incapacidad para representar el perfil disciplinar de las universidades, al presentar normalmente una única tabla global. Aunque algunos rankings presentan una amplia cobertura y otros ya incluyen tablas por campos científicos, no existe actualmente ningún ranking internacional que cumpla estos dos requisitos. Con el objetivo de superar este obstáculo desde el punto de vista de los gestores de política científica, este trabajo pretende analizar la posibilidad de emplear rankings nacionales para complementar la información que ofrecen los internacionales. Para ello, analizamos el sistema universitario español y presentamos los Rankings I-UGR de universidades españolas por campos y disciplinas científicas. Comparamos los resultados obtenidos en dicho rankings con los que ofrecen el Ranking de Shanghai, el Ranking QS, el Ranking de Leiden y el Ranking NTU, ya que todos ellos tienen características comunes que permiten dicha comparación. Concluimos recomendando el uso de rankings nacionales para complementar a los rankings internacionales. No obstante, esto debe hacerse con cautela, ya que emplean distintas metodologías tanto para el cálculo de indicadores como para la construcción de los campos y disciplinas científicas.

Aportación 2. What do rankings by fields rank? Exploring discrepancies between the organizational structure of universities and bibliometric classifications

Normalmente, los rankings de universidades por campos científicos se basan en la actividad investigadora de las universidades. No obstante, los gestores universitarios y los consumidores de estas herramientas suelen interpretar dichos campos como un reflejo de la estructura organizacional de su propia institución. En este trabajo analizamos este problema interpretativo mediante el desarrollo de perfiles de investigación de las unidades organizacionales de dos universidades españolas: la Universidad de Granada y la Universidad Pompeu Fabra. Empleamos dos sistemas de clasificación científica, el sistema de categorías temáticas desarrollado por Thomson Scientific, que suele emplearse en los estudios bibliométricos, y las 37 disciplinas empleadas en el desarrollo de los Rankings I-UGR de universidades españolas, que se construyen a partir de la agregación de las categorías temáticas de Thomson Scientific. También describimos los problemas técnicos a los que hay que hacer frente cuando se emplea una perspectiva *top down*. Mostramos las diferencias entre distintas estructuras universitarias. Concluimos señalando que los rankings por campos deben incluir en su metodología cómo construyen los campos en los que se basan. También mostramos cómo la construcción de perfiles de investigación puede ser una solución plausible para identificar los niveles de concordancia entre las distintas unidades organizativas de una universidad y los campos científicos mostrados por los rankings.

Aportación 3. On the use of biplot analysis for multivariate bibliometric and scientific indicators.

El mapeo de la ciencia y las técnicas de visualización científica representan uno de los principales pilares sobre los que se sustenta la bibliometría. Tradicionalmente se han empleado técnicas como el escalamiento multidimensional, el análisis de componentes principales, o el análisis de correspondencia. En este trabajo presentamos la metodología de visualizaciión análisis Biplot para representar indicadores bibliométricos y de ciencia y tecnología. Un Biplot es una representación gráfica de datos multivariantes, donde se representa los elementos de la matriz de datos mediante puntos y vectores que se asocian con las filas y las columnas de la matriz. En este trabajo exploramos la posibilidad de emplear el análisis Biplot como herramienta para asesorar en política científica. Para ello, presentamos y describimos en primer lugar la metodología. En segundo lugar, analizamos las fortalezas y debilidades de esta metodología analizando tres unidades distinas: países, universidades y campos científicas. Para ello, empleamos el análisis JK-Biplot. Finalmente comparamos la representación Biplot con otras técnicas de análisis multivariante. Concluimos resaltando la utilidad de esta técnica en el cmapode la bibliometría y su potencial como herramienta de apoyo en política científica.

Aportación 4. Mapping Academic Institutions According to Their Journal Publication Profile: Spanish Universities as a Case Study

Introducimos la nueva metodología para el mapeo de instituciones académicas basada en su perfil de publicación en revistas científicas. Consideramos que las revistas en las que publican los investigadores de las instituciones académicas puede ser un buen identificador para representar las relaciones que se establecen entre instituciones y establecer comparaciones. Sin embargo, al utilizar las revistas como elemento relacional, es necesario introducir distinciones entre las mismas basándonos en su valor como descriptores institucionales. Esto nos lleva a asignar pesos a las revistas que se asocian a los identificadores de las instituciones. Partiendo de la base de que una revista en la que publican investigadores de muchas instituciones es un mal predictor de la similaridad entre dos instituciones académicas, resulta razonable asignar los pesos de acuerdo a la frecuencia con que los investigadores de diferentes instituciones publican en dicha revista. De este modo podemos aplicar análisis de cluster para establecer grupos de instituciones académicas e ilustrar dichos grupos mediante dendrogramas empleando una agrupación jerárquica de conglomerados. Para contrastar dicha metodología, utilizamos una muestra de universidades españolas como estudio de caso. Primero elaboramos mapas de la ciencia basados en la totalidad de la producción científica de dichas universidades. Luego nos centramos en dos campos científicos (Tecnologías de la Información y la Comunicación y Medicina y Farmacia). De este modo, mostramos la aplicación práctica de esta metodología, no sólo para analizar la producción científica de las universidades a nivel global, sino también centrándonos en contextos disciplinarios concretos.

Aportación 5. Análisis de redes de las universidades españolas de acuerdo a su perfil de publicación en revistas por áreas científicas

Este trabajo presenta un análisis descriptivo de las universidades españolas de acuerdo a su perfil de publicación en revistas científicas en cinco áreas de conocimiento para el periodo 2007-2011. Dos universidades tienen un perfil de publicación en revistas similar cuando publican en un alto número de revistas comunes. Siguiendo este principio es posible crear mapas de universidades que ofrezcan una visión enriquecedora del sistema universitario español. Para ello, analizamos las áreas de Ciencias Sociales, Ciencias Exactas, Ingeniería y Tecnología, Ciencias de la Vida y Ciencias de la Salud. Además, utilizamos el indicador de centralidad del análisis de redes sociales para identificar aquellas universidades que muestran un rol más destacado en cada área al tener un mayor número de conexiones directas con el resto de universidades. Finalmente, discutimos la aplicación de esta metodología en un contexto de política científica de cara a la búsqueda de colaboraciones científicas potenciales.

Referencias

Aghion, P., Dewatripoint, M., Hoxby, C., Mas-Colell, A., & Sapir, A. (2010). The governance and performance of universities: Evidences from Europe and the US. *Economic Policy, 25*(61), 7-59.

Altbach, P.G., Reisberg, L., & Rumbley, L.E. (2010). Tracking a global academic revolution. *Change: The Magazine of Higher Education, 42*(2), 30-39.

Bornmann, L., Moritz, S., Moya-Anegón, F., & Mutz, R. (2014). Ranking and mapping of universities and research-focused institutions worldwide based on highly-cited papers: A visualisation of results from multi-level models. *Online Information Review, 38*(1), 43-58.

Calero-Medina, C., López-Illescas, C., Martijn, V.S., & Moed, H.F. (2008). Important factors when interpreting bibliometric rankings of world universities: An example from oncology. *Research Evaluation, 17*(1), 71-81.

Docampo, D. (2011). On using the Shanghai ranking to assess the research performance of university systems. *Scientometrics, 86*(1), 77-92.

Docampo, D. (2013). Reproducibility of the Shanghai academic ranking of world universities results. *Scientometrics, 94*(2), 567-587.

Estermann, T., Nokkola, T., & Steinel, M. (2011). *University autonomy in Europe II. The scorecard* [Informe técnico]. European University Association.

Garfield, E. (1964). Science Citation Index: A new dimension in indexing. *Science, 144*(3619), 649-654.

Geisler, E. (1995). When whales cast are cast ashore: The conversion to relevancy of American universities and basic science. *IEEE Transactions on Engineering Management, 42*(1), 3-8.

Hakanen, E.A. (2002). Lists as a social grid: Ratings and rankings in everyday life. *Social Semiotics, 12*(3), 245-254.

Hazelkorn, E. (2011). *Rankings and the reshaping of higher education. The battle for world-class excellence.* Houndmills, UK: Palgrave MacMillan.

Hicks, D. (2012). Performance-based university research funding systems. *Research Policy, 41*(2), 251-261.

Ishikawa, M. (2009). University rankings, global models, and emerging hegemony. Critical analysis from Japan. *Journal of Studies in International Education, 13* (2), 159-173.

Marginson, S. (2008). Global field and global imagining: Bourdieu and worldwide higher education. *British Journal of Sociology and Eduction, 29*(3), 303-315.

Moed, H.F. (2008). UK Research Assessment Exercises: Informed judgments on research quality or quantity? *Scientometrics, 74*(1), 153-161.

Noyons, E.C.M. (2005). Science maps within a science policy context. In: H.F. Moed, W. Glänzel & U. Schmoch (eds). *Handbook of quantitative science and technology research.* Dordrecht: Kluwer Academic Publishers, pp. 237-255.

Sauder, M., & Espeland, W.N. (2009). The discipline of rankings: Tight coupling and organizational change. *American Sociological Review, 74*(1), 63-82

Torres-Salinas, D., Delgado López-Cózar, E., Moreno-Torres, J.G., Herrera, F. (2011). Rankings ISI de las universidades españolas según campos científicos: Descripción y resultados. *El profesional de la información, 20*(1), 111-122.

Van Raan, A.F.J. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics, 62*(1), 133-143.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E.C.M., Tijssen, R.J.W., van Eck, N.J., van Leeuwen, T.N., van Raan, A.F.J., Visser, M.S., & Wouters, P. (2012). The Leiden ranking 2011-2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology, 63*(12), 2419-2432.

Table of contents

PART 1. Framework, background and research questions

1. Introduction

1.1. The phenomenon of university rankings

The obsession with rankings and league tables seems to be a widely spread phenomenon that goes beyond higher education or research policy. Humans are inevitably attracted to rankings. The act of comparing is a human notion intrinsictly related with our nature. The music industry as well as the publishing industry have long acknowledged the power of rankings and top chart lists as marketing tools to capture consumers' attention. The popular Billboard music charts released their first Hot 100 list in 1958. The first bestseller list goes as far back as 1895 ranking books according to sales. Rankings and league tables allow consumers to easily visualize the performance of a particular unit of analysis and compare it with other similar such units. On the one hand, they introduce a quantitative perspective to an issue such as reading preferences: ranked lists introduce the notion of *what is best* or *which option is better* by sorting them. On the other hand, they seem to reduce the complexity of the objects that are compared, hence facilitating the consumers' choice. The novelist Nick Hornby (1996) succinctly portrays the obsession with rankings and how they change our way of understanding our surroundings in his novel 'High Fidelity' when describing one of the main characters:

> *But his conversation is simply enumeration: if he has seen a good film, he will not describe the plot, or how it made him feel, but where it ranks in his best-of-year list, his best-of-all-time list, his best-of-decade list – he thinks and talks in tens and fives, and as a consequence, Dick and I do too.*

Here two relevant issues are highlighted. First, rankings divert our attention away from the object we are analyzing; instead of trying to understand it or learn more about it, we focus on where it stands in comparison with the rest. The notion of finding out where something or someone *stands* or *belongs* is emphasized by Hakanen (2002) who analyzes the social role lists play as signifiers of our perception. He states that rankings have become part of our social code, to help us understand our context and relate to others. Second, even if our first reaction is to disregard rankings, after being exposed to them we quickly internalize their existence and include them in our own way of perceiving things (*as a consequence, Dick and I do too*). This is confirmed by Sauder and Espeland (2009) who analyze the reaction of managers of Law

Schools to rankings and describe how they have reacted in order to control and buffer the effects of rankings on their institutional image. These authors describe how research managers first ignored rankings only to later introduce strategies to mitigate their effect.

In this sense, rankings have had a deep impact on university management and the establishment of institutional and national strategic planning (Hazelkorn, 2011). Although they are external tools, rankings have become the yardstick of the international higher education landscape. They greatly affect peer and student perceptions of the reputation of a given institution (Bastedo & Bowman, 2010), which in turn affects its potential for attracting talent and funds. This translates into what Hazelkorn (2011) defines as 'policy-making by numbers', that is, considering the methodology and indicators employed by rankings when establishing strategic planning. This is supported by Adams and Baker (2010) who conducted a global survey of university managers and stakeholders, reporting that 70% of the respondents considered that rankings 'make institutions focus on numerical comparisons rather than on educating students'. Although malpractices have been reported (Sauder & Espeland, 2009), the first step university managers have taken is to influence those indicators which can be easily improved. For instance, some institutions have developed policies directed at instructing their academic staff to include their affiliation information correctly in their published papers in order to overcome data retrieval issues (Hazelkorn, 2011).

Thomson Reuters and Elsevier have become strategic players on providing bibliometric data for the analysis of institutional performance through their multidisciplinary databases Web of Science and Scopus. They have both established partnerships with the main international university rankings as data providers and developed their own tools (Incites and Scival respectively) which they sell to universities to monitor and assess their research performance. In the same vein, many research assessment consultancies have emerged offering professional bibliometric assistance in order to ease decision-making. This has had a positive effect, urging universities to be more transparent and to care about the quality of their institutional data. Also, some countries have been forced to reconsider their national university system and introduce changes in their funding and staffing systems that will allow them to compete in an international context.

But the simple and simplistic interpretation derived from rankings when using them to understand the composition of higher education has also had its pitfalls. Most of these have to do with limitations on bibliometric issues. As they try to reduce to a number the overall research performance of an institution they are forced to surpass many limitations which are well acknowledged by bibliometricians (van Raan, 2005). As they rely on international databases which are biased towards the English language, they favor universities from English-speaking countries (Altbach, Reisberg & Rumbley, 2010). Consequently, this affects the fields of the Social Sciences and Humanities. Universities that specialize in these fields do not usually perform well in rankings as these rely on citation indices which are of limited use when evaluating research performance in the Social Sciences and Humanities (Ishikawa, 2009).

1.2. Purpose, objectives and structure of the thesis

Academic institutions and universities constitute one of the main units of analysis in research evaluation. There is a strong belief among countries that industrial cooperation with academia enhances economic competitiveness (Geisler, 1995). This assumption is based on three factors. First, the post cold-war era ends with military research-funding which must be transferred to other players. Universities are seen as the perfect candidate to convert military research-funding to civilian uses. Second, the economic crisis of the 1980s leads to New Public Management, and to demand a return on investment to universities. In order for universities to adapt to this new context, there is an emphasis on converting university research into 'a portfolio more relevant and applicable to industrial needs' (Geisler, 1995). Finally, the rise of the global economy (Castells, 1996), where a knowledge-based economy investing in intangibles and human capital emphasizes the importance of developing world-class universities. This new scenario in higher education is characterized by 1) cross-border flows which allow international collaboration and the migration of talent, and 2) patterns of differences which channel and limit global flows by the stratification of national university systems (Marginson, 2008).

This context leads to the introduction of national research performance-based funding systems and the success of international world-class university rankings (Hazelkorn, 2011). The former derives from governmental initiatives directed at encouraging competition and allocating funds. International rankings have no direct purpose or aim, but have stirred the higher education scenario. Since their emergence at the beginning of the 21st century, their success has been immediate. This success lies in the lack of tools to analyze and offer an overall view of the state of higher education worldwide. The fascination exerted by top lists and league tables as well as the demand of savvy students to learn which are the best institutions to pursue their studies have played a crucial part in this success. However, the most noteworthy aspect of the impact of international rankings is their current relevance and presence in the agendas of research managers, politicians and policy-makers (Hazelkorn, 2011).

International university rankings mainly reflect research performance, neglecting other activities such as education or knowledge transfer, and are afflicted by many shortcomings regarding data retrieval and methodological choices (van Raan, 2005). The bibliometric community has repeatedly shown its concerns about their use in research policy (i.e., Calero-Medina et al., 2008; Moed et al., 2011; Zitt & Filliatreau, 2007), but it has not been until recently that ranking producers have adopted a more reflexive attitude, conceding the limitations of their tools and suggesting further improvements. The work presented here aims at developing methodologies and tools that allow research policy-makers to analyze the disciplinary focus of universities and detect similarities between academic institutions which will allow them to better understand the information provided by rankings. In this sense, one of the most recent improvements in university rankings has been the inclusion of league tables by fields (Robinson-García et al., 2014). The importance of considering the disciplinary profile of universities is crucial to understand their positioning in rankings and the relevance of the comparisons made between universities.

The motivation for this approach is derived from my involvement as a member of the team that develops the I-UGR Rankings of Spanish universities by scientific fields and disciplines[1]. Despite having a national scope, the I-UGR rankings were developed because of the existence of international rankings. The Spanish system suffers from low levels of autonomy (Estermann, Nokkola & Steinel, 2011) and a high dependence on public funding which historically has been based on the number of students. The lack of competition between universities has led to a system where there are fewer differences between universities than within them. However, this country has not been relieved from the impact of international rankings. This puts universities in a difficult position: unable to compete, they are under the spotlight of politicians who press them to improve their visibility in rankings. Their presence in these

[1] http://rankinguniversidades.es

rankings is very limited, which leaves university managers without the tools to learn where Spanish universities stand.

The I-UGR Rankings aimed to fill this gap and, at the same time, provide a comprehensible tool that would consider disciplinary differences between universities, offering league tables by fields and disciplines (Torres-Salinas et al., 2011). My contribution to this project is reflected in this thesis. It begins from the foundation established by my colleagues and goes on to further study the understanding of university rankings and the development of complementary tools that improve assessment with regard to disciplinary differences between universities.

Part 1 of the thesis presents the theoretical framework and background and states the research questions that will be addressed here. It is structured in three chapters. Chapter 1 serves as an introduction to the topics that will later be developed. Chapter 2 sets the framework and background of the work presented. It is structured in four sections, each of them presenting the different topics that will be discussed later. Section 2.1 is a general overview of the field of bibliometrics and its relevance in research policy. Sections 2.2 and 2.3 refer to institutional evaluation, following a chronological order which starts with the description of national research performance-based funding systems and ends with the appearance of international rankings and a description of the main exponents. Section 2.4 offers a practical guide to science mapping and classification which will be the main techniques employed in this thesis. In chapter 3 we first resume the main conclusions reached in part 1. Then we briefly describe the Spanish university system and we pose the research questions of this thesis.

Part 2 constitutes the main component. It is formed of five chapters, each of them being published papers that contribute to fulfill the objectives presented. Chapters 4 and 5 directly address the I-UGR Rankings, while the other three make use of data derived from the rankings. Chapter 4 aims to justify the role of a national ranking within an international framework. Chapter 5 analyzes some interpretation issues that have been observed when using rankings by fields in order to show disciplinary differences among universities. In order to surpass these issues, we take a different approach by making use of science mapping to develop methodologies that allow us to identify different university profiles. In chapter 6 we present biplot analysis as a plausible solution when using multivariate data and we offer three different case studies in which this technique may be of use. The first one analyzes the scientific efforts of countries by combining bibliometric data with Science and Technology indicators. The second one explores the profile of the top 25 universities in the THE Ranking according to four of the variables it records. The third focuses on a single university and displays its performance in different fields according to indicators of a different nature. The discussion of this third case study is supported by the positioning of the university in the I-UGR Rankings, which prove to be a reliable source to complement the interpretation of the results.

In chapter 7 we present the *journal publication profile* methodology which identifies similar universities based on their publishing in common journals. This methodology seems promising as it not only relates similar research topics, but it is not necessarily based on a classification system (although, it is advisable to focus on areas in order to better analyze disciplinary differences). It also reflects collaboration between universities as well as geographical proximity. We further analyze the use of this methodology in chapter 8 and establish four classes of Spanish universities according to their journal publication profile. Finally, in part 3 we present our main conclusions and discuss relevant issues which remain unanswered suggesting further research.

References

Adams, J., & Baker, K. (2010). *Global opinion survey – New outlooks on institutional profiles.* Thomson Reuters, accessible at http://ip-science.thomsonreuters.com/m/pdfs/Global_Opinion_Survey.pdf

Altbach, P.G., Reisberg, L., & Rumbley, L.E. (2010). Tracking a global academic revolution. *Change: The Magazine of Higher Education, 42*(2), 30-39.

Bastedo, M.N., & Bowman, N.A. (2010). U.S. News & World Report College Rankings: Modeling institutional effects on organizational reputation. *American Journal of Education, 116*(2), 163-183.

Calero-Medina, C., Lopez-Illescas, C., Visser, M.S., & Moed, H.F. (2008). Important factors when interpreting bibliometric rankings of world universities: An example from oncology. *Research Evaluation, 17*(1), 71-81.

Castells, M. (1996). *The rise of the network society.* Blackwell Publishers.

Estermann, T., Nokkala, T., & Steinel, M. (2011). *University autonomy in Europe II. The scorecard* [Technical report]. European University Association

Geisler, E. (1995). When whales are cast ashore: The conversion to relevancy of American universities and basic science. *IEEE Transactions on Engineering Management, 42*(1), 3-8.

Hakanen, E.A. (2002). Lists as a social grid: Ratings and rankings in everyday life. *Social Semiotics, 12*(3), 245-254.

Hazelkorn, E. (2011). *Rankings and the reshaping of higher education. The battle for world-class excellence.* Houndmills, UK: Palgrave MacMillan.

Hornby, N. (1996). *High fidelity.* Penguin books.

Ishikawa, M. (2009). University rankings, global models, and emerging hegemony: Critical analysis from Japan. *Journal of Studies in International Education, 13*(2), 159-173.

Marginson, S. (2008). Global field and global imagining: Bourdieu and worldwide higher education. *British Journal of Sociology and Education, 29*(3), 303-315.

Moed, H.F., Moya-Anegón, F., López-Illescas, C., & Visser, M.S. (2011). Is concentration of university research associated with better performance? *Journal of Informetrics, 5*(4), 649-658.

Robinson-García, N., Torres-Salinas, D., Delgado López-Cózar, E., & Herrera, F. (2014). An insight into the importance of national university rankings in an international context: The case of the I-UGR Rankings of Spanish universities. *Scientometrics,* doi:10.1007/s11192-014-1263-1.

Sauder, M., & Espeland, W.N. (2009). The discipline of rankings: Tight coupling and organizational change. *American Sociological Review, 74*(1), 63-82.

Torres-Salinas, D., Delgado López-Cózar, E., Moreno-Torres, J.G., & Herrera, F. (2011). Rankings ISI de las universidades españolas según campos científicos: Descripción y resultados. *El profesional de la información, 20*(1), 111-122.

Van Raan, A.F.J. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics, 62*(1), 133-143.

Zitt, M., & Filliatreau, G. (2007). Big is (made) beautiful. Some comments about the Shanghai ranking of world-class universities. In: J. Sadlak & N.C. Liu (eds). *The World-Class University and Ranking: Aiming beyond status.* UNESCO-CEPES/ Cluj University Press; Shanghai: Institute of Higher Education, Shanghai Jiao Tong University. pp. 141-160.

2. Literature review

In this chapter we analyze the role of bibliometric methodologies for research evaluation and research policy making. The field of bibliometrics is relatively new with less than 100 years of history. However, it has played an increasingly important role on the allocation of research funds and the structuring of academic institutions. Here we present a state-of-the-art of the main developments upon the research presented in this thesis builds on. For this, we first review the development of the field, since Lotka's and Gross and Gross contributions in the early 1920s to the important contributions of De Solla Price, Garfield and Merton who provided the theoretical and practical basis for the later development of the field. We mention alternative *proxies* to citation data as measures of research impact and quality and discuss the main bibliometric indicators and their application to the different levels of analysis (macro, meso and micro).

Sections 2.2 and 2.3 deal with the tools employed by the state and market to regulate and assess on the performance of universities. They also discuss the consequences and changes the globalization of higher education has had on research evaluation. We present the three missions of universities (teaching, research and innovation) focusing on the second one which is the one under study in this thesis. Then we deal with the different types of national university systems in order to review the national performance-based research funding systems introduced in many countries and reflect on their effect on universities. Then we introduce international university rankings as part of the effect of globalization, internationalization, diversification of funds and masssification of education. We revise their origins and review the methodologies employed by the main rankings. Finally we describe the response of different trans-national organizations to minimize the ill-effects of rankings in institutional diversity and the concerns shown from the research community towards the threats rankings expose to the diversification of the disciplinary profile of universities.

Finally, section 2.4 addresses methodological issues towards the development of classification and visualization techniques for assessing research policy makers. We provide a brief historical review of the role played by science mapping within bibliometrics. We offer a short how-to guide on the process flow for ellaborating maps of science. Then we analyze recent attempts to classify and profile academic institutions focusing on the methodologies employed.

2.1. The use of bibliometrics as an evaluation tool in research policy

Since the 1970s, the development of bibliometric methodologies and indicators to assess the science policy system has excited much interest (Delanghe, Sloan & Muldur, 2010). The need to account for the investment made by governments and other funding bodies has in many ways stimulated the development of the field. The ever-tightening grip of policy-making on academic life has created an internal process that distributes funding which is similar to that impused on academia from outside of the universities. This interest can be seen on the reports commissioned by the US National Science Board in the 1970s (Elkana et al., 1978; National Science Board, 1973) or, in the 1980s, the collaboration between the European Commission and the Centre for Science and Technology Studies (CWTS) in Leiden (The Netherlands) (de Bruin, Straathof & Moed, 1992; Moed et al., 1992). However, the last decades have witnessed an abuse and misuse of the application of these methods of evaluation, raising concerns about their capacity as suitable means of assessing research activity (Glänzel & Schoepflin, 1994).

In spite of criticisms, the evolution of bibliometrics, conceived of as the analysis of bibliographic information (Borgman & Furner, 2002; van Leeuwen, 2005), is intrinsically related with the evolution of scholarly communication itself and of how researchers relate to each other. Hence, and despite possible flaws, the field of bibliometrics is not primarily directed at the evaluation of research but is multidisciplinary in nature and encompasses different paradigms (Moed, 2005) aimed at quantitatively defining the relations between scientists and the evolution of scientific knowledge. In order to use and correctly interpret the methodologies and indicators developed in the field, we need a deep understanding of these paradigms, contextualizing their meaning and reach as research policy tools.

In this section we will try to achieve this by briefly reviewing the development of bibliometrics and its progressive focus on research policy. To do so, we will first define the scientific article as the main channel of scientific communication and the citation as foundation of citation analysis. We will present a brief overview of the development of the field of bibliometrics and elaborate on the efforts made by the community to establish a unified citation theory. Then, we will describe the main bibliometric indicators used in evaluation and the different levels of analysis. Finally, we will focus on the use of bibliometrics for research evaluation at a meso-level in order to relate this with the evaluation of national systems, which we refer to in Section 2.2.

The citation as a proxy for research impact

Rather than the citation itself, the main pillar of the discipline is the scientific article. This is defined as the first acceptable, publicly-available manuscript containing sufficient information to make it the object of evaluation (peer review), to show reproducible results, and to evaluate the intellectual processes undertaken during the research study in order to justify the conclusions reached (Day, 1998). Its purpose is threefold: 1) it serves as the main communication vehicle of scientific discovery, 2) it ensures quality control through the peer review system creating a valid, trustworthy corpus of scientific knowledge on which scientists can build (Bornmann, 2011), and 3) it distributes credit and recognition through authorship and citations. Although under perpetual scrutiny, the fact is that this threefold purpose is acknowledged and accepted by bibliometricians and by the scientific community as a whole, which despite its criticisms actually follows and abides by it (Harley, 2013).

Unlike other document types, the journal article is characterized by a particular structure and writing style, and by continuous need to support any argument by referencing previous work that demonstrates the facts referred to. This way of presenting research results aims not only to satisfy the information readers' need to contrast the conclusions reached but also to convince them of the arguments presented. Here, citations play a pivotal role as authors use them to support their arguments and, at the same time, to pay tribute to and acknowledge the contribution of their peers. Hence, the field of bibliometrics relies on the assumption that citations serve to distribute credit and recognition among scientists. The first studies analyzing the use of publication and citation data hearken back to the 1920s and the works of Lotka and Gross & Gross. While the former analyzed the frequency distribution of scientific productivity, setting forth his famous law (Lotka, 1926), the latter suggested the statistical use of references in order to determine the importance of contributions (Gross & Gross, 1927).

However, it was not until the 1950s when Derek J. de Solla Price, Robert K. Merton and Eugene Garfield developed the theoretical aspects of the field. Derek J. de Solla Price suggested the potential of publication analysis to measure the growth of science (de Solla Price, 1951) and the use of citations to establish networks and relations between papers (de Solla Price, 1965). On the other hand, Robert K. Merton offered, from a sociological viewpoint, a theoretical framework which considered the role of citations as a means to recognize the work of peers and the cumulative advantage of prestigious researchers on what he named the *Mathew effect* (Merton, 1973). He defined the *scientific ethos* which characterizes the motivation of scientists and science itself: universalism, organized skepticism, communalism, ethically neutral, and disinterest. These premises allowed the acceptance of citations as a proxy for the recognition and scholarly impact of peers' contributions to the development of scientific knowledge.

This view was soon challenged not only by Kuhn's paradigm (Kuhn, 1996) but also by other sociologists who considered that the *scientific ethos* presented by Merton had little to do with the actual motives of scientists. Indeed, they challenged Merton's statement that research was actually driven by individualism, organized dogmatism, secrecy, and self-interest (see e.g., Rothman, 1972). This view was further formulated by Bruno Latour in what is known as the 'social constructivist' or 'Latourian' view (Lunberg, 2006) and further stressed in the works of MacRoberts & MacRoberts (1986;1987;1996). Here it is suggested that citations and the scientific method as a whole, constitute some kind of rhetorical exercise in which researchers do not simply present their results but try to persuade their readers. From this perspective, research is a confluence of different lines of thought. This debate remains open and serious evidence that questions the citation and publication practices have been presented by Sokal (1996), and more recently by Delgado López-Cózar, Robinson-Garcia and Torres-Salinas (2014).

However, Eugene Garfield made the most crucial contribution to the development of the field (Cronin, 1984) when he created the *Science Citation Index* in 1963 (Garfield, 1964). Following the example set by the citation index for the field of Law developed by Shepard in the late 1870s, it gave future bibliometricians a basic resource from which to explore the potential of bibliometric data. As Wouters (1999) states in his review of the creation of the *Science Citation Index*:

> *It promised to make an old dream come true: the application of "the scientific method" to science itself, an idea central to the science of science.*

It also implicitly expanded the normative view of science formulated by Merton, putting citations under the spotlight of sociologists and bibliometricians, and of the whole scientific community too. Also, it led to a closer focus in its use evaluating research performance, leaving aside the study of science.

Bibliometric indicators for evaluation purposes - Beyond the Impact Factor

Although the field of bibliometrics encompasses much more, there is no doubt that most of the research community has focused its attention on the use of bibliometric indicators and methodologies for evaluation purposes. The interest of agencies such as the US National Science Foundation or the European Commission in bibliometric indicators (Delanghe et al., 2010) along with contributions such as Moed, Glänzel and Schmoch (2004), Moed (2005) or Vinkler (2010) have enabled the subfield of *Evaluative Bibliometrics* to develop (Narin, 1976). Also, the launch of the journal *Scientometrics* in 1979 greatly contributed to the creation of a scientific corpus devoted to the development of the field. Traditionally, bibliometric indicators have been built on two measures: publications and citations. However, Science & Technology indicators have also been used and combined with them when developing an evaluative framework for academic institutions (Bordons et al., 2010; Sanz-Casado et al., 2013; van Vught & Ziegele, 2011). Due to the constant scrutiny bibliometric indicators are subjected to, other measures and methodological approaches have been explored as substitutes or, at least, as supplementary to citations and publications. These have been possible thanks to the development of technology and the capacity to process large amounts of data. Without being exhaustive, following are some of the most important contributions:

- **Webometrics**. First formulated by Almind and Ingwersen (1997), the field of 'Webometrics' has now become a research front of its own. Based on the premise that the WWW offers the largest repository of documents in the world, it tracks links in order to establish relations between academic-related documents in the same way that citations track relations between research papers (Thelwall, 2008). Links are defined in various ways such as weblinks or web citation, for instance.

- **Usage indicators**. Bollen and colleagues (Bollen, Rodriguez & Van de Sompel, 2007; Bollen et al., 2005; Bollen & Van de Sompel, 2006) were the first using downloads as usage indicators. The term usage indicators refer to data on user access to scientific literature through journal platforms. The underlying idea is that usage indicators show an enhanced scenario of research communication by comparison to that offered by citation data which relies solely on the act of publication and neglects research activities that are not associated with publishing (Kurtz & Bollen, 2010).

- **Acknowledgments**. Cronin and Weaver (1995) coined the term *Influmetrics* to refer to the influence exerted by other scholars and reflected in one's own publications through the acknowledgments' section in scientific papers. They consider that this section reflects a kind of sub-authorship of organizations and individuals who have played a secondary role in the development of the research study. Hence, they suggest analyzing this field in order to close what they term as the *reward triangle*, made up of authorship, citations and acknowledgments. Costas and van Leeuwen (2012) further analyzed this field in order to study two types of acknowledgment: financial support and peer interactive communication. They concluded that a deeper understanding of the so-called *reward triangle* would benefit our understanding of scholarly recognition and widen the theoretical framework of bibliometrics.

- **Altmetrics**. The most recent buzzword in the field, altmetrics has raised expectations as to its potential. Intended to provide 'measures of scholarly impact drawn from Web 2.0 data' (Priem & Hemminger, 2010), altmetrics still have serious shortcomings which should be overcome before they can be considered supplements or complements to traditional bibliometric indicators (Wouters & Costas, 2012). Among other limitations, we would highlight their evanescent nature, normalization problems, difficulties in aggregating data from different sources (Torres, Cabezas & Jiménez, 2013) and the impossibility (at least currently) of scaling to units of analysis other than those found at the article-level.

Here, we will focus on traditional bibliometric indicators, that is, those derived from publication and citation analysis. Although the Impact Factor is the best-known indicator, the bibliometric community has developed many others as well as adapting those drawn from fields other than their own. Nonetheless, the Impact Factor has historically had the greatest impact on the research community. Despite its original focus on the analysis of scientific journals for library management (Archambault & Larivière, 2009), its use shifted to the evaluation of individuals and publications themselves as a proxy for their potential impact (Bollen et al., 2005). Alternatives for measuring journals' prestige have been proposed such as the Scimago Journal Rank indicator (González-Pereira, Guerrero-Bote & Moya-Anegón, 2010), the Source-Normalized Impact per Paper indicator (Moed, 2010) or the Eigenfactor (Bergstrom, West & Wiseman, 2008). But still, the popularity of the Impact Factor has led to a constant debate and opposition to the use of bibliometric indicators for evaluation purposes. Many researchers have been especially critical of this indicator (Simons, 2008) and the latest flare-up occurred in late Spring 2013 when journal editors and scientific associations joined forces to ask research evaluation agencies not to use bibliometric indicators in general, and the Impact Factor in particular, for research evaluation. To do this, they signed and promoted a petition known as the San Francisco Declaration on Research Assessment or DORA (Way & Ahmad, 2013).

Although most of the debate has to do with the misuse of the indicator to assess units of analysis it was not originally designed for (Garfield, 1998), the truth is that many indicators have been based on the Impact Factor. The reason for this has to do with the difficulty researchers face due to the competition to publish in journals with a high Impact Factor. It is based on the idea that the higher the Impact Factor a journal has, the more difficult it will be to publish in it. Authors will have to surpass an extremely demanding review process exerted by highly reputable colleagues. These indicators are based on the scientific community's consideration of journals. Therefore, indicators based on the Impact Factor would refer more to the visibility of research papers (papers published in journals with a high Impact Factor will potentially have a wider audience than those with a lower Impact Factor) rather than with the use the research community will actually make of them (as measured by citations).

Figure 2.1. Screenshot of the *Journal Citation Reports Social Science Edition* highlighting the bibliometric indicators displayed and a brief definition of each of these.

Total Cites	Number of citations received in a given year
Impact Factor	Average number of citations to articles published in the last two years
5-Year Impact Factor	Average number of citations to articles published in the last five years
Immediacy Index	Average number of citations to articles published in the same year of publication
Articles	Total number of papers published in a given year
Cited Half-life	Median age of articles cited in a given year
Eigenfactor Score	Number of times articles from the past five years have been cited in a given year, taking into account which journals have contributed these citations
Article Influence Score	Average influence of articles over the first five years after publication

Some of these indicators are based on the position of the journal in which a given paper is published in a ranking of journals of the same discipline. Others are based on comparisons between what is referred to as the 'observed' impact of publications (number of citations received) and their 'expected' impact (Impact Factor). The former indicators - those derived from the position of journals in rankings - are based on the journal rankings published annually by Thomson Reuters through the *Journal Citation Reports* (hereafter JCR). This product produces lists of journals indexed in the *Science Citation Index* and the *Social Science Citation Index* organized by subject categories. These categories follow the classification scheme designed by Thomson Reuters which groups journals on heuristic criteria based on citation data (Pudovkin & Garfield, 2002). The JCR offers a series of bibliometric indicators for each journal (Figure 2.1) and the Impact Factor is one of them. Although the user may choose any of the other available indicators, the Impact Factor is normally the main criterion used. Some of the indicators based on the Impact Factor are: share of papers published in journals ranked in the top 25% of their category (hereafter Q1 journals) (Bordons et al., 2005), weighting journals according to the decile they belong to (Lewison & Dawson, 1998) or using percentile rank classes (Bornmann, Leydesdorff & Mutz, 2013).

Indicators based on comparisons between 'observed' and 'expected' impact are used less frequently and are not recommended. The reason for this is that the relationship between the journal Impact Factor and the citations received by a given article is not as clear as one might think at first (Lozano, Larivière & Gingras, 2012). This is due to the characteristics of the citation distribution of papers. Citations, unlike other phenomena which follow a normal distribution, are usually concentrated on a few papers, showing a skewed distribution (Albarrán et al., 2011; Seglen, 1992). This skewness of the citation distribution of papers has important consequences not only on the adequacy of certain indicators such as the use of mean or median scores, but also on decision making. Here we summarize some of these:

1. Citations follow a similar distribution to the productivity distribution (Lotka, 1926). This not only confirms Price's impressions (de Solla Price, 1965) but means that most citations are concentrated on a core of highly cited papers. This challenges the 'Ortega Hypothesis' which states that eminent scientists rely on the works of mediocre ones, showing a stratification of science - highly cited papers will tend to cite other highly cited papers - thus validating the so-called *Mathew Effect* (Cole & Cole, 1972; Bailón-Moreno et al., 2007). In the same vein, because this skewness is linear, the distinction between mediocre and eminent researchers becomes somewhat arbitrary, as no clear line can be drawn between them.

2. As a direct consequence of the skewness of the citation distribution, one cannot avoid uncitedness as it is an inherent part of the distribution. Removing uncited papers will only lead to fewer publications as the distribution depends on the number of references papers cite and erasing a proportion of these papers would also mean removing their references (Seglen, 1992).

3. A common request in the research policy arena is to develop ways of comparing research impact between aggregates of a different nature, for instance, benchmarking individuals belonging to different fields, institutions with different disciplinary foci, etc. The main problem in doing this is that different fields have different citation and reference patterns, so making a direct comparison between, say, the number of citations received, would have serious shortcomings (Moed, 2005). Bibliometricians have developed different approaches to establish these comparisons. The first would be to use indicators based on the position of the journal in the subject category

as mentioned above. Another way would be to develop normalized citation scores based on the fact that the citation distribution for different fields follows the same model and, therefore, differences in the citation score could be considered no more than a matter of scale (Albarrán et al., 2011; Glänzel, 2011; Radicchi, Fortunato & Castellano, 2008).

Various approaches to the normalization of citation scores between fields have been analyzed. For instance, Glänzel, Schubert and Czerwon (1999) and Moed, de Bruin and van Leeuwen rely on a classification system. Moed (2010) and Zitt (2010) base their indicators on source normalization, whereas Leydesdorff and Bornmann (2011) use fractional counting. This has led to an on-going debate as to which indicator best solves the normalization issue (Leydesdorff et al., 2013; Li et al., 2013).

When choosing which indicator to use, we should consider both the purpose of the analysis and, the level of aggregation and unit of analysis. Bibliometric indicators are reasonably consistent at the macro-level (countries, fields, etc.) and the meso-level (institutions, journals, disciplines), but less consistent when focusing on the micro- or individual level (Costas & Bordons, 2005). Although research policy-makers have shown special interest in developing tools that would allow us to evaluate researchers with objective indicators, we must be especially aware that we cannot rely on single indicators to compress the research performance of individuals.

Here we must mention the success of the h-index (Hirsch, 2005) which has drawn much attention from both bibliometricians, and the research community as a whole (Rousseau, García-Zorita & Sanz-Casado, 2013). The h-index can be defined as follows: a researcher will have an h index of h when a number of h publications has received at least h citations each and the rest of their publications have each less than h+1 citations. Despite its popularity and continuous use, this indicator has been harshly criticized for its shortcomings (Costas & Bordons, 2007; Waltman & van Eck, 2012). The principle objections are: 1) it is a size-dependent indicator, 2) it is misleading when comparing researchers at different stages of their careers, and 3) it shows inconsistencies in the way researchers are ranked, functioning in a counterintuitive way.

In this regard, Jiménez-Contreras, Robinson-García and Cabezas-Clavijo (2011) recommend using citation thresholds to refer to different levels of performance and maintaining a fixed production window, which would allow fair comparisons. On the other hand, Costas, van Leeuwen and Bordons (2010) suggest classifying individuals into profiles or classes of researchers according to their productivity and impact. This would allow comparisons to be made between individuals belonging to the same class.

Here we will pay special attention to the bibliometric analysis of units at a meso-level, particularly to the analysis of academic institutions. In this regard, although bibliometric indicators seem to be more stable at this level of analysis, other issues surface which must be taken into account. The main issue here, has to do with the inability of any bibliometric indicator to encompass all aspects of institutional research performance. This is due to the influence exerted by the disciplinary focus of institutions (Visser, Calero-Medina & Moed, 2007), as a consequence of their heterogeneity (Collini, 2011). Our main objective is to study these disciplinary differences in depth, offering methodologies that may help to overcome them and studying the role university rankings play in assessing decision makers in the research policy arena.

The great interest research agencies have in evaluating research institutions and particularly, universities can be easily seen on the development of research evaluation exercises such as those conducted in the United Kingdom or in the popularity of university rankings. Bibliometric indicators tend to be seen by policy-makers as a more *objective* means of allocating funds and directing research. But they also consider universities as homogeneous systems which can be either better or worse than others, without considering differences between them. In the following two sections we will discuss the development of evaluation exercises and the development of university rankings, summarizing their methodology, impact and shortcomings.

2.2. Towards institutional research performance accountability at the national level

Despite following different paths, the phenomenon of rankings is intimately related with government interest in fiscalizing the research outputs of universities. This interest extends back to the beginning of the 20th century and has to do with two different factors. First, there has been a shift in the perception of the main mission of universities. Originally conceived of as elite centres for teaching and training, their role as generators of research and scientific knowledge has increased over the last century. Secondly, but related with the first factor, the way research is conducted has also changed. Research studies undertaken by small teams or individuals have evolved into what Price named *Big Science*, where multidisciplinary teams conduct large-scale studies using expensive resources (de Solla Price, 1963). As a consequence of these two factors, new forms of managerialism have been introduced into the higher education system (Morris, 2002). These have led to an increasing interest from funding bodies in developing external evaluation processes to allocate funds instead of the traditional peer review system (Hicks, 2009).

In this section we will analyze the changing role of universities and the growing interest on developing quantitative indicators of their performance. We will briefly describe different national university models and systems, focusing on their differences regarding the control exerted by public and private parties. Then, we will follow by descriing the evaluation schemes developed according to these different systems and models and the way they have expanded in different parts of the world. Here, we will focus on the role played by the field of bibliometrics as a plausible option for the development of 'objective' indicators of research performance. Next we will study the consequences of performance-based research funding systems (hereafter PRFS) in the higher education system, and the challenges that research policy makers, researchers and bibliometricians are currently facing. Finally we will look forward to the next section by referring to the rise of a global higher education landscape which blurs differences between national systems.

The evolving mission of universities: Theoretical university models versus actual university systems

The development of universities as educational and research institutions is a history of controversies, ideals and the constant revision of concepts and goals due to discrepancies over their main role. Originally conceived of as elite centres for highly qualified education, they have evolved into complex institutions that combine traditional roles inherited by their own history with the demands of current times that impose flexible and dynamic entities capable of generating wealth to benefit their context. The dramatic changes research has experienced during the 19th and 20th centuries have not only influenced the way we live and perceive reality, but they have also had deep consequences on how science is produced. The utility of science and the benefits from research are evident, building on an unprecedented reputation of the scientific enterprise. However, the expenses entailed are higher than ever, requiring financial support for equipment, staff and training.

Such phenomena have produced deep changes in the mission and expectations universities raise in society. Their historical role as education providers has prevailed for centuries, originally leaving research and scientific progress to other institutions such as national societies and academies (Manjarrés-Henríquez, 2009). This is the case of the United Kingdom, where the Royal Society channelled scholarly communication within the research community. Another variant is that set by the French Academy of Sciences, which adopted a key role coordinating and controlling research development in France and expanding its model to other European countries.

But it is at the end of the 19th Century, with the rise of the German universities (Barnes as cited by Torres-Salinas, 2007) when research enters higher education, establishing itself as the second mission of universities. This role has grown over time and universities have become the main motor of scientific progress. This highlights their potential as generators of wealth and has entailed their shift from the ivory towers in which they were embedded to the heart of society. This new concept of university as a research and educational institution is known as the German or Humboldt's Model. Etzkowitz (1990) refers to this change as the first academic revolution, which will lead to a need to assess and evaluate the research performance of universities. A once minimal interest in student satisfaction, it turns into a real concern to understand and monitor the behavior of universities in order to optimize and assess the production of knowledge and the human capital involved (Geuna, 1999).

Hence, we observe different variants of two higher education models expanding throughout Europe and North America. The first one would be the French or Napolean Model, where a national entity keeps its leading role as coordinator of the country's research development (the Spanish system developed in a similar fashion until the 1970s with the establishment of the Spanish National Research Council). The second one is that set by the German universities, where the government takes charge of developing a legal framework while universities are responsible for the development of scientific progress. Here, although the Max Planck Institutes play a leading role in research, they work independently to universities. This model will later lead to different variants depending on the levels of governance universities exert. It will deeply influence countries such as the United Kingdom or the United States, where the importance of universities as generators of scientific knowledge has prevailed. As a result, there is a constant tension between the State and the universities. The former, the main source of universities' funding, tries to direct them as a tool to achieve prosperity. On the other hand, the universities are immersed in an identity crisis trying to maintain their independence but also knowing their importance as key players in the economic and political development of their countries.

The second mission of universities was finally favoured over teaching as research advancement became an icon of nationals' supremacy. Here, two views have been adopted. First, the *scientific ethos* set by Merton, which defines the theoretical motivations of researchers to pursue scientific knowledge. Secondly, the idea that scientific progress will lead to technological and industrial knowledge (Bush, 1945). This latter notion has lately expanded with an on-going discussion of a *third* mission of universities (Nedeva, Barker & Osman, 2014). Here, universities are not only expected to generate knowledge and offer academic training, but also to apply and exploit this knowledge outside of academia, having a direct impact on society (Laredo, 2007). At this point, a third player emerges: the market. This new context shapes a landscape in which countries develop distinctive national university systems determined by the relations between market, state and university.

These tensions are presented through the Triple Helix theory formulated by Etzkowitz and Leydesdorff (2000). This conceives of the relation between these three entities as a network in which each actor reinforces and influences the other two, establishing an 'endless transition' by which knowledge is generated and then used as a source for production and distribution. The increasing importance of knowledge in the innovation process introduces universities into a system in which the only actors at the time were Industry and the state. This introduction involves a series of processes that transform the way production, information exchange and the use of knowledge take place. According to (Etzkowitz, 2000) these processes would be: 1) an internal transformation of each institution with regard to the way in which they interconnect, 2) a greater influence of each institution over the other two actors, 3) the emergence of new entities resulting from the institutionalization of the established networks and 4) a drift from the disinterested generation of science as conceived by Merton (1973) to a science based on goals.

As a result, since the 1980s most countries have introduced PRFSs into their system as a means of monitoring the research performance of their academic institutions, following the example set mainly by two countries: the United States and the United Kingdom. The first was a pioneer in the development of non-governmental university rankings, highlighting the publication of the 1st edition of the America's Best Colleges university ranking in 1983. The latter, in 1986, launched the first Research Assessment Exercise, an example of governmental action to monitor research performance. At the time, both countries were reaching the end of the huge economic recession that followed the end of the 'Cold War'. With the conservative parties in power, Reagan in the United States and Thatcher in the United Kingdom undertook important reforms in all sectors, including Higher Education. It was a time to reflect when public funding had to be justified by any means and return on investment became a priority. Hence, it is not surprising to see the emergence of a highly competitive environment in which universities had to prove their worth. These two events flag the beginning of two initiatives that began separately but which illustrated the response of government and market to the economic constraints, and the social pressure they would ultimately put on universities.

The introduction of national multi-university research evaluation systems

While governmental PRFSs saw an earlier expansion, especially during the last decades of the 20th Century, the globalization of Higher Education, the massification of universities and the emergence of the entrepreneurial university model led to the final outbreak of world-class university rankings at the beginning of the 21st Century. For the sake of clarity, here we will follow the chronological order in which each evaluation system expanded, focusing now on the former and dedicating Section 2.3 to the latter.

According to Abramo, Cicero and D'Angelo (2011), national agencies and governments implement such evaluative measures by two possible routes which are not exclusive and may co-exist. The first one has to do with the introduction of national R&D programs for evaluating research project grants. An example of such a system can be observed in Spain where, since 1988, the government launches a National R&D Plan every three years for allocating funds for research projects through a peer review process (Cabezas-Clavijo et al., 2013; Jiménez-Contreras et al., 2011), following the same pattern that the National Science Foundation in the United States. The second and most widely extended channel takes place through the implementation of PRFSs. Most of these are focused at the institutional level and take place periodically. The first country to introduce these systems was the United Kingdom as we will later discuss. It is normally implemented in national university systems where there is internal competition within institutions to attract talent and funding, leading to a stratified system. Here, agencies assess institutional performance and universities are the ones in charge of the evaluation of their personnel. However, there is another variant which focuses on the evaluation of individuals at a national level, adapting the British model to non-competitive higher education systems where researchers may be widely dispersed throughout universities (Abramo et al., 2011; Jiménez-Contreras, Moya-Anegón, Delgado López-Cózar, 2003). Here we will focus on PRFSs centered on the institutional level. In general terms, performance-based research funding systems can be characterized as follows (Hicks, 2012):

- They are focused mainly (and in most cases solely) on research performance, ignoring other facets of universities such as teaching or innovation.
- They use an *ex post* evaluation, meaning that they are based on the principle that past performance is the best predictor of future production.
- They do not include input in their evaluations but only the research output.
- Their aim is to distribute funds. Hence the final budget of universities will depend to a certain extent on the results of this evaluation.
- They are national, highly focused on the particularities and characteristics of their university system.
- They are dynamic systems which are constantly updated and redesigned.
- They analyze and evaluate the research performance of academic institutions, ignoring other research centers.

The first system implemented with these characteristics and probably the one which has gained most attention is that undertaken in the United Kingdom: the Research Assessment Exercise[1] (RAE, 1989-2008), now converted into the Research Excellence Framework[2] (REF, expected implementation 2014). The RAE has been subjected to many studies describing its methodology and evolution (e.g., Barker, 2007), analyzing its advantages and disadvantages (e.g., Clerides, Pashardes & Polycarpou, 2011; Geuna & Martin, 2003) or discussing its effect on researchers and universities (e.g., Elton, 2000; Moed, 2008; Yokoyama, 2006)). This system affects both the budgets of universities and their image when attracting top researchers, as high ratings may promote departments, serving as a 'marketing' tool (Clerides et al., 2011). It rates research units from one to five, five being the highest score, and employs what

[1] http://www.rae.ac.uk/

[2] http://www.ref.ac.uk/

is known as an 'informed peer review' approach (van Raan, 1996). That is, the process is based on peers' judgments which use bibliometric indicators as a basis for their decision. Hence, research units are classified into 'Units of Assessment' (UoA), each of them having their own panels formed by 10 to 15 experts. The members of each panel are then asked to nominate reviewers from outside the agency. The idea is that each active research staff member of each unit must submit up to four of their 'best' publications to be reviewed.

Indeed, similar experiences have been reporter in countries such as Australia, The Netherlands, Norway or Denmark for instance (Auranen & Nieminen, 2010; Geuna & Martin, 2003; Hicks, 2012). The role played by bibliometrics when implementing them has been undisputable (Hicks, 2009), leaving traditional peer review in the background. However, there seem to be mixed feelings about results (Anon, 2010) and contradictory views as to the effect these evaluation systems have had on individual researchers' performance (Abramo et al., 2011). The main problems have to do with the misuse of bibliometric methodologies, a perceived mistrust of peer review among policy managers and a simplistic, deterministic view of the effect economic incentives have on research performance and on the transfer from basic research to innovation. The bibliometric community has often warned of the limitations of bibliometric indicators, indicating the need for good practices, quality data and to combine these methodologies with peer review (i.e., Moed, 2007; van Leeuwen, 2007). However, these have mostly been ignored, turning the use of bibliometrics as a research policy tool into a controversial issue.

The effects and prospects of national performance-based research funding systems

The relevance given to bibliometrics by research policy-makers has increased over time. This has had important consequences not only on the structure of national systems but on the behavior of researchers. In general, the use of bibliometric indicators leaves a sour taste in mouth of the research community, who are not convinced by seeing their performance reduced to numbers (Abbott et al., 2010). In contrast, research managers are relatively satisfied with their use, as they provide 'objective' measures which seem easy to interpret. This allows them to partially remove or at least combine them with peer review, which is seen with apprehension (Schneider, 2009). This replacement can be seen in the evolution of the criteria used to evaluate research performance in many countries (Hicks, 2009). As far as bibliometricians are concerned, they believe in the potential of bibliometrics as a research policy tool but consider that PRFSs have both advantages and disadvantages.

Geuna and Martin (2003) comment on some of these limitations, such as their bias towards the basic sciences, as well as their neglect of interdisciplinary research. Most of the criticisms have been directed to the British RAE. However, others have questioned this negative opinion offering a more reasoned analysis of performance. For instance, Clerides and colleagues (2011) insist that researchers' perceptions are in close harmony with the ratings. Moed (2008) concludes that the RAE is producing an increase in output, especially in journals with higher visibility (higher Impact Factor) which may lead to indirect growth in terms of citations. However he has not found a direct correlation between.

Indeed, the criteria employed will be a crucial factor in determining the success or failure of the PRFS. A good example of failure is that reported in Australia, where priority was given to publication counting rather than other measures intended to quantify 'impact' as a proxy for research quality (Butler, 2003). Another important issue has to do with determining which indicators best identify impact. As discussed in Section 2.1, two basic approaches can be taken. The first puts an emphasis on citations of individual papers as these recognize the value a given publication has had for other colleagues. The second approach puts an emphasis on the Impact Factor of the journal in which it is published, considering that highly reputed journals will receive more manuscripts, employ prestigious and highly experienced reviewers and, hence, publish high quality research. An example of such approach is the one undertaken by the Norwegian PRFS (Schneider, 2009). Here, the criticism would go to the relevance given to journal rankings, which are not alleged to be good 'proxies' of the future impact of a particular publication. However, the argument given has to do with the instability of citation analysis at the micro level and the need to introduce time frames.

All in all, in general terms, PRFSs seem to be beneficial at first until the costs to maintain them surpass their benefits (Geuna & Martin, 2003). The implementation of these systems is extremely expensive, especially when they include peer review. In the case of the United Kingdom, Martin (2011) considers that the RAE lost purpose and meaning right after the third evaluation took place. PRFSs have an immediate effect on the budgets of universities affecting researchers' resources and scientific careers. This is why they feel very strongly about them, involving themselves in heated debates on the advantages and disadvantages these systems offer. In principle, these exercises should be perceived positively, as they are based on meritocratic criteria. Also, they represent a good opportunity to link research with policy, which makes it easier to argue in favor of the need for research investment when facing funding bodies. But in order to have an efficient system, it must be relatively cheap, transparent and constantly evolving to reflect the changing needs of national university systems. So far, performance-based research systems have benefited the countries that have implemented them, increasing their research output and visibility (Jiménez-Contreras et al., 2003; Martin, 2011; Moed, 2008) but there is a perceived urge to reformulate them in order to keep them useful.

The main problem we find here has to do with the purpose of these systems. There is a clear need to evaluate and assess the performance of universities. But there will always be a critical mass against their use due to the inherent difficulties of evaluating as complex an institution as the university. In this sense, Geuna and Martin (2003) identified the following shortcomings of PRFSs: 1) the economic and time-consuming costs necessary to obtain reliable data, 2) the application of a common evaluation criteria shapes universities homogeneously threatening institutional diversity, and 3) as PRFSs are focused mainly on research output they may affect the teaching performance of universities.

From a bibliometric point of view, the configuration of these systems should also be reconsidered. Although they have succeeded in their original goals, PRFSs must now move forward into adjusting to the current needs of national governments and university systems. In this sense, some practices should be discarded such as the reliance on publication counting and journal impact factors (Moed, 2007). Also, a focus on the evaluation of research groups rather than individuals or institutions would be desirable, as it is the most basic research unit in which citation analysis is meaningful (Moed, 2007). Along similar lines, it is advisable to combine multiple bibliometric indicators rather than relying on a single one, 'opening up' the research evaluation process and adapting it to specific strategic goals (Rafols et al., 2012). In this sense, Martin (2011) suggests looking at the third mission of universities by using proxies such as the number of spin-offs a university produces. The other alternative would be to take a more groundbreaking approach and question the need of government to distribute such an important budget among universities, urging them to follow a market-oriented funding model as occurs in the United States.

2.3. Universities going global: The rise of international university rankings

Although they are quite common in bibliometric studies, university rankings have not considered the use of bibliometric indicators until recently. Moreover, despite their success in the United States, they have had little presence on the European research policy scenario. For this reason, the great popularity world-class university rankings currently enjoy among research managers has caught bibliometricians off guard. Their success can only be understood when looking at the bigger picture and analyzing the evolution of the higher education landscape during the second half of the 20th Century. Here we summarize the main key factors which may help to understand their wholehearted adoption as research policy tools:

- **The globalization of the research-oriented model.** The increasing importance given to the so-called second mission leads to the emergence of a university model focused on research excellence. This enhances the globalization of higher education in which technology, knowledge, people, and ideas flow across national borders.
- **The internationalization of universities.** As a consequence of globalization, universities adopt an international model for promoting joint ventures and a race for talent developing international recruitment strategies.
- **Diversification of funds.** Governmental dependence weakens as universities adopt a diversified funding model, searching for private as well as public investment.
- **Massification of education.** The expansion of the mass education model from primary and secondary education reaches higher education which has to combine top research with massified teaching.

These factors lead to the 'Emerging Global University Model' which seeks to develop the World-Class Universities which compete on an international scenario, ignoring all national barriers (Mohrman, Ma & Baker, 2008). In this section we will analyze the use of international rankings in research policy, as they have become the main yardstick for benchmarking these new research-oriented super-universities (Baker, 2007). For this, we divided this section in four parts. The first two describe the origins of university rankings and the concerns of the bibliometric community with regard to the main international university rankings. The next two parts are focused on the consequences rankings have for two particular aspects of universities: national higher education systems and specialization. These two topics represent the main research challenges presented in this thesis. That is, the need to offer rankings which apply to different national systems while allowing international comparisons and complementary tools which allow us to analyze and develop university profiles. This latter aspect will lead us to Section 2.4 in which we focus on the development of science mapping and information visualization techniques to profile institutional diversity.

The origins of rankings: From reputational to bibliometric measures

The first scientist who popularized the idea of using rankings in order to measure and compare research performance was James McKeen Catell, who published in 1906 his *American Men of Science: A Biographical Directory* (van Noorden, 2010). There, he presented a ranked list of researchers by fields sorted according to the judgment of experts, following a rationale similar to that of peer review. However some antecedents can be tracked to end of the 19th Century in the United States (Orduña-Malea, 2012). Rankings have historically been the result of the initiative of individuals. But in the last decades the media and governmental institutions have turned their interest on them, becoming part of their research policy 'toolbox' (Salmi & Saroyan, 2007).

The globalization and democratization of higher education in the 1990s rushes different countries into a race for the most talented students and researchers worldwide, breaking down all traditional barriers (Hazelkorn, 2011). While this scenario well suits market-oriented university systems such as that in the United States or, to a certain extent, in the United Kingdom, European universities start to be perceived differently by potential students and professors, creating a fragmentation in the European research system (Nedeva, 2013). What has been developed as a horizontal system, in which all universities guaranteed a certain level of quality assurance, starts to break into a vertical system endangering institutional diversity. Potential students and researchers start to seek out the best institutions and it is in this context where rankings take a predominant role.

At the dawn of the new century, all national barriers start to collapse. Such fundamental change shifting from national higher education systems to a globalized scenario can be easily resumed in the great impact the launch of the Shanghai Ranking in 2003 had on research policy makers. Although it originally aimed at showcasing the position of Chinese universities in comparison with the United States (Zitt & Filliatreau, 2007), its impact worldwide has converted it into a yardstick used by research managers to measure the health of academic institutions and systems. Since then, others have followed their initiative developing their own university rankings using different methodologies and building on them new characteristics and functionalities (the methodologies employed by the main international rankings are reviewed in Aguillo et al., 2010; Buela-Casal et al., 2007).

An overview of the main international university rankings

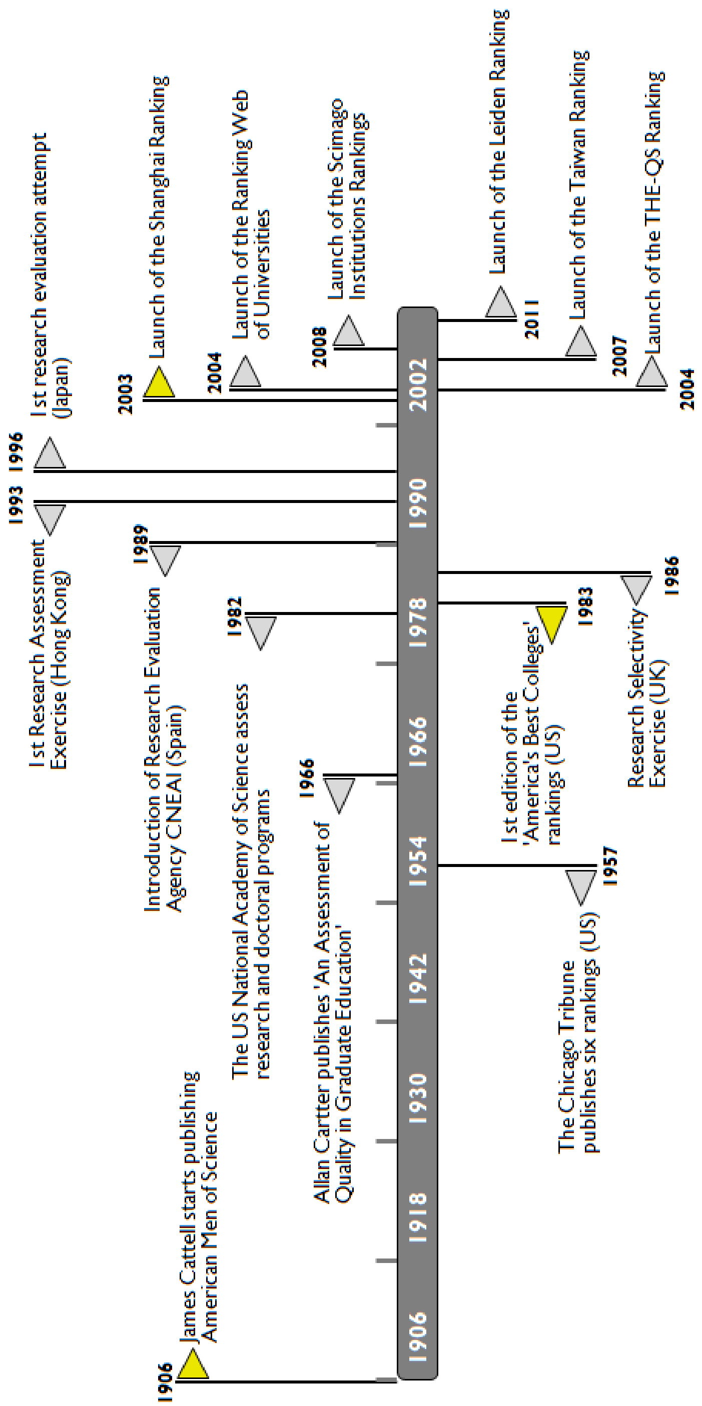
Here we will briefly review the methodologies employed by the main international university rankings and discuss the literature generated around them, focusing on the bibliometric perspective. In table 2.1 we include a list of the rankings that will be discussed indicating the launch year, frequency of publication, type of data and source used to retrieve the bibliometric data and the weight bibliometric indicators have in the final score.

The first ranking launched was the Shanghai Ranking in 2003, followed the next year by the Ranking Web of Universities (Aguillo, Ortega & Fernández, 2008) and the joint version of the THE-QS Ranking, which in 2007 split into two independent rankings. The NTU Ranking appeared that same year, originally developed by the Higher Education Evaluation and Accreditation Council of Taiwan and since 2012, published by the National University of Taiwan. In the subsequent years the Leiden Ranking appeared followed by the Scimago Institutions Rankings. However, the former was published at irregular intervals and it was not until 2011 when it adopted its current annual frequency and methodology. Regarding this latter aspect, all rankings are published annually with the exception of the Ranking Web of Universities which is published every six months. This has to do with the nature of the ranking which is based on webometric indicators, making its results less stable. All rankings referenced in Table 2.1 make use of bibliometric data, however, the weight of this data varies greatly between rankings. We find three rankings that rely exclusively on bibliometric data, while the Ranking Web of Universities only employs it for 16.7% of the final score. This data is retrieved exclusively from two distinct databases: the Web of Science to which we referred before, and Scopus, which is owned by Elsevier. This selection is expected as these databases represent the main citation indexes for international scientific literature (Gavel & Iselid, 2008).

Table 2.1. General description of the main international university rankings.

	Launch Year	**Frequency**	**Type of data**	**Bibliometric data source**	**Weight of bibliometric indicators**
Shanghai Ranking	2003	Annual	Bibliometric and reputational	Web of Science	90%
THE World Universities Rankings	2004	Annual	Bibliometric, surveys and manpower	Web of Science	60%
QS World University Rankings	2004	Annual	Bibliometric, surveys and manpower	Scopus	40%
Ranking Web of Universities	2004	Semi annual	Bibliometric and webometric	Scopus	16.7%
NTU Rankings	2007	Annual	Bibliometric	Web of Science	100%
Leiden Rankings	2008	Annual	Bibliometric	Web of Science	100%
Scimago Institutions Rankings	2009	Annual	Bibliometric	Scopus	100%

Figure 2.2. Timeline with the main milestones in the development of university rankings and performance-based research funding systems. In yellow, the most significant milestones refering to rankings. **Source:** Adapted and updated from Table I (Salmi & Saroyan, 2007)



Regarding other types of data used for ranking universities, the Shanghai Ranking includes reputational data based on Nobel Prizes or other field medal awards. The THE World Universities Rankings and the QS World University Rankings both use survey data to determine the quality of teaching or the perception certain so-called experts have of universities. They also use national statistics when available in order to determine the number of international students for instance and other similar variables. Finally, we find that the Ranking Web of Universities uses webometric data. The reason for this lies in the purpose of this ranking which is not to benchmark the perceived *quality* of universities, but to encourage Open-Access and web presence.

Hence, if we ignore this latter ranking which has a completely different nature, we observe two types of rankings: those which aim to analyze all dimensions of universities and those which only focus on the research mission of universities. We also observe that this conceptual difference follows a chronological order, establishing an analogy with the shift from peer review to bibliometric analysis in research evaluation as described in Section 2.2. In this case, bibliometric rankings represent the response of the bibliometric community to the emergence of the first international university rankings which were heavily criticised by the former.

Indeed, the launch of the THE-QS rankings and, especially, of the Shanghai Ranking caught from the beginning the interest of the bibliometric community. This was due mainly to three reasons: 1) it employed bibliometric indicators as part of the criteria for ranking academic institutions, 2) the authors had no bibliometric background and ignored previous literature and 3) the ranking suffered from serious shortcomings in the data collection as well as from an arbitrary methodology. Such strong opossition is resumed in the correspondence maintained between Professor van Raan from Leiden University and the creators of the Shanghai Ranking (Liu, Cheng & Liu, 2005; van Raan, 2005ab). Here, van Raan (2005a) highlights serious methodological and technical concerns which are later emphasized by others (i.e., Billaut, Bouyssou & Vincke, 2009). In Figure 2.3 we include the methodology employed by the Shanghai Ranking. Although van Raan's comments are directed at this ranking, many of the criticisms are also applied to the THE World Universities Rankings and the QS World University Rankings. For clarification, these can be summarized in four points:

1) **The criteria selected are not supported.** The indicators chosen for measuring reputation are not scientifically validated or justified. In the case of

the Shanghai Ranking, issues such as institutional splitting and merging (for instance, the split into two of the historical University of Berlin) when including Nobel prizes or field medal awards, or introducing an indicator based on publications in Nature and Science, which are journals biased towards certain scientific fields, are especially worrying. Regarding the THE World Universities Rankings and the QS World University Rankings, the main issues are the use of surveys, which create a positive feedback loop benefiting well-known universities (Waltman et al., 2012), and data availability and retrieval regarding manpower data, which is extremely complex (Bonaccorsi et al., 2007).

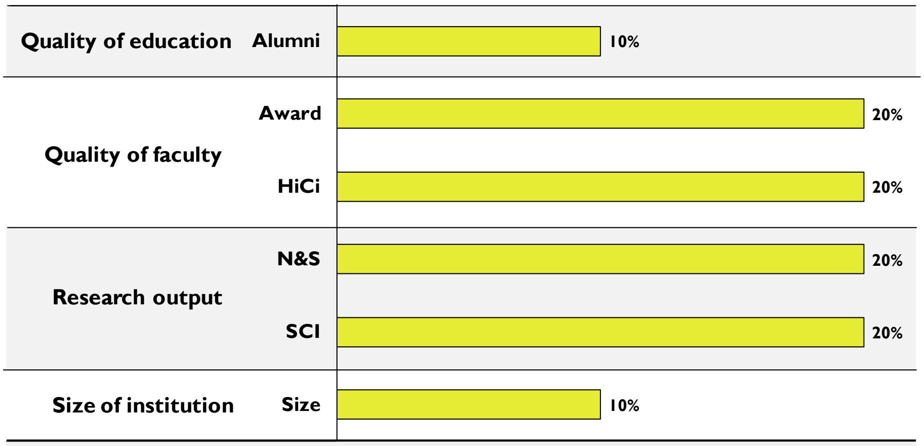
2) **The assignment of weights to the indicators used is arbitrary.** As Buela-Casal and colleagues (2007) explain, there is no justification for the weights assigned by the producers of the rankings. What is more, this assignment varies from one ranking to another, showing the lack of consensus among ranking producers.

3) **The lack of transparency.** The impossibility of reproducing the methodology employed by the authors of the rankings according to the indications they provide represents a severe threat to their credibility among the research community. This has been especially denounced in the case of the Shanghai Ranking (Florian, 2007). Despite this ambiguity, some researchers have succesfully revealed the insides of such ranking, showing the arbitrarieness of some of the decisions taken (Docampo, 2013).

4) **There is a bias towards universities with a strong research profile in certain fields.** Because they employ bibliometric indicators based on journal articles and use citation indexes to account for research performance, they inherit the shortcomings of citation analysis, especially with regard to the Social Sciences and Humanities (Moed, 2005). Indeed, these fields are neglected mainly for three reasons: 1) rankings ignore monographs as research output which are important document types in these fields (van Leeuwen et al., 2001), 2) rankings select indicators which benefit institutions that are strong in the Natural Sciences (for instance, the use of the N&S indicator by the Shanghai Ranking), and 3) rankings use Scopus and the Web of Science which are strongly biased towards the English-language, again to the detriment of the Social Sciences and Humanities (Torres-Salinas et al., 2013).

Thus, we see how the bibliometric community, which first viewed these new research evaluation tools with skepticism, embraced them into their literature either criticizing their methodologies, analyzing their capabilities or building on them more enhanced tools and suggesting their own solutions. The main problems derived from the methodologies employed by these rankings are conceptual. These conceptual issues can be resumed in two points. First, they consider universities as homogeneous units of analysis that can be compared, ignoring institutional diversity (Collini, 2011). Secondly, they aim at synthesizing data of a very different nature in a global indicator which presumes to represent in a single number the overall performance of such complex institutions.

Figure 2.3. Indicators and weight of the Shanghai Ranking since 2003. Based on Liu & Cheng (2005). Weights are stable over time (2003-2013 according to the methodology displayed for each edition at http://www.shanghairanking.com



Alumni	Alumni of an institution winning Nobel Prizes and Medal Fields
Award	Staff of an institution winning Nobel Prizes and Medal Fields
HiCi	Highly cited researchers in 21 broad subject categories
N&S	Articles published in Nature or Science
SCI	Articles indexed in SCI or SSCI
Size	Academic performance relative to size

In order to overcome such limitations, the subsequent three rankings which we here review (Leiden Ranking, Scimago Institutions Rankings and NTU Ranking), adopt a very different perspective. First, they focus only on research, the second mission of universities, acknowledging the inability to rigorously measure other dimensions of universities such as teaching or innovation. Secondly, they offer a very restrictive definition of what is considered research performance. Hence, research output is defined as journal articles mainly indexed in the corresponding database used for data retrieval produced by a given university during a fixed time frame. A good example of such restrictions is the complete name of the NTU Ranking: Performance Rankings of Scientific Papers for World Universities. Finally, regarding the use of a global indicator, the NTU Ranking is the only one of these three which employs it, while the other two include a battery of indicators acknowledging the impossibility of compressing into a single number the research performance of an institution. Once again, this shows discrepancies as to which methodology and choices are better.

Still, there are other technical issues which are acknowledged in many cases by the authors of bibliometric rankings and are difficult to avoid. Waltman and colleagues (2012) mention some of them when presenting the Leiden Ranking. These have to do with the collection of data which is mainly based on the *address field* of bibliographic records and with methodological decisions made for calculating the indicators. Here we summarize them:

1. **Data retrieval.** Not only institutional name changes and restructuring may affect the quality of the data retrieved, but also the lack of normalization of this field may lead to false positives (publications mistakenly assigned to a given institution) and false negatives (publications not assigned to the right institution). Robinson-García and Calero-Medina (2014) thoroughly analyze the *address* field in the Web of Science database offering a complete description on the many problems one may encounter when using this field for evaluation purposes.
2. **Methodological decisions.** Other differences between bibliometric rankings have to do with the battery of indicators shown as well as some methodological decisions taken when calculating them. These decisions have to do with the counting of publications (full or fractional counting), the inclusion of document types or coverage differences derived from the use of the Web of Science or Scopus databases.

Rankings for evaluating National Higher Education Systems

The development of proper methodologies for the elaboration of research rankings is an on-going research front in which many variables and questions still remain unsolved. But still, if we rely on the imperfect picture drawn here, one may question why these tools have had such success with research managers despite their many caveats. The answer is twofold. On the one hand, international university rankings are powerful and persuasive tools for demonstrating university performance to external parties from the university management such as potential students, the media and politicians (Bornmann, 2014). The interest does not rely necessarily on their use for developing research strategies rather than for convincing others on the success of previous decisions. On the other hand, international rankings, based completely or partially on bibliometric data, provide accountable and seemingly objective metrics on an international scale. This allows research managers to benchmark their own performance with others, disregarding any methodological concern (Hazelkorn, 2008).

International rankings serve as yardsticks for measuring the pulse of a rising global higher education market. They are the only tools for acknowledging what until their emergence was a blurred world-scale picture in which only certain super-universities stood out (Harvard, Oxford or MIT for instance) while the rest remained hidden (Marginson & van der Wende, 2007). Such visibility accelerates the globalization of the research-oriented university model discussed at the beginning of this section. They also encourage competition not only at the institutional level, but at the national level, urging governments to offer reliable national statistics and to establish incentives for universities to stimulate competition. But this institutional stratification finally benefits those countries in which universities were already regulated in such a way, that is, mainly the United States and the United Kingdom (Robinson-Garcia et al., 2014).

In this sense, the supremacy of the Shanghai Ranking above the rest as an influential player in the higher education landscape is undisputable. It originally aimed to compare Chinese universities' positions with World-Class universities but, due to the relevance it has gained, it is now used by research managers, students, researchers and the media all over the world. It is also one of the few which has been described in peer-reviewed literature by its authors (Liu & Cheng, 2005). Among others, research managers and national governments have used this ranking to evaluate the *health* of their national university systems (i.e., Docampo, 2011; Marginson & van der Wende, 2007). This type of exercise aims to rethink and reformulate national university systems, especially from European countries as, in the words of Aghion and colleagues (2010), 'there is little point on promoting competition among universities if they do not have sufficient autonomy to respond with more productive, inventive, or efficient programs'.

There is a perceived notion of the inferiority in terms of research excellence of European versus US research (Herranz & Ruiz-Castillo, 2013) which is clear when analyzing the nationality of the universities in international rankings (Robinson et al., 2014). However, European research managers have shown a special sensitivity towards the shortcomings of these rankings in general, and specifically towards the threat to institutional diversity that they represent.

Therefore, efforts have been directed at establishing the bases for promoting *horizontal diversitication* (universities with different profiles and missions) while allowing the *vertical stratification* that university rankings introduce (benchmarking and comparing) (van Vught & Ziegele, 2012). For this, the European Union developed two related projects each of aimed at these two instances: the U-Map and the U-Multirank:

- **U-Map[3].** The U-Map is a project for developing a profiling and classification tool similar to the Carnegie Classification in the United States. It is funded by the European Union and is led by the Centre for Higher Education Policy Studies (CHEPS) from the University of Twente (The Netherlands). Rauhvargers (2011) provides full details of the methodology it follows. In brief, it provides 29 indicators based on six different aspects of universities: 1) teaching and learning, 2) student profile, 3) research involvement, 4) knowledge exchange, 5) regional engagement and 6) international orientation. The idea is to provide non-hierarchical profiles of universities in order to allow fair comparisons. According to the information provided on its website, the project will soon be publicly available after profiling more than 300 European universities.
- **U-Multirank[4].** Funded by the European Union, it is an on-going project conducted by the Consortium for Higher Education and Research Performance Assessment (CHERPA Network), led by the Centre for Higher Education (Germany) and CHEPS. Other members are the Centre for Science and Technology Studies CWTS (The Netherlands), Elsevier, the Bertelsmann Foundation (Germany) *folge3*, a German software developer. The aim of this project is to provide a multidimensional, user-driven ranking. The idea is that users select the criteria to rank universities while only those with similar profiles are compared.

Another perspective is the one adopted by the *Universitas 21* international network of universities, which develops a rankings of national higher education systems, considering that rather than comparing universities individually, what may be in the interest of research policy makers is to benchmark which national system works best (Williams et al., 2013). This is the position adopted by Calero-Medina and colleagues (2008), who consider that 'bibliometric outcomes of an individual university can only be interpreted properly when one takes into account the structure of the national academic system in which it is embedded', and later confirmed by Bornmann, Mutz and Daniel (2013)

[3] http://www.u-map.eu/
[4] http://www.u-multirank.eu/

who prove that the location of a university plays an important role in performance. Such initiatives reveal two important facts: the final consolidation of international university rankings and a global higher education landscape, and the need to preserve institutional diversity at the national level despite external pressures.

The disciplinary issue

The other important drawback on which we will focus is the inability of university rankings to represent the variability of a university's performance by fields (Calero et al., 2008). Reducing to a single measure the activity of universities is a simplistic view even when focusing on a single dimension such as research. Even if they offer a battery of indicators instead of a single one, these do not allow the reader to see the fields in which a university is strong or to interpret which fields influence most the final position of a given university. Also, due to publication and citation differences, it is possible that certain university profiles may have an advantage over others, leading to 'unfair' comparisons. Indeed, it seems that universities specialized in certain scientific domains may perform better in university rankings even when using normalized indicators (Bornmann, Moya-Anegón & Mutz, 2013), and especially if their research is of a multi-disciplinary nature (Moed et al., 2011).

In order to overcome this important limitation, many of the main international university rankings have started to offer in the last few years rankings by fields. The first one to do so was the Shanghai Ranking in 2007, including five rankings by broad fields and in 2009, five more rankings by specific disciplines. Subsequently, the THE Ranking, the QS Rankings and the NTU Ranking have followed ther example. The Leiden Ranking has been the last one to follow this trend, offering in its 2013 edition five rankings by broad fields. In principle, these rankings would offer a more complete picture than global rankings, showing a multiple perspective of the performance of a university. But still, they do not allow the reader to easily detect university profiles. For this, López-Illescas, Moya-Anegón and Moed (2011) suggest using graphs and other complementary tools along with rankings in order to provide policy managers with a more precise picture.

The development of visualization tools for research policy is a research front in itself in bibliometrics (Noyons, 2005). Many efforts have been made along these lines but until recently (i.e., Bornmann et al., 2014; García et al., 2013) none focused on complementing university rankings. For example, we here highlight the works of Bornmann and Leydesdorff (2012), Ortega and colleagues (2008) or Leydesdorff and Persson (2010). In section 2.4 we will review the most relevant literature on the development of classification and visualization techniques for analyzing university disciplinary profiles focusing on the methodological aspects and approaches taken.

2.4. Classifying and visualizing disciplinary profiles

The development of science mapping techniques has been within the interest of the field of research evaluation since its conception. Nevertheless, Derek de Solla Price already envisioned the potential use of publications and citations as their main link to manifest the overall structure of scientific fields (Price, 1965). Since then and especially thanks to the advances in technology, this specific field has evolved expanding its use, not only focusing on domain analysis (Hjørland & Altbrechtsen, 1995) but also on research assessment (Noyons, Moed & Luwel, 1999). Science mapping techniques allow us to analyze the structure of scientific domains, define relations between the units of analysis or classify and identify research profiles (Ingwersen, Larsen & Noyons, 2001; Shiffrin & Börner, 2004). Visualization techniques offer easy-to-read solutions to rapidly establish the structure of a given set of objects identifying the main elements in it. Also, social network analysis strengthens its interpretation allowing us to apply network theory and explain the patterns followed by the model represented in the graph (Calero-Medina, 2012; Vargas-Quesada & Moya-Anegón, 2007).

In this section we will briefly discuss the main issues regarding classification and science mapping techniques. We will first review the historical development of the field of science mapping within the field of bibliometrics and focus on its application to research assessment. Then, we will address the process flow one must follow to elaborate science maps. Here we will rely on the review presented by Börner, Chen & Boyack (2003) in which they give full account of the different aspects that must be considered. We will draw attention to some specific techniques employed for dimensionality reduction as well as the most common measures and units of analysis. Since the 1930s, the theory of networks has been largely employed in the social sciences in general (Borgatti et al., 2009) and specifically in the field of bibliometrics (Calero-Medina, 2012). We will give the key points for understanding network theory and social network analysis. Finally, we will go through recent literature using visualization and classification techniques for grouping or profiling research institutions in order to link this with chapter 3.

Science mapping as a tool for research evaluation: Historical roots and major milestones

Probably, the first attempt to represent scientific progress through mapping techniques was that performed by Eugene Garfield. He aimed to analyze the history of science using citation data and presented as a case study the discovery of DNA based on Isaac Asimov's book 'The Genetic Code' (Garfield, Sher & Torpie, 1964). The idea was to test if citation data was a good measure for the development of topological maps of science. Using the bibliography employed by Asimov in his book, he elaborated longitudinal maps that could reflect the links between the key contributions that would ultimately lead to Watson and Crick's discovery (Garfield, 2004).

During the 1960s, 1970s and 1980s the main bases of science mapping and the measures for linking papers, authors, institutions, countries and disciplines were formulated. Indeed, one year before Garfield presented his historical map on DNA, Kessler (1963) suggested the concept of bibliographic coupling as a measure of relation between scientific papers, that is, grouping documents in terms of the number of common references they have. Henry Small and his team continued during the next decades further exploring the use of citation data for science mapping. They developed the concept of co-citation, which is defined as the relationship established between two documents which are cited by a third paper (Small, 1973; Small & Garfield, 1985; Small & Griffith, 1974). In their works, Small and his team already highlight the potentials of science mapping for research evaluation. For instance, they consider that these techniques offer a new perspective on the use of citation data for assessing research on the relevance and relative value of papers, authors and institutions (Griffith et al., 1974).

At the end of the 1970s and based on another of Price's ideas, the evolution of the scientific enterprise from individuals working alone to the development of large research teams (de Solla Price, 1963), Beaver and Rosen (1979) conducted their seminal study on scientific collaboration. Here they introduced co-authorship as a measure within research evaluation leading to a need to understand the meaning and motives behind scientific collaboration (Katz & Martin, 1997). Also, their contribution offers a unique opportunity to extend the social network theory to the field by developing science maps based on co-authorship (Newman, 2001; Schubert & Braun, 1990).

Collaboration networks at the individual level may lead to the discovery of *invisible colleges* (Crane, 1972), while aggregating at the meso- and macro-level leads to the analysis of the social structure of science by studying the relations between countries and institutions and changes overtime (Glänzel & Schubert, 2005).

The latest milestone took place with the contributions of Callon and colleagues on co-word analysis. Here, the link between the units of analysis are words themselves, which are considered the best representations of the contents of research contributions. Callon and colleagues (1983) justify their proposal in the need to overcome the shortcomings co-citation analysis introduces when developing science maps. They subscribe their opinions to those of Cole and Cole (1963) and their criticisms of the Mertonian perspective on the sociology of science (for further details see section 2.1). Their concerns on co-citation analysis rely on the fact that 'scientists, whoever they are, can only have a partial and biased view of the social and intellectual organization of their disciplines' (Callon et al., 1983). They stress their concerns about co-citation analysis on the fact that the distinctions between the research community and other communities participating on research advancement are not clear, hence citation practices are not universal. Also, they indicate that using citation data 'consolidates the controversial notion of specialty areas because it attempts to reveal cognitive structures, whose theoretical status is not clear, on the basis of social practices (citations) whose meaning is not properly understood' (Callon et al., 1983).

Thus, the three main measures for establishing links in science are formulated and almost simultaneously employed for research evaluation purposes, where their application seems evident. While co-authorship seems to show explicit relations within the scientific community, co-citation and co-word analysis seem to be two means to identify implicit relations. In this sense, the open confrontation between the two methods is soon settled. The Centre for Science and Technology Studies CWTS (The Netherlands), which played an important role in the 1980s on the development of the field of research evaluation (see section 2.1), appears on the scene, introducing an integrated model for developing maps of science by combining both measures (Braam, Moed & van Raan, 1991; Noyons, Moed & Luwel, 1999). However, with few exceptions (López-Herrera et al., 2009; Ruiz-Baños et al., 1999), co-citation analysis seems to have become the most common measure for science mapping (i.e., Boyack, Klavans & Börner, 2005; Waltman, van Eck & Noyons, 2010; Moya-Anegón et al., 2004) along with co-authorship, relegating co-word analysis to the background.

A short how-to guide on the process flow for science mapping

Now we will review the basic aspects on the construction of maps of science. For this we will follow the structure presented by Börner, Chen and Boyack (2003). In their review, they establish six steps which must be followed for developing any type of information visualization. Here we will pay special attention to the techniques, measures and methodologies which are more common in research evaluation. This is crucial, as science mapping aims to offer an overall picture of the structure that is being analyzed. This means that large amounts of data will be displayed with a limited resolution and space. Also, we often turn to these techniques when analyzing multivariate data. This data will usually end up being reduced to a two dimensional graph, which means that the technique used for reducing the dimensionality may be crucial for the analysis. These techniques are especially interesting for identifying similar profiles and classifying objects. Figure 2.4 includes a summary of the process flow. We will now describe this process flow, explaining each of the steps and the different options one may encounter:

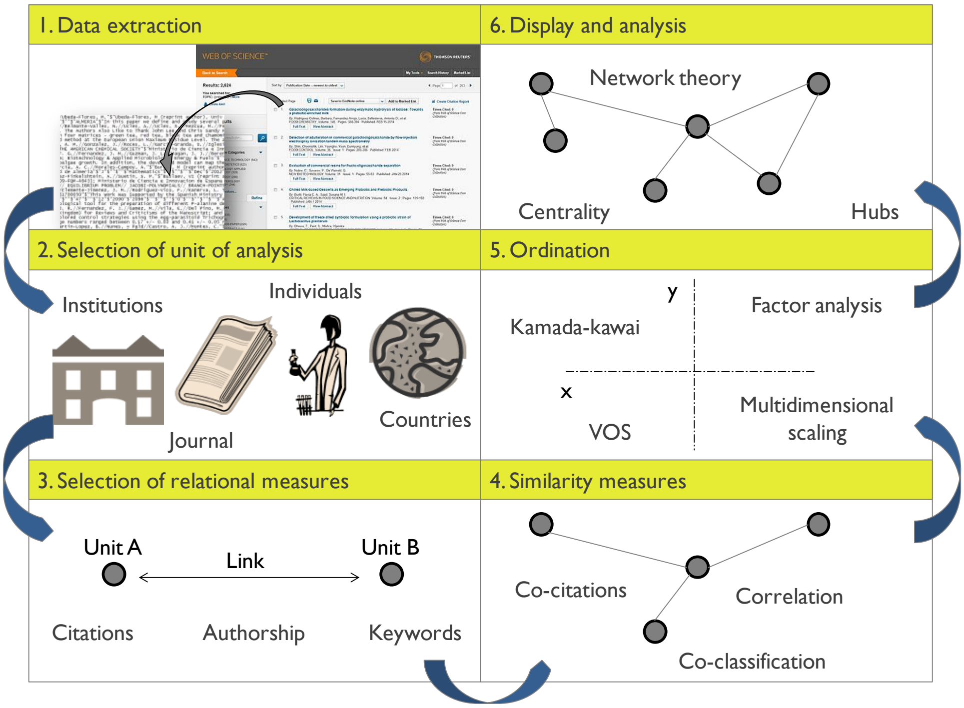
1. **Data extraction.** Although in the last decade other alternative data sources have appeared (see discussion in section 2.1), the most common databases for retrieving bibliometric information are the Thomson Reuters Web of Science and Elsevier's Scopus. Not only do they include citation data, but their multidisciplinary nature, coverage and selective inclusion criteria for indexing journals have led them to become the most relevant databases currently (Gavel & Iselid, 2008). Although other alternative databases have emerged during the last years which seem very promising, they are still not considered as reliable and manageable as these two. Here we must highlight the emergence of Google Scholar which seems the most firm alternative (Harzing, 2013; Meho & Yang, 2007; Kulkarni et al., 2009). But issues regarding its lack of transparency and instability disregard its current use for bibliometric purposes (Aguillo, 2012; Bohannon, 2014; Delgado López-Cózar, Robinson-García & Torres-Salinas, 2013; 2014). The main reason for looking for alternative databases from a bibliometric point of view, rely on the well-known biases these databases may introduce in the analyses, especially when studying the Social Sciences and Humanities (Moed, 2005).

2. **Unit of analysis.** Once we have retrieved and processed the data, we must select the unit of analysis on which we are interested. Noyons (2005) considers that the most important fields that may be subjected to analysis are:
    - Authorship
    - Name of the publication
    - Publication name
    - Year of publication
    - Affiliation
    - Abstract

    However, other fields such as references, keywords or publisher may be subjected to analysis. Also fields such as affiliation may be used for identifying and processing different units of analysis such as countries or academic institutions. In this thesis our focus will inevitably be on this latter unit.
3. **Relational measures.** At the beginning of this section we reviewed the main relational measures employed in bibliometric studies. These are: co-citations, co-words and co-authorship. However, many other measures may be used such as inter-citations, direct citations or bibliographic coupling. Measures using the prefix *inter* are defined as the count of times that any unit links another one as well as itself in a matrix. The prefix *co* is used for defining co-occurrences.

    Although relational measures are usually related to specific fields of bibliographic records, one may use specific bibliometric indicators for relating units of analysis. For instance, García and colleagues (2013) use the Impact Factor of journals in which academic institutions publish to develop heliocentric benchmarking maps for identifying institutions similar to a given one. Another alternative would be to analyze cited author names instead of co-citations (Noyons, 2005). Also, one may choose to combine different relational measures following the example set by Braam, Moed and van Raan (1991). For this, they first cluster documents using co-citation analysis and then develop word-profiles based on co-word analysis to name these clusters. This methodology was recently employed by Waltman and van Eck (2012) in order to develop a publication-level classification system for science.

Figure 2.4. Process flow for the construction of science maps based on Börner, Chen & Boyack, 2003.



4. **Similarity.** In order to develop similarity measures between the units of analysis one must produce a similarity matrix representing the degree of similarity between them. We may use scalar values such as direct counts of citations for instance, or normalize the matrix when analyzing co-occurrences, especially if we aim to later apply a multivariate analysis technique (van Eck & Waltman, 2009). Originally, the technique employed for this task was the Pearson *r* correlation. However, during the 2000s a fierce debate originated questioning the use of this measure and other alternative measures were proposed such as the chi-square distance, Salton's cosine similarity (derived from the vector space model in information retrieval) or the Jaccard index (Ledesdorff, 2008; Schneider & Borlund, 2007). In a more recent study, van Eck and Waltman (2009) suggest the use of probabilistic measures.

5. **Ordination.** Here we refer to the algorithms used to position the units of analysis that are similar close together, and place those which are dissimilar

far apart. These algorithms and techniques are employed for reducing the dimensionality of the representation without losing much quality (goodness of fit). Traditionally, the main classifying methodologies employed for representing bibliographic data have been those based on multivariate analysis such as Multi-Dimensional Scaling (MDS), Principal Component Analysis (PCA), Correspondence Analysis or Hierarchical Cluster Analysis, for instance.

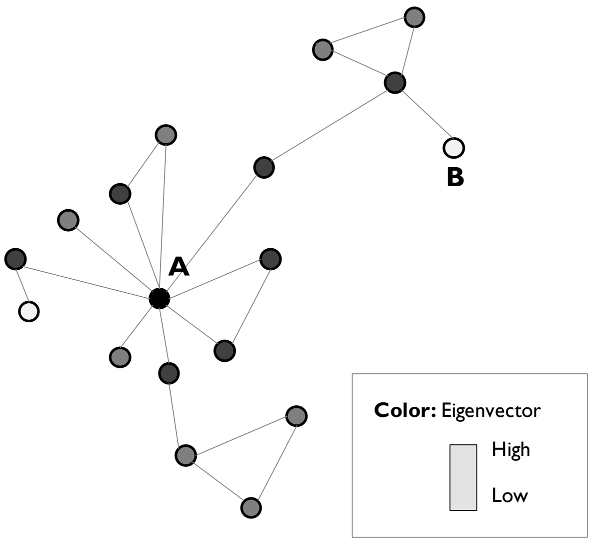
MDS is a visualization technique used for exploring similarities and dissimilarities contained in a distance matrix. It places each unit in an n-dimensional space trying to preserve the distances between objects as well as possible. Then, it assigns coordinates to each unit in n-dimensions. Usually the MDS algorithm is performed in a two-dimensional space in order to represent it on a scatterplot. PCA is a mathematical methodology that uses orthogonal transformation converting a set of cases of possibly correlated variables into a set of values of uncorrelated variables which are known as principal components aiming at reducing the number of variables. CA is a multivariate statistical methodology similar to PCA, providing the means to display and summarize a set of data in a two-dimensional graph. Finally, Hierarchical Cluster Analysis is a type of cluster analysis for grouping a set of objects which are considered to share common attributes and hence belong to the same class (or cluster).

6. **Display and analysis.** The last step for the elaboration of the science map and one of the most important ones is the display stage (Noyons, 2005). Here, the most common visualization technique is the social network (i.e., Groh & Fuchs, 2011) by placing nodes according to the Kamada Kawai algorithm or the Fruchterman and Reingold algorithm (Vargas-Quesada & Moya-Anegón, 2007) for instance. Other visualization techniques employed have been Pathfinder Networks (White, 2003), visualization of similarities VOS (van Eck et al., 2010) or Kohonen self-organizing maps SOM (Moya-Anegón, Herrero-Solana & Jiménez-Contreras, 2006).

   The importance in the selection of the display is crucial as it will condition the analysis and interpretation of the data displayed in the map. We will now identify the key aspects related to network theory in order to interpret this type of visualization. For this we will follow the introduction offered by Calero-Medina (2012). To illustrate the explanation, in figure 2.5 we highlight the points we will subsequently refer to. Network graphs consist of nodes connected by edges. Each node represents an object while the edges represent the relations between objects. Hence if we were analyzing an institutional collaboration map, the nodes would represent universities. Following our example, we would say that university A and university B have a common set of co-authored papers, while A and C have no co-authored publications.

The interest in network theory relies on the network measures one can infer from them. These indicators will provide further understanding of the role played by each node within the network. Here the main indicators are those which measure the centrality of the nodes. The simplest one is the degree centrality which is defined as the number of edges that connect a given node with the rest. In our example, the degree of node A is 9 while the degree of node B is 2. If we are working with directed networks (i.e., node A cites B but B does not cite A) we would have two types of degree: the in-degree (i.e., citations received by A) and the out-degree (i.e., references made by A).

Figure 2.5. Example of social network. Colors represent the eigenvector centrality.



Another centrality indicator is betweenness. A node will have a high betweenness if it appears often on the shortest path that connects any other two nodes. The betweenness indicator is employed in chapter 8. A third centrality indicator is closeness. Here we consider that the most central node will be the one that connects with a higher number nodes with no intermediate nodes (this is used in chapter 7). Finally, we must refer to eigenvector centrality (Pinski & Narin, 1976). In this case, nodes are weighted according to the 'prestige' of the nodes to which they are connected, which passes on to them and is defined by the number of connections with the rest of the network. In our example, we observe that node A is the one with the highest eigenvector as it is central to the structure. We find that nodes connected to A have a higher centrality value than node C for instance, as this node is connected to a node with a lower eigenvector value.

Finally we will review the following properties of social networks :

- **Small-world effect.** This effect states that the mean geodesic distance between pairs of nodes is very short compared to the size of the whole network. The geodesic distance is defined as the shortest path between two nodes. This property was first explored by Pool and Kochen (1978-1979) and means that all nodes are just a few 'connections' away from any other node in the network.
- **Transitivity.** This property indicates that two nodes which are very similar to a third node have a high probability of being very similar to each other (Watts & Strogatz, 1998). Thanks to this property, network analysis allows us to reveal communities as these are usually detected by identifying highly clustered regions of the network.
- **Exclusiveness.** This property is probably a consequence of the former. The degree of exclusiveness indicates the difficulty of an outside node to access highly clustered regions of the network with strong ties to each other.

Mapping and classifying research institutions for research policy assessment

Finally, we will review recent literature attempting to classify academic institutions or defining institutional research profiles. Referring to the former, Shin (2009) employs Hierarchical Cluster Analysis to classify South Korean universities. He develops a mission-level classification based on research performance establishing five different types of university. However he warns against the heavy reliance of research managers on his classification as it is sensitive to disciplinary specialization. Ortega, López-Romero & Fernández (2011) perform a similar exercise to classify the 109 research institutes of the Spanish National Research Council. For this, they apply three different techniques (Principal Components Analysis, Agglomerative Hierarchical Cluster Analysis and Linear discriminant analysis). In this case, the resulting classification identifies disciplinary aspects of the institutes and defines three types: technological, humanistic and scientific. What is more, they are able to assign publication practices regarding document types to each institutional profile. This has important consequences as it permits the development of specific research evaluation exercises for each type of institution.

Another perspective is to analyze university profiles and characterize them instead of establishing a classification of universities. For instance, developing inverse research profiles (Calero-Medina & van Leeuwen, 2012) by breaking down the subject categories or field areas into the institutions which contribute most to each of these areas according to the overall university system. Adams, Gurney and Marshall (2007) propose creating impact profiles by comparing a given university or country to the world average for instance. Carpenter and colleagues (1988) suggest a model for profiling universities in which a battery of indicators such as the Activity Index, number of papers or % papers with influence are given to each field in which a university performs all with reference to a given average (for instance, the whole system's average or the world average).

Last but not least, we must mention the studies conducted by Bordons and colleagues (2010) or Sanz-Casado and colleagues (2013) where they suggest using input and contextual data along with bibliometric data in order to analyze and establish institutional profiles as this information will allow us to better capture the characteristics of universities.

References

Abbott, A., Cyranosk, D., Jones, N., Maher, B., Schiermeier, Q., & van Noorden, R. (2010). Metrics: Do metrics matter? *Nature, 465*(7300), 860-862.

Abramo, G., Cicero, T., & D'Angelo, C.A. (2011). The dangers of performance-based research funding in non-competitive higher education systems. *Scientometrics, 87*(3), 641-654.

Adams, J., Gurney, K., & Marshall, S. (2007). Profiling citation impact: A new methodology. *Scientometrics, 72*(2), 325-344.

Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A., & Sapir, A. (2010). The governance and performance of universities: Evidence from Europe and the US. *Economic Policy, 25*(61), 7-59.

Aguillo, I.F. (2012). Is Google Scholar useful for bibliometrics? A webometric analysis. *Scientometrics, 91*(2), 343-351.

Aguillo, I.F., Bar-Ilan, J., Levene, M., & Ortega, J.L. (2010). Comparint university rankings. *Scientometrics, 85*(1), 243-256.

Aguillo, I.F., Ortega, J.L., & Fernández, M. (2008). Webometric Ranking of World Universities: Introduction, methodology, and future developments. *Higher Education in Europe, 33*(2-3), 233-244.

Albarrán, P., Crespo, J.A., Ortuño, I., & Ruiz-Castillo, J. (2011). The skewness of science in 219 sub-fields and a number of aggregates. *Scientometrics, 88*(2), 385-397.

Almind, T., & Ingwersen, P. (1997). Informetric analyses on the World Wide Web: Methodological approaches to 'Webometrics'. *Journal of Documentation, 53*(4), 404-426.

Anon (2010). Assessing assessment. *Nature, 465*(7300), 845.

Archambault, É., & Larivière, V. (2009). History of the Journal Impact Factor: Contingencies and consequences. *Scientometrics, 79*(3), 635-649.

Auranen, O., & Nieminen, M. (2010). University research fundin and publication performance - An international comparison. *Research Policy, 39*(6), 822-834.

Bailón-Moreno, R., Jurado-Alameda, E., Ruiz-Baños, R., Courtial, J.P., & Jiménez-Contreras, E. (2007). The pulsing structure of science: Ortega y Gasset, Saint Matthew, fractality and transfractality. *Scientometrics, 71*(1), 3-24.

Baker, D. Mass higher education and the super research university : A symbiotic relationship. *International Higher Education, 49*, 9-10.

Barker, K. (2007). The UK Research Assessment Exercise: The evolution of a national research evaluation system. *Research Evaluation, 16*(1), 3-12.

Beaver, D., & Rosen, R. (1978). Studies in scientific collaboration. Part 1. The professional origins of scientific co-authorship. *Scientometrics, 1*(1), 65-84.

Bergstrom, C.T., West, J.D., & Wiseman, M.A. (2008). The Eigenfactor[tm] metrics. *The Journal of Neuroscience, 28*(45), 11433-11434.

Billaut, J.C., Bouyssou, D., & Vincke, P. (2009). Should we believe the Shanghai ranking? An MCDM view. *Scientometrics, 84*(1), 237-263.

Bohannon, J. (2014). Google wins raves - But can it be trusted? *Science, 343*(6166), 14.

Bollen, J., Rodriguez, M.A., & Van de Sompel, H. (2007). MESUR: Usage-based metrics of scholarly impact. In: *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries*, pp. 474.

Bollen, J., Van de Sompel, H., Smith, J.A., & Luce, R. (2005). Toward alternative metrics of journal impact: A comparison of download and citation data. *Information Processing & Management, 41*(6), 1419-1440.

Bollen, J., & Van de Sompel, H. (2006). Mapping the structure of science through usage. *Scientometrics, 69*(2, 227-258.

Bonaccorsi, A., Daraio, C., Lepori, B., & Slipersaeter, S. (2007). Indicators on individual higher education institutions: Addressing data problems and comparability issues. *Research Evaluation, 16*(2), 66-78.

Bordons, M., Morillo, F., Fernández, M.T., Gómez, I., León, M., & Diego, D.M. (2005). *La investigación matemática española de difusión internacional. Estudio bibliométrico (1996-2001).* Madrid: Consejo Superior de Investigaciones Científicas.

Bordons, M., Sancho, R., Morillo, F., & Gómez, I. (2010). Perfil de actividad científica de las universiddes españolas en cuatro áreas temáticas: Un enfoque multifactorial. *Revista Española de Documentación Científica, 33*(1), 9-33.

Borgatti, S.P., Mehra, A., Brass, D.J., & Labianca, G. (2009). Network analysis in the Social Sciences. *Science, 323*(5916), 892-895.

Borgman, C.L., & Furner, J. (2002). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology, 36*, 3-72.

Börner, K., Chen, C., & Boyack, K.W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology, 37*, 179-255.

Bornmann, L. (2011). Scientific peer review. *Annual Review of Information Science and Technology, 45*, 197-245.

Bornmann, L. (2014). On the function of university rankings. *Journal of the Association for Information Science and Technology, 65*(2), 428-429.

Bornman, L., & Leydesdorff, L. (2012). Which are the best performing regions in information science in terms of highly cited papers? Some improvements of our previous mapping approaches. *Journal of Informetrics, 6*(2), 336-345.

Bornmann, L., Leydesdorff, L., & Mutz, R. (2013). The use of percentiles and percentile rank classes in the analysis of bibliometric data: Opportunities and limits. *Journal of Informetrics, 7*(1), 158-165.

Bornmann, L., Moritz, S., Moya-Anegón, F., & Mutz, R. (2014). Ranking and mapping or universities and research-focused institutions worldwide based on highly-cited papers: A visualisation of results from multi-level models. *Online Information Review, 38*(1), 43-58.

Bornmann, L., Moya-Anegón, F., & Mutz, R. (2013). Do universities or research institutions with a specific subject profile have an advantage or a disadvantage in institutional rankings? *Journal of the American Society for Information Science and Technology, 64*(11), 2310-2316.

Bornmann, L., Mutz, R., & Daniel, H.-D. (2013). Multilevel-statistical reformulation of citation-based university rankings: The Leiden ranking 2011-2012. *Journal of the American Society for Information Science and Technology, 64*(8), 1649-1658.

Boyack, K.W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics, 64*(3), 351-374.

Braam, R.R., Moed, H.F., & van Raan, A.F.J. (1991). Mapping science by combined co-citation and word analysis. II: Dynamical aspects. *Journal of the American Society for Information Science, 42*(4), 252-266.

Buela-Casal, G., Gutiérrez-Martínez, O., Bermúdez-Sánchez, M.P., & Vadillo-Muñoz, O. (2007). Comparative study of international academic rankings of universities. *Scientometrics, 71*(3), 349-365.

Bush, V. (1945). *Science: The endless frontier*. Washington, DC: US GPO.

Butler, L. (2003). Explaining Australia's increased share of ISI publications - The effects of a funding formula based on publication counts. *Research Policy, 32*(1), 143-155.

Cabezas-Clavijo, Á., Robinson-García, N., Escabias, M., & Jiménez-Contreras, E. (2013). Reviewers' ratings and bibliometric indicators: Hand in hand when assessing over research proposals? *PLOS One, 8*(6), e68258.

Calero-Medina, C. (2012). *Links in science: Linking network and bibliometric analyses in the study of research performance* [Doctoral thesis]. Leiden: Leiden University.

Calero-Medina, C., López-Illescas, C., Martijn, V.S., & Moed, H.F. (2008). Important factors when interpreting bibliometric rankings of world universities: An example from oncology. *Research Evaluation, 17*(1), 71-81.

Calero-Medina, C., & van Leeuwen, T.N. (2012). Reorganizing research with the help of bibliometric collaboration networks. In: C. Calero-Medina. *Links in science: Linking network and bibliometric analyses in the study of research performance* [Doctoral thesis]. Leiden: Leiden University, pp. 45-66.

Callon, M., Courtial, J.-P., Turner, W.A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information, 22*(2), 191-235.

Carpenter, M.P., Gibb, F., Harris, M., Irvine, J., Martin, B.R., & Narin, F. (1988). Bibliometric profiles for British academic institutions: An experiment to develop research output indicators. *Scientometrics, 14*(3-4), 213-233.

Clerides, S., Pashardes, P., & Polycarpou, A. (2011). Peer review vs. metric-based assessment: Testing bias in the RAE ratings of UK economics departments. *Economica, 78*(311), 565-583.

Cole, S., & Cole, J.R. (1967). Scientific output and recognition: A study in the operation of the reward system in science. *American Sociological Review, 32*(3), 377-390.

Cole, J.R., & Cole, S. (1972). The Ortega hypothesis. *Science, 178*(1972), 368-375.

Collini, S. (2011). *What are universities for?* London: Penguin Books.

Costas, R., & Bordons, M. (2005). Bibliometric indicators at the micro-level: Some results in the area of natural resources at the Spanish CSIC. *Research Evaluation, 14*(2), 110-120.

Costas, R., & Bordons, M. (2007). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics, 1*(3), 193-203.

Costas, R., & van Leeuwen, T.N. (2012). Approaching the "reward triangle": General analysis of the presence of funding acknowledgments and "peer interactive communication" in scientific publications. *Journal of the American Society for Information Science and Technology, 63*(8), 1647-1661.

Costas, R., van Leeuwen, T.N., & Bordons, M. (2010). A bibliometric classificatory approach for the study and assessment of research performance at the individual level: The effects of age on productivity and impact. *Journal of the American Society for Information Science and Technology, 61*(8), 1564-1581.

Crane, D. (1972). *Invisible colleges: Diffussion of knowledge in scientific communities.* Chicago: University of Chicago Press.

Cronin, B. (1984). *The citation process. The role and significance of citations in scientific communication.* London: Taylor Graham.

Cronin, B., & Weaver, S. (1995). The praxis of acknowledgment: From bibliometrics to influmetrics. *Revista Española de Documentación Científica, 18*(2), 172-177.

Day, R.A. (1998). *How to write and publish a scientific paper.* Phoenix, AZ: Oryx Press.

De Bruin, R.E., Straathof, A., & Moed, H.F. (1992). Measuring the effects of European science policy: The assessment of scientific output and international scientific cooperation in CEC target fields, 1985-1990. In: *Science and Technology in a Policy Context. Selected Proceedings of the Joint EC-Leiden Conference on Science and Technology Indicators.* Leiden, The Netherlands, pp. 453-467.

De Solla Price, D.J. (1951). Quantitative measures of the development of science. *Archives Internationales d'Histoire des Sciences, 14*, 85-93.

De Solla Price, D.J. (1963). *Little science, big science.* New York: Columbia University Press.

De Solla Price, D.J. (1965). Networks of scientific papers. *Science, 149*(3683), 510-515.

Delanghe, H., Sloan, B., Muldur, U. (2010). European research policy and bibliometric indicators, 1990-2005. *Scientometrics, 87*(2), 389-398.

Delgado López-Cózar, E., Robinson-García, N., & Torres-Salinas, D. (2013). Science communication: Flawed citation indexing. *Science, 342*(6163), 1169.

Delgado López-Cózar, E., Robinson-García, N., & Torres-Salinas, D. (2014). The Google Scholar experiment: How to index false papers and manipulate bibliometric indicators. *Journal of the Association for Information Science and Technology, 65*(3), 446-454.

Docampo, D. (2011). On using the Shanghai ranking to assess the research performance of university systems. *Scientometrics, 86*(1), 77-92.

Docampo, D. (2013). Reproducibility of the Shanghai academic ranking of world universities results. *Scientometrics, 94*(2), 567-587.

Elkana, Y., Lederberg, J., Merton, R.K., Thackray, A., & Zuckerman, H. (1978). *Toward a metric of science. The advent of science indicators.* New York: Wiley.

Elton, L. (2000). The UK Research Assessment Exercise: Unintended consequences. *Higher Education Quarterly, 54*(3), 274-283.

Etzkowitz, H. (1990). The second academic revolution: The role of the research university in economic development. In: S.E. Cozzens, P. Healey, A. Rip & J. Ziman (eds). *The research system in transition.* Springer.

Etzkowitz, H. (2000). The future of the university and the university of the future: Evolution of ivory tower to entrepreneurial paradigm. *Research Policy, 29*(2), 313-330.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: From national systems and "Mode 2" to a Triple Helix of university-industry-government relations. *Research Policy, 29*(2), 109-123.

Florian, R.V. (2007). Irreproducibility of the results of the Shanghai academic ranking of world universities. *Scientometrics, 72*(1), 25-32.

García, J.A., Rodriguez-Sanchez, R., Fdez-Valdivia, J., Robinson-García, N., & Torres-Salinas, D. (2013). Benchmarking research performance at the university level with theoretic measures. *Scientometrics, 95*(1), 438-452.

Garfield, E. (1964). Science Citation Index: A new dimension in indexing. *Science, 144*(3619), 649-654.

Garfield, E. (1998). The multiple meanings of impact factors. *Journal of the American Society for Information Science, 49*(8), 768.

Garfield, E. (2004). Historiographic mapping of knowledge domains literature. *Journal of Information Science, 30*(2), 119-145.

Garfield, E., Sher, I.H., & Torpie, R.J. (1964). *The use of citation data in writing the history of science* [Unpublished report]. Air Force Office of Scientific Research. Available at: http://www.garfield.library.upenn.edu/papers/useofcitdatawritinghistofsci.pdf

Gavel, Y., & Iselid, L. (2008). Web of Science and Scopus: A journal title overlap study. *Online Information Review, 32*(1), 8-21.

Geuna, A. (1999). *The economics of knowledge production: Funding and the structure of university research.* Edward Elgar Publishing.

Geuna, A., & Martin, B.R. (2003). University research evaluation and funding: An international comparison. *Minerva, 41*(4), 277-304.

Glänzel, W. (2011). The application of characteristic scores and scales to the evaluation and ranking of scientific journals. *Journal of Information Science, 37*(1), 40-48.

Glänzel, W., & Schoepflin, U. (1994). Little scientometrics, big scientometrics... and beyond? *Scientometrics, 30*(2-3), 375-384.

Glänzel, W., & Schubert, A. (2005). Analysing scientific networks through co-authorship. In: H.F. Moed, W. Glänzel & U. Schmoch (eds). *Handbook of Quantitative Science and Technology Research.* Dordrecht: Kluwer Academic Publishers, pp. 257-276.

Glänzel, W., Schubert, A., & Czerwon, H. (1999). An item-by-item subject classification of papers published in multidisciplinary and general journals using reference analysis. *Scientometrics, 44*(3), 427-439.

González-Pereira, B., Guerrero-Bote, V.P., & Moya-Anegón, F. (2010). A new approach to the metric of journals' scientific prestige: The SJR indicator.

Griffith, B.C., Small, H.G., Stonehill, J.A., & Dey, S. (1974). The structure of scientific literatures II: Toward a macro- and microstructure for science. *Social Studies of Science, 4*(4), 339-365.

Groh, G., & Fuchs, C. (2011). Multi-modal social networks for modeling scientific fields. *Scientometrics, 89*(2), 569-590.

Gross, P.L., & Gross, E.M. (1927). College libraries and chemical education. *Science, 66*(1713), 385-389.

Harley, D. (2013). Scholarly communication: Cultural contexts, evolving models. *Science, 342*(6154), 80-82.

Harzing, A.-W. (2013). A preliminary test of Google Scholar as a source for citation data: A longitudinal study of Nobel prize winners. *Scientometrics, 94*(3), 1057-1075.

Hazelkorn, E. (2008). Learning to live with league tables and ranking: The experience of institutional leaders. *Higher Education Policy, 21*(2), 193-215.

Hazelkorn, E. (2011). *Rankings and the reshaping of higher education. The battle for world-class excellence.* Houndmills, UK: Palgrave MacMillan.

Herranz, N., & Ruiz-Castillo, J. (2013). The end of the "European Paradox". *Scientometrics, 95*(1), 453-464.

Hicks, D. (2009). Evolving regimes of multi-university research evaluation. *Higher Education, 57*(4), 393-404.

Hicks, D. (2012). Performance-based university research funding systems. *Research Policy, 41*(2), 251-261.

Hirsch, J. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), 16569-16572.

Hjørland, B., & Albrechtsen, H. (1995). Toward a new horizon in information science: Domain-analysis. *Journal of the American Society for Information Science, 46*(6), 400-425.

Ingwersen, P., Larsen, B., & Noyons, E. (2001). Mapping national research profiles in social science disciplines. *Journal of Documentation, 57*(6), 715-740.

Jiménez-Contreras, E., Moya-Anegón, F., & Delgado López-Cózar (2003). The evolution of research activity in Spain: The impact of the National Commission for the Evaluation of Research Activity (CNEAI). *Research Policy, 32*(1), 123-142.

Jiménez-Contreras, E., Robinson-García, N., & Cabezas-Clavijo, Á. (2011). Productivity and impact of Spanish researchers: Reference thresholds within scientific areas. *Revista Española de Documentación Científica, 34*(4), 505-525.

Katz, J.S., & Martin, B.R. (1997). What is research collaboration? *Research Policy, 26*(1), 1-18.

Kessler, M.M. (1963). Bibliographic coupling between scientific papers. *American Documentation, 14*(1), 10

Kuhn, T.S. (1996). *The structure of scientific revolutions.* University of Chicago Press.

Kulkarni, A.V., Aziz, B., Shams, I., & Busse, J.W. (2009). Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. *Journal of the American Medical Association, 302*(10), 1092-1096.

Kurtz, M.J., & Bollen, J. (2010). Usage bibliometrics. *Annual Review of Information Science and Technology, 44*, 3-64.

Laredo, P. (2007). Revisiting the third mission of universities: Toward a renewed categorization of university activities? *Higher Education Policy, 20*(4), 441-456.

Lewison, G., & Dawson, G. (1998). The effect of funding on the outputs of biomedical research. *Scientometrics, 41*(1), 17-27.

Leydesdorff, L. (2008). On the normalization and visualization of author co-citation data: Salton's cosine versus the Jaccard Index. *Journal of the American Society for Information Science and Technology, 59*(1), 77-85.

Leydesdorff, L., & Bornmann, L. (2011). How fractional counting of citations affects the impact factor: Normalization in terms of differences in citation potentials among fields of science. *Journal of the American Society for Information Science and Technology, 62*(2), 217-229.

Leydesdorff, L., & Persson, O. (2010). Mapping the geography of science: Distribution patterns and networks of relations among cities and institutes. *Journal of the American Society for Information Science and Technology, 61*(8), 1622-1634.

Leydesdorff, L., Radicchi, F., Bornmann, L., Castellano, & de Nooy, W. (2013). Field-normalized impact factors (IFs): A comparison of rescaling and fractionally counted IFs. *Journal of the American Society for Information Science and Technology, 64*(11), 2299-2309.

Li, Y., Radicchi, F., Castellano, C., & Ruiz-Castillo, J. (2013). Quantitative evaluation of alternative field normalization procedures. *Journal of Informetrics, 7*(3), 746-755.

Liu, N.C., & Cheng, Y. (2005). The Academic Ranking of World Universities. *Higher Education in Europe, 30*(2), 127-136.

Liu, N.C., Cheng, Y., & Liu, L. (2005). Academic ranking of world universities using scientometrics - A comment to the "Fatal attraction". *Scientometrics, 64*(1), 101-109.

López-Herrera, A.G., Cobo, M.J., Herrera-Viedma, E., Herrera, F., Bailón-Moreno, R., & Jiménez-Contreras, E. (2009). Visualization and evolution of the scientific structure of fuzzy sets research in Spain. *Information Research, 14*(4), paper 421.

López-Illescas, C., Moya-Anegón, F., & Moed, H.F. (2011). A ranking of universities should account for differences in their disciplinary specialization. *Scientometrics, 88*(2), 563-574.

Lotka, A.J. (1926). The frequency distribution of scientific productivity. *Journal of Washington Academy Sciences, 16*, 317-323.

Lozano, G.A., Larivière, V., & Gingras, Y. (2012). The weakening relationship between the impact factor and papers' citations. *Journal of the American Society for Information Science and Technology, 63*(11), 2140-2145.

Lundberg, J. (2006). *Bibliometrics as a research assessment tool - Impact beyond the impact factor* [Doctoral thesis]. Stockholm: Karolinska Institut.

MacRoberts, M., & MacRoberts, B. (1986). Quantitative measures of communication in science - A study of the formal level. *Social Studies of Science, 16*(1), 151-172.

MacRoberts, M., & MacRoberts, B. (1987). Another test of the normative theory of citing. *Journal of the American Society for Information Science, 38*(4), 305-306.

MacRoberts, M., & MacRoberts, B. (1996). Problems of citation analysis. *Scientometrics, 36*(3), 435-444.

Manjarrés-Henríquez, L.A. (2009). *Las relaciones universidad empresa y su efecto sobre la segunda misión universitaria* [Doctoral thesis]. Valencia: Universidad Politécnica de Valencia.

Marginson, S., & van der Wende, M. (2007). To rank or to be ranked: The impact of global rankings in higher education. *Journal of Studies in International Education, 11*(3-4), 306-329.

Martin, B.R. (2011). The Research Excellence Frameword and the 'impact agenda': Are we creating a Frankenstein monster? *Research Evaluation, 20*(3), 247-254.

Meho, L.I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology, 58*(13), 2105-2125.

Merton, R.K. (1973). *The sociology of science: Theoretical and empirical investigations.* University of Chicago Press.

Moed, H.F. (2005). *Citation analysis in research evaluation.* Springer.

Moed, H.F. (2007). The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. *Science and Public Policy, 34*(8), 575-583.

Moed, H.F. (2008). UK Research Assessment Exercises: Informed judgments on research quality or quantity? *Scientometrics, 74*(1), 153-161.

Moed, H.F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics, 4*(3), 265-277.

Moed, H.F., de Bruin, R.E., Nederhof, A.J., van Raan, A.F.J., & Tijssen, R.J.W. (1992). State-of-the-art of bibliometric macro-indicators: Demand side. In: *Science and Technology in a Policy Context. Selected Proceedings of the Joint EC-Leiden Conference on Science and Technology Indicators.* Leiden, The Netherlands, pp.1-16.

Moed, H.F., de Bruin, R.E., & van Leeuwen, T.N. (1995). New bibliometric tools for the assessment of national research performance - Database description, overview of indicators and first applications. *Scientometrics, 33*(3), 381-422.

Moed, H.F., Glänzel, W., & Schmoch, U. (2004). *Handbook of quantitative science and technology research.* Dordrecht: Kluwer Academic Publishers.

Moed, H.F., Moya-Anegón, F., López-Illescas, C., & Visser, M.S. (2011). Is concentration of university research associated with better research performance? *Journal of Informetrics, 5*(4), 649-658.

Mohrman, K., Ma, W., & Baker, D. (2008). The research university in transition: The emerging global model. *Higher Education Policy, 21*(1), 5-27.

Morris, N. (2002). The developing role of departments. *Research Policy, 31*(5), 817-833.

Moya-Anegón, F., Herrero-Solana, V., & Jiménez-Contreras, E. (2006). A connectionist and multivariate approach to science maps: The SOM, clustering and MDS applied to library and information science research. *Journal of Information Science, 32*(1), 63-77.

Moya-Anegón, F., Vargas-Quesada, B., Herrero-Solana, V., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., & Muñoz-Fernández, F.J. (2004). A new technique for building maps of large scientific domains based on the cocitation of classes and categories. *Scientomtrics, 61*(1), 129-145.

Narin, F. (1976). *Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity.* Washington DC: Computer Horizons.

National Science Board (1973). *Science indicators 1972.* Washington, DC: National Science Board, National Science Foundation.

Nedeva, M. (2013). Between the global and the national: Organising European science. *Research Policy, 42*(1), 220-230.

Nedeva, M., Barker, K., & Osman, S.A. (2014). Policy pressures and the changing organization of university research. In: P.N. Teixeira (ed). *Reforming higher education.* Springer, pp. 175-188.

Newman, M.E.J. (2001). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E, 64*(1), 016131.

Noyons, E.C.M. (2005). Science maps within a science policy context. In: H.F. Moed, W. Glänzel & U. Schmoch (eds). *Handbook of quantitative science and technology research.* Dordrecht: Kluwer Academic Publishers, pp. 237-255.

Noyons, E.C.M., Moed, H.F., & Luwel, M. (1999). Combining mapping and citation analysis for evaluative bibliometric purposes: A bibliometric study. *Journal of the American Society for Information Science, 50*(2), 115-131.

Orduña-Malea, E. (2012). *Propuesta de un modelo de análisis redinformétrico multinivel para el estudio sistémico de las universidades* [Doctoral thesis]. Valencia: Universidad Politécnica de Valencia.

Ortega, J.L., Aguillo, I.F., Cothey, V., & Scharnhost, A. (2008). Maps of academic web in the European Higher Education Area - An exploration of visual web indicators. *Scientometrics, 74*(2), 295-308.

Ortega, J.L., López-Romero, E., & Fernández, E. (2011). Multivariate approach to classify research institutes according to their outputs: The case of the CSIC's institutes. *Journal of Informetrics, 5*(3), 323-332.

Pinski, G., & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. *Information Processing & Management, 12*(5), 297-312.

Pool, I.S., & Kochen, M. (1978). Contacts and influence. *Social Networks, 1*(1), 5-51.

Priem, J., & Hemminger, B.H. (2010). Scientometrics 2.0: New metric of scholarly impact on the social web. *First Monday, 15*(7). doi:10.5210/fm.v15i7.2874

Pudovkin, A., & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technoloy, 53*(13), 1113-1119.

Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences of the United States of America, 105*(45), 17268-17272.

Rafols, I., Ciarli, T., van Zwanenberg, P., & Stirling, A. (2012). Towards indicators for 'opening up' science and technology policy. In: *Proceedings of 17th International Conference on Science and Technology Indicators, 2,* pp. 663-674.

Rauhvargers, A. (2011). *Global university rankings and their impact.* Brussels: European University Association.

Robinson-García, N., & Calero-Medina, C. (2014). What do university rankings by fields rank? Exploring discrepancies between the organizational structure of universities and bibliometric classifications. *Scientometrics, 98*(3), 1955-1970.

Robinson-García, N., Torres-Salinas, D., Delgado López-Cózar, E., & Herrera, F. (2014). An insight into the importance of national university rankings in an international context: The case of the I-UGR rankings of Spanish universities. *Scientometrics,* doi: 10.1007/s11192-014-1263-1.

Rothman, R.A. (1972). A dissenting view on the scientific ethos. *The British Journal of Sociology, 23*(1), 102-108.

Rousseau, R., García-Zorita, C., & Sanz-Casado, E. (2013). The h-bubble. *Journal of Informetrics,* 7(2), 294-300.

Ruiz-Baños, R., Bailón-Moreno, R., Jiménez-Contreras, E., & Courtial, J.-P. (1999). Structure and dynamics of scientific networks. Part I: Fundamentals of the quantitative model of translation. *Scientometrics, 44*(2), 217-234.

Salmi, J., & Saroyan, A. (2007). League tables as policy instruments: Uses and misuses. *Higher Education Management and Policy, 19*(2), 31-68.

Sanz-Casado, E., García-Zorita, C., Serrano-López, A.E., Efrain-García, P., & de Filippo, D. (2013). Rankings nacionales elaborados a partir de múltiples indicadores frente a los de índices sintéticos. *Revista Española de Documentación Científica, 36*(3), e012.

Schneider, J.W. (2009). An outline of the bibliometric indicator used for performance-based funding of research institutions in Norway. *European Political Science, 8*(3), 364-378.

Schneider, J.W., & Borlund, P. (2007). Matrix comparison, Part 1: Motivation and important issues for measuring the resemblance between proximity measures or ordination results. *Journal of the American Society for Information Science and Technology, 58*(11), 1586-1595.

Schubert, A., & Braun, T. (1990). International collaboration in the sciences 1981-1985. *Scientometrics, 19*(1-2), 3-10.

Seglen, P.O. (1992). The skewness of science. *Journal of the American Society for Information Science, 43*(9), 628-638.

Shiffrin, R.M., & Börner, K. (2004). Mapping knowledge domains. *Proceedings of the National Academy of Sciences, 101*(suppl 1), 5183-5185.

Shin, J.C. (2009). Classifying higher education institutions in Korea: A performance-based approach. *Higher Education, 57*(2), 247-266.

Simons, K. (2008). The misused Impact Factor. *Science, 322*(5899), 165.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science, 24*(4), 265-269.

Small, H., & Garfield, E. (1985). The geography of science: Disciplinary and national mappings. *Journal of Information Science, 11*(4), 147-159.

Small, H., & Griffith, B.C. (1974). The structure of scientific literatures I: Identifying and graphing specialties. *Science Studies, 4*(1), 17-40.

Sokal, A.D. (1996). A physicist experiments with cultural studies. *Lingua Franca, 6*(4), 62-64.

Thelwall, M. (2008). Bibliometrics to webometrics. *Journal of Information Science, 34*(4), 605-621.

Torres-Salinas, D. (2007). *Diseño de un sistema de información y evaluación científica: Análisis cienciométricos de la actividad investigadora de la Universidad de Navarra en el área de ciencias de la salud, 1999-2005* [Doctoral thesis]. Granada: Editorial de la Universidad de Granada.

Torres, D., Cabezas, Á., & Jiménez, E. (2013). Altmetrics: New indicators for scientific communication in web 2.0. *Comunicar,* (41), 53-60.

Torres-Salinas, D., Rodriguez-Sanchez, R., Robinson-García, N., Fdez-Valdivia, J., & García, J.A. (2013). Mapping citation patterns of book chapters in the Book Citation Index. *Journal of Informetrics, 7*(2), 412-424.

Van Eck, N.J., & Waltman, L. (2009). How to normalize cooccurrence data? An analysis of some well-known similarity measures. *Journal of the American Society for Information Science and Technology, 60*(8), 1635-1651.

Van Eck, N.J., Waltman, L., Dekker, R., & van den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology, 61*(12), 2405-2416.

Van Leeuwen, T.N. (2005). Descriptive versus evaluative bibliometrics. In: H.F. Moed, W. Glänzel & U. Schmoch (eds). *Handbook of quantitative science and technology research*. Dordrecht: Kluwer Academic Publishers, pp. 373-388.

Van Leeuwen, T.N. (2007). Modelling of bibliometric approaches and importance of output verification in research performance assessment. *Research Evaluation, 16*(2), 93-105.

Van Leeuwen, T.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S., & van Raan, A.F.J. (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics, 51*(1), 335-346.

Van Noorden, R. (2010). Metrics: A profusion of measures. *Nature, 465*(7300), 864-866.

Van Raan, A.F.J. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics, 36*(3), 397-420.

Van Raan, A.F.J. (2005a). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics, 62*(1), 133-143.

Van Raan, A.F.J. (2005b). Reply to the comments of Liu et al. *Scientometrics, 64*(1), 111-112.

Van Vught, F., & Ziegele, F. (2011). *U-Multirank. The design and testing the feasibility of a multidimensional global university ranking. Final report.* Consortium for Higher Education and Research Performance Assessment, CHERPA-Network.

Van Vught, F., & Ziegele, F. (2012). *Multidimensional Ranking: The design and development of U-Multirank.* Springer.

Vargas-Quesada, B., & Moya-Anegón, F. (2007). *Visualizing the structure of science.* Springer.

Vinkler, P. (2010). *The evaluation of research by scientometric indicators.* Chandos Publishing.

Visser, M.S., Calero-Medina, C.M., & Moed, H.F. (2007). Beyond rankings: The role of large research universities in the global scientific communication system. In: *Proceedings of 11th Conference of the International Society for Scientometrics and Informetrics*, pp. 761-765.

Waltman, L., & van Eck, N.J. (2012). The inconsistency of the h-index. *Journal of the American Society for Information Science and Technology, 63*(2), 406-415.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E.C.M., Tijssen, R.J.W., van Eck, N.J., van Leeuwen, T.N., van Raan, A.F.J., Visser, M.S., & Wouters, P. (2012). The Leiden ranking 2011-2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology, 63*(12), 2419-2432.

Waltman, L., & van Eck, N.J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology, 63*(12), 2378-2392.

Watts, D.J., & Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*(6684), 440-442.

Way, M., & Ahmad, S.A. (2013). The San Francisco Declaration on Research Assessment. *Journal of Cell Science, 126*(9), 1903-1904.

White, H.D. (2003). Pathfinder networks and author cocitation analysis: A remapping of paradigmatic information scientists. *Journal of the American Society for Information Science and Technology, 54*(5), 423-434.

Williams, R., de Rassenfosse, G., Jensen, P., & Marginson, S. (2013). *U21 Ranking of national higher education systems 2013*. Melbourne Institute of Applied Economic and Social Research.

Wouters, P. (1999). *The citation culture* [Doctoral thesis]. University of Amsterdam.

Wouters, P., & Costas, R. (2012). Users, narcissism and control - Tracking the impact of scholarly publications in the 21st Century. In: *Proceedings of 17th International Conference on Science and Technology Indicators, 2*, pp. 847-857.

Yokoyama, K. (2006). The effect of the Research Assessment Exercise on organisational culture in English universities: Collegiality versus managerialism. *Tertiary Education and Management, 12*(4), 311-322.

Zitt, M. (2010). Citing-side normalization of journal impact: A robust variant of the audience factor. *Journal of Informetrics, 4*(3), 392-406.

Zitt, M., & Filliatreau, G. (2007). Big is (made) beautiful. Some comments about the Shanghai ranking of world-class universities. In: J. Sadlak & N.C. Liu (eds). *The World-Class University and Ranking: Aiming beyond status.* UNESCO-CEPES/ Cluj University Press; Shanghai: Institute of Higher Education, Shanghai Jiao Tong University. pp. 141-160.

3. First conclusions and research questions

In part I of this thesis we have started by reviewing the development of the field of bibliometrics, emphasizing its relevance for research evaluation and research policy-making. We have presented the main indicators employed and their advantages and shortcomings when used for evaluation purposes. Then, we have given an overview of their increasing use for evaluating academic institutions. For this, we have followed a historical perspective, starting with the use of bibliometric indicators by governments when developing national performance-based research funding systems. We have discussed their consequences in the higher education landscape and the challenges that must still be overcome. From a national context we have expanded our vision to the international scenario, highlighting the role of top-class university rankings as the new yardsticks for evaluating the performance of universities. We highlight two concerns with regard to the use of rankings: the threat they impose towards institutional diversity and their counter-productive effect in certain national university systems. Finally, we have focused on methodological aspects regarding the application of mapping and classification techniques applied in bibliometrics and some recent studies in which these are used to analyze universities and other research institutions.

Based on an analysis on the debates and issues reported above and public and private trans-national and national initiatives for developing bibliometric tools for assessing research policy, some tentative conclusions are drawn.

- The role of bibliometrics in research evaluation and research policy has increased over time since its inception in the 1970s (Narin, 1976; Delanghe, Sloan & Murdur, 2010). There is a perceived mistrust by research managers of peer review, the traditional evaluation system in science, turning towards quantitative measures (Hicks, 2009). Research policy-makers' demands for transparent and accountable measures (Garcia-Aracil, 2013a) have led them to focus on the use of bibliometric indicators, which seem to offer an objective solution for the measurement of research impact and productivity. This interest in the field has allowed it to expand greatly during the second half of the 20th century and the beginning of the 21st century. Also, it has led to the worrisome development of *amateur bibliometrics*, that is, sloppy practices when using bibliometric methodologies and techniques for research evaluation purposes (Bornmann et al., 2008).

As a consequence, a significant proportion of the research community has a negative perception of the use of bibliometrics in research evaluation (Anon, 2010). In order to develop a more critical perspective, the bibliometric community has actively participated in the academic debate warning against the limitations of bibliometric indicators (i.e., van Leeuwen et al., 2001; van Raan, 1995; van Raan, 1996).

- Today's higher education landscape is currently overruled by university rankings at the international level and performance-based research funding systems at the national level. Globalization has led to the emergence of competition among World-class universities for attracting funds and talent which is reflected in the great success university rankings have with research managers (Hazelkorn, 2008). On the other hand, internationalization has led to the final establishment of research as the main proxy to evaluate their performance. Research has become a priority to the detriment of teaching due to the difficulties to quantify it, as observed from the analysis of the methodology employed by rankings. This imposes serious threats in non-competitive university systems in which universities do not have sufficient autonomy to develop their own policies and compete with each other. Also, because rankings oversimplify the complex nature of universities considering them as homogeneous (Collini, 2011), they discourage institutional diversity, as those universities which have a profile which adapts better to the criteria of the ranking, position themselves better than those which do not (Bornmann, Moya-Anegón & Mutz, 2013; López-Illescas, Moya-Anegón & Moed, 2011). Finally, performance-based research funding systems seem to have reached their peak. The evaluation framework these systems present has served its original purposes reasonably well (García-Aracil, 2013; Jiménez-Contreras, Moya-Anegón & Delgado López-Cózar, 2003), but the new global scenario along with new means for scholarly communication and knowledge transfer suggest rethinking this framework (Martin, 2011).

- Ranking producers are sensitive to the criticisms their tools have received and have adopted a good disposition towards improvement. In this sense, the decision adopted by most of the main international rankings on offering rankings by fields is a good example. On their part, bibliometricians, who first dismissed university rankings due to their limitations, have ended up accepting their consolidation in research managers' agendas and have engaged in the development of more meaningful rankings (i.e., Waltman et al., 2012) and developing tools that complement the information provided by them (i.e., Bornmann et al., 2014; García et al., 2013). Also, bibliometricians have understood the need to develop goal-oriented research evaluation exercises rather than general and standard evaluations (Moed & Plume, 2011). In this sense, bibliometricians are not only exploring new variables and indicators widening the number of variables considered for the construction of bibliometric indicators and methodologies (Moed & Halevi, 2014), but also 'opening up' the outcomes of the research evaluations they perform (Rafols et al., 2012). The objective is to avoid the typical simplistic answer research managers expect of bibliometric assessments (A is improving or B has a poor performance) and offer them different options and solutions according to the aims they seek to achieve.

From the above, many questions arise especially regarding the implementation of improvements in research evaluation at the institutional level. Here we will address those related with new methodologies for identifying university profiles according to their disciplinary focus. We assume the need to preserve or, at least, respect national differences and hence focus on the evaluation of universities within a given national university system. The country used in our case study is Spain, which presents a non-competitive university system formed mainly by public universities. In the next section we will describe the Spanish university system and review efforts in the development of rankings and evaluation exercises at the national level. In this sense, the bibliometric community has long studied the effects of research policy strategies as well as involving themselves in the development of meaningful university rankings. Next, in section 3.2, we will present the specific research question of this thesis and will summarize the published articles presented in part 2.

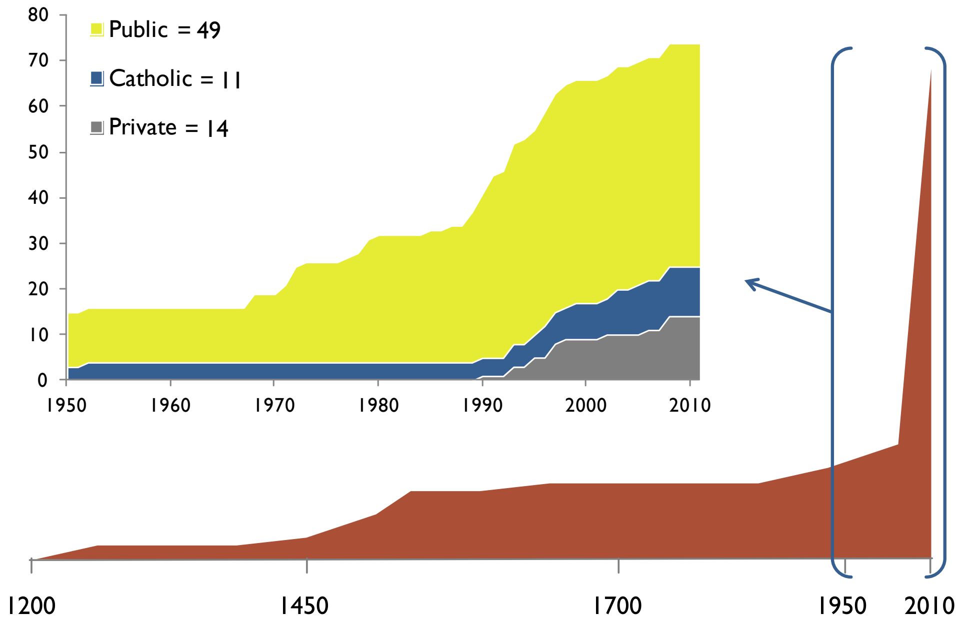
3.1. The Spanish university system as a case study

According to the latest report of the Spanish National Institute of Statistics[1], the Spanish system is formed by 74 universities. 49 of them are publicly-funded, while the other 25 are funded by private parties. In Figure 3.1, we show the chronology of the foundation of universities in Spain. We observe that there are twelve universities which were created before the 20th century. However, most of the universities were created during the last two centuries.

[1] http://www.ine.es

Within the 20th century and at the beginning of the present one, 62 universities have appeared in the scene. Regarding public universities, we observe an increase in the creation of new institutions between the 1960s and 1990s. The last two decades are characterized by an increase on the emergence of privately-funded universities.

Figure 3.1. Temporal evolution of the foundation of universities in Spain. Detail for the 1950-2011 period with types of universities. Source: Spanish National Institute of Statistics 2010-2011.



This has to do with changes and restructuring within the national legal framework which affected the university system. In the 1970s there is a shift from the Napoleonic model to one more like the German university model (for a discussion of university models see section 2.2). This transformation is reflected in the passing of the University Reform Act (*Ley de Reforma Universitaria*) in 1983. This reform had the following consequences: 1) universities became more autonomous, 2) many responsabilities were transferred to the regional governments, 3) it opened up the opportunity to create private universities and 4) it partially transferred power within the universities to non-academic staff (García-Aracil, 2007). A later reform of the system took place in 2001 (*Ley de Ordenación Universitaria*) aiming to restructure the system according to the Bologna Declaration. Among the main novelties this law introduced, we would highlight the creation of a national accreditation system as a prerrequisite to work as academic staff and for study programmes through the creation of the National Agency for Quality Assessment and Accreditation (ANECA).

Regarding the allocation of funds, regional governments are in charge of its distribution. The most common approach is that based on the number of students. However, since the 1980s some formulas are being introduced based on research performance (García-Aracil, 2007). In this decade the national government sees the need to stimulate the second mission of universities by creating voluntary individual research assessments which take place every six years (*sexenios*). If a researcher is evaluated positively, they have a slight increase in their salary (Jiménez-Contreras, Moya-Anegón & Delgado López-Cózar, 2003). At the national level, formulas such as the *Campus de Excelencia*, which was introduced at the end of the 2000s, aim to promote competition and establish ties between universities with common interests (Docampo et al., 2012). These efforts are intended to introduce competitive elements within the system, although they do not involve significant structural changes.

Still, the Spanish university system can be considered non-competitive: universities do not benefit from levels of autonomy sufficient to establish their own policies directed at improving their research performance. According to a recent report commissioned by the European University Association, the autonomy of Spanish universities is regarded as medium-low at the organizational, financial, staffing and academic levels (Estermann, Nokkola & Steinel, 2011). For instance, regarding organizational autonomy, they are considered to be 'severely limited' in the appointment of external members to their governing bodies. With respect to funding and academic autonomy, tuition fees are set by external authorities and programmes must also be negotiated with government. But the most troublesome issue is staffing autonomy, where civil servant status limits salaries, and promotions are heavily regulated by external authorities. This lack of competition and autonomy at the institutional level leads to an inconsistent institutionalization of university missions (García-Aracil, 2013b) which holds the country back when compared with other national systems (Aghion, et al., 2010).

The emergence and popularity of rankings however, seems to have had a positive effect in the sense that it has alerted research managers to how these tools affect their institutional image (Bornmann, 2014). Rapidly, and due to the poor presence of Spanish universities in international rankings, national rankings have been developed to analyze the state of the national university system (Torres-Salinas et al., 2011a). Unlike the first international rankings that emerged (see section 2.3), in this case, the initiative for developing these tools has come from the research community. This means that special care has been taken to overcome the limitations present in the former rankings. For instance, Buela-Casal and colleagues (2009;2010;2011) developed rankings of public universities considering input indicators (i.e., tenured staff, number of PhDs) and other types of output indicators as well as publication data (i.e., number of patents) to construct a research performance-based ranking. Torres-Salinas and colleagues (2011ab) develop rankings by fields based on bibliometric data from the Web of Science aggregating subject categories. Finally, others have focused on analyzing and thoroughly understanding the methodology employed by international rankings in order to predict and analyze the factors that affect the positioning of Spanish universities (Docampo & Torres-Salinas, 2013).

An interesting perspective is the one adopted by Sanz-Casado and colleagues (2013) who have made a remarkable effort to retrieve not only bibliometric data, but also data such as tenured staff, number of *sexenios*, national prizes, patents, spin-offs, etc. of all Spanish universities. In this case, the intention is not to aggregate these variables into a global indicator but offer all of them separately in order to allow further analyses to improve our understanding of the national system. An example of this is the study conducted by Filippo and colleagues (2012) to analyze the effect of an alliance between four Spanish universities on their performance. This perspective introduces a much wider framework that goes beyond bibliometric data by introducing other research outputs and broadening the number variables considered.

In this line of thought, others have attempted to analyze variables or characteristics that may influence the performance of some Spanish universities over others. Luque-Martínez (2013) analyzes whether issues such as institutional age or size may affect university performance. Bordons and colleagues (2010) go a step further and offer a methodology for profiling universities according to these characteristics among others. These types of studies are ultimately directed at identifying the institutional factors that favor efficiency at the institutional level, profiling and classifying universities according to these factors. However, regarding the disciplinary focus of universities, no study has been found analyzing university profiles in Spain. This thesis intends to fill this gap by developing methodologies for classifying and visualizing the disciplinary profile of universities in Spain.

3.2. Research questions and summary of publications

This thesis is built on the need for national university rankings to complement international ones and the importance of analyzing institutional profiles when assessing the performance of universities. These issues are presented in the following two research questions.

RESEARCH QUESTION 1. Are national university rankings necessary in an international context?

The idea of using national university rankings stems from the worrisome use research managers make of international rankings to evaluate national systems. International rankings try to show league tables of world-class universities. World-class universities are the result of an 'Emerging Global University Model' consisting of a small group of universities which present a very different profile to the rest as they are not necessarily subjected to national barriers (Mohrman, Ma & Baker, 2008). Hence, using these tools for decision-making may negatively affect non-competitive university systems such as the Spanish one.

In chapter 4 we suggest the use of national rankings by fields as a complement to international rankings in order to provide the full picture of the Spanish national system. In order to overcome the limitations of rankings with regard to the neglect of disciplinary differences, we suggest the use of rankings by fields, disregarding global league tables. Also, we consider that only an analysis of the information provided by all main international rankings can lead to an accurate picture in the research performance of Spanish universities (Robinson-García, Jiménez-Contreras & Delgado López-Cózar, 2013). Hence, we compare the results of the I-UGR Rankings of Spanish universities by scientific fields and disciplines with those of the Leiden Ranking, QS Rankings, Shanghai Ranking and NTU Ranking. We conclude by stressing the role of national university rankings especially in non-competitive higher education systems and the need to acknowledge methodological differences between rankings. These differences are derived from 1) the selection of the data source, 2) the calculation of the global indicator and 3) the construction of fields and disciplines.

RESEARCH QUESTION 2. How can we develop tools for research policy-makers that can allow them to identify disciplinary profiles?

Although rankings by fields seem to be a plausible solution to analyze the research performance of universities whilst avoiding specialization differences between universities, they still seem to have some limitations. These limitations are discussed in chapter 5 and affect any bibliometric study which is built on a subject category classification system. They are derived from the misconception of the unit of analysis used to build the classification system. In the case of rankings by fields, readers usually believe that these fields are based on the specialization of researchers or the organizational structure of universities. However, in most cases these rankings are built on journal classifications such as those provided by Thomson Reuters. This has to do with the impossibility of adopting a *bottom up* approach in the data retrieval process.

This leads us to turn our attention to science mapping and classification techniques which have proven to be useful tools for research policy (Noyons, 2005). Chapters 6 and 7 are methodological studies in which different techniques are developed and tested. In chapter 6 we introduce Biplot analysis for visualizing multivariate data. We provide three case studies as examples of potential use all of them related to research evaluation of universities.

In chapter 7 we introduce the journal publication profile methodology for establishing relations between universities based on the co-occurrence of journals in which they publish. This methodology allows us to identify collaboration and disciplinary similarities. Based on this methodology, in chapter 8 we attempt to describe the Spanish national university system using network analysis. We conclude by developing a classification of four types of Spanish universities based on their specialization, research intensity and publication strategy.

References

Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A., & Sapir, A. (2010). The governance of universities: Evidence from Europe and the US. *Economic Policy, 25*(61), 7-59.

Anon (2010). Assessing assessment. *Nature, 465*(7300), 845.

Bordons, M., Sancho, R., Morilla, F., & Gómez, I. (2010). Perfil de actividad científica de las universidades españolas en cuatro áreas temáticas: Un enfoque multifactorial. *Revista Española de Documentación Científica, 33*(1), 9-33.

Bornmann (2014). On the function of university rankings. *Journal of the Association for Information Science and Technology,65*(2), 428-429.

Bornmann, L., Moya-Anegón, F., & Mutz, R. (2013). Do universities or research institutions with a specific subject profile have an advantage or a disadvantage in institutional rankings? *Journal of the American Society for Information Science and Technology, 64*(11), 2310-2316.

Bornmann, L., Mutz, R., Neuhaus, C., & Daniel, H.-D. (2008). Citations counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics, 8*(1), 93-102.

Bornmann, L., Stefaner, M., Moya-Anegón, F., & Mutz, R. (2014). Ranking and mapping of universities and research-focused institutions worldwide based on highly-cited papers: A visualisation of results from multi-level models. *Online Information Review, 38*(1), 43-58.

Buela-Casal, G., Bermúdez, M.P., Sierra, J.C., Quevedo-Blasco, R., & Castro, A. (2009). Ranking de 2008 en productividad en investigación de las universidades públicas españolas. *Psicothema, 21*(2), 304-312.

Buela-Casal, G., Bermúdez, M.P., Sierra, J.C., Quevedo-Blasco, R., & Castro, A. (2010). Ranking de 2009 en investigación de las universidades públicas. *Psicothema, 22*(2), 171-179.

Buela-Casal, G., Bermúdez, M.P., Sierra, J.C., Quevedo-Blasco, R., Castro, A., & Guillén-Riquelme, A. (2011). Ranking en producción y productividad en investigación de las universidades públicas españolas. *Psicothema, 23*(4), 527-536.

Collini, S. (2012). *What are universities for?* London: Penguin Books.

Delanghe, H., Sloan, B., & Muldur, U. (2010). European research policy and bibliometric indicators, 1990-2005. *Scientometrics, 87*(2), 389-398.

Docampo, D., Herrera, F., Luque-Martínez, T., & Torres-Salinas, D. (2012). Efecto de agregación de universidades españolas en el Ranking de Shanghai (ARWU): Caso de las comunidades autónomas y los campus de excelencia. *El profesional de la información, 21*(4), 428-432.

Docampo, D., & Torres-Salinas, D. (2013). La nueva lista de investigadores altamente citados de Thomson Reuters y el Ranking Shanghai: Situación de España y mapa universitario. *El profesional de la información, 22*(3), 264-272.

Estermann, T., Nokkala, T., & Steinel, M. (2011). *University autonomy in Europe II. The scorecard* [Technical report]. European University Association

Filippo, D., Casani, F., García-Zorita, C., Efraín-García, P., & Sanz-Casado, E. (2012). Visibility in international rakings. Strategies for enhancing the competitiveness of Spanish universities. *Scientometrics, 93*(3), 949-966.

García, J.A., Rodríguez-Sánchez, R., Fdez-Valdivia, J., & Robinson-García, N., & Torres-Salinas, D. (2013). Benchmarking research performance at the university level with information theoretic measures. *Scientometrics, 95*(1), 435-452.

García-Aracil, A. (2007). Expansion and reorganization in the Spanish higher education system. In: A. Bonaccorsi & C. Daraio (eds). *Universities and strategic knowledge* *creation: Specialization and performance in Europe.* Edward Elgar Publishing, pp. 376-402.

García-Aracil, A. (2013a). Understanding productivity changes in public universities: Evidence from Spain. *Research Evaluation, 22*(5), 351-368.

García-Aracil, A. (2013b). What is it that differentiates the universities in Spain? A qualitative perspective. *Regional and Sectoral Economic Studies, 13*(3), 61-76.

Hazelkorn, E. (2008). Learning to live with league tables and ranking: The experience of institutional leaders. *Higher Education Policy, 21*(2), 193-215.

Hicks, D. (2009). Evolving regimes of multi-university research evaluation. *Higher Education, 57*(4), 393-404.

Jiménez-Contreras, E., Moya-Anegón, F., & Delgado López-Cózar, E. (2003). The evolution of research activity in Spain: The impact of the National Commission for the Evaluation of Research Activity (CNEAI). *Research Policy, 32*(1), 123-142.

López-Illescas, C., Moya-Anegón, F., & Moed, H.F. (211). A ranking of universities should account for differences in their disciplinary specialization. *Scientometrics, 88*(2), 563-574.

Luque-Martínez, T. (2013). La actividad investigadora de la universidad española en la primera década del siglo XXI: La importancia del tamaño de la universidad. *Revista Española de Documentación Científica, 36*(4), e026.

Martin, B.R. (2011). The Research Excellence Framework and the 'impact agenda': are we creating a Frankenstein monster? *Research Policy, 20*(3), 247-254.

Moed, H.F., & Halevi, G. (2014). Research assessment: Review of methodologies and approaches. *Research Trends, 36*, 3-5.

Moed, H.F., & Plume, A. (2011). The multi-dimensional research assessment matrix. *Research Trends, 23*, 5-7.

Mohrman, K., Ma, W., & Baker, D. (2008). The research university in transition: The emerging global model. *Higher Education Policy, 21*(1), 5-27.

Narin, F. (1976). *Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity.* Washington DC: Computer Horizons.

Noyons, E.C.M. (2005). Science maps within a science policy context. In: H.F. Moed, W. Glänzel & U. Schmoch (eds). *Handbook of quantitative science and technology research.* Dordrecht: Kluwer Academic Publishers, pp. 237-255.

Rafols, I., Ciarli, T., van Zwanenberg, P., & Stirling, A. (2012). Towards indicators for 'opening up' science and technology policy. In: *Proceedings of 17th International Conference on Science and Technology Indicators, 2,* pp. 663-674.

Robinson-García, N., Jiménez-Contreras, E., & Delgado López-Cózar, E. (2013). Las universidades españolas en los principales rankings de universidades. 2013. *EC3 Working Papers, 14*. Available at: http://hdl.handle.net/10481/29083

Sanz-Casado, E., García-Zorita, C., Serrano-López, A.E., Efraín-García, P., & Filippo, D. (2013). Rankings nacionales elaborados a partir de múltiples indicadores frente a los de índices sintéticos. *Revista Española de Documentación Científica, 36*(3), e012.

Torres-Salinas, D., Delgado López-Cózar, E., Moreno-Torres, J.G., & Herrera, F. (2011a). Rankings ISI de las universidades españolas según campos científicos: Descripción y resultados. *El profesional de la información, 20*(1), 111-122.

Torres-Salinas, D., Moreno-Torres, J.G., Robinson-García, N., Delgado López-Cózar, E., & Herrera, F. (2011b). Rankings ISI de las universidades españolas según campos y disciplinas científicas. *El profesional de la información, 20*(6), 701-711.

Van Leeuwen, T.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S., & van Raan, A.F.J. (2001). Language biases in the coverage of th Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics, 51*(1), 335-346.

Van Raan, A.F.J. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics, 36*(3), 397-420.

Van Raan, A.F.J. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics, 62*(1), 133-143.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E.C.M., Tijssen, R.J.W., van Eck, N.J., van Leeuwen, T.N., van Raan, A.F.J., Visser, M.S., & Wouters, P. (2012). The Leiden ranking 2011/212:Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology, 63*(12), 2419-2432.

PART 2. Publications

4. An insight into the importance of national university rankings in an international context: The case of the I-UGR Rankings of Spanish Universities

Paper published in Scientometrics doi:10.1007/s11192-014-1263-1 authored by N. Robinson-García, D. Torres-Salinas, E. Delgado López-Cózar and F. Herrera

The great importance international rankings have achieved in the research policy arena warns against many threats consequence of the flaws and shortcomings these tools present. One of them has to do with the inability to accurately represent national university systems as their original purpose is only to rank world-class universities. Another one has to do with the lack of representativeness of universities' disciplinary profiles as they usually provide a unique table. Although some rankings offer a great coverage and others offer league tables by fields, no international ranking does both. In order to surpass such limitation from a research policy viewpoint, this paper analyzes the possibility of using national rankings in order to complement international rankings. For this, we analyze the Spanish university system as a study case presenting the I-UGR Rankings for Spanish universities by fields and subfields. Then, we compare their results with those obtained by the Shanghai Ranking, the QS Ranking, the Leiden Ranking and the NTU Ranking, as they all have basic common grounds which allow such comparison. We conclude that it is advisable to use national rankings in order to complement international rankings, however we observe that this must be done with certain caution as they differ on the methodology employed as well as on the construction of the fields.

4.1. Introduction

Since the launch of the first edition of the Shanghai Ranking in 2003, interest has grown on the development of tools for benchmarking and comparing academic and research institutions. As a result of the massification of higher education, the race for excellence and a fierce battle for research funding, universities now strive for positioning themselves in these international rankings (Hazelkorn 2011). These tools have gained an undisputable position in research managers' 'toolkit' for measuring the state of health of higher education institutions and the main resource for many universities and countries when taking decisions in a research policy context (Marginson & van der Wende, 2007). The great effect they have- not only in the media and the public but also for research managers, politicians and decision makers - relies on the perception that highly ranked institutions are usually more productive, produce higher quality research and teaching and contribute best to society than the rest of universities (Shin & Toutkoushian, 2011).

However, despite their advantages as easy-to-read tools, they also have many inconsistencies and shortcomings that warn against a careless use (Delgado López-Cózar, 2012). In this sense, we can identify five major issues which must be addressed: 1) methodological and technical errors and difficulties such as the recollection of reliable and standardized data (Toutkoushian & Webber, 2011); 2) the criteria for selecting the indicators are not scientifically supported (Van Raan, 2005); 3) the multidimensional nature of universities (Orduña-Malea, 2012; Waltman et al., 2012) leads to a wide heterogeneity among institutions (Collini 2011); 4) using a unique table to rank universities neglects their disciplinary focus (Visser et al., 2007); and 5) international rankings cannot reflect the state of national higher education systems as they usually cover just the top universities of each country (Torres-Salinas et al., 2011a).

While the issue of data reliability still remains a major shortcoming and there is no consensus yet on which indicators represent better the nature and quality of universities, the other issues have been somehow surpassed using approaches which do not solve completely their dangers but, at least, diminish the flaws. For instance, rankings such as the Leiden Ranking (Waltman et al., 2012) or the Scimago Institutions Rankings (henceforth SIR) have emerged focusing uniquely on the research dimension of universities to the neglect of other aspects such as innovation or teaching. Others now publish, along with a global ranking, rankings by subjects and fields, which offer a better picture of universities' performance (García et al, 2012). Also, some rankings such as the SIR or the Ranking Web of World Universities cover now not just top-class universities but the former includes more than 3,000 research institutions and the latter, more than 19,000.

Rankings have not been fully developed and still have serious shortcomings (van Raan, 2005). But their dominance as decisive factors in research policy (Hazelkorn, 2011) at national and supranational level puts them in the spotlight. One of the most important threats rankings entail is that they ignore universities' diversity, which can affect seriously the health of higher education systems and lead to dangerous and simplistic conclusions when interpreting and developing ranking systems (e.g., Moed et al., 2011). These differences affect institutions at two levels, at their organizational structure, and in the national configuration of higher education systems, affecting their multidisciplinary nature and diversity (Orduña, 2011). The phenomenon of university rankings has influenced deeply all university systems, even those that were not conceived at first to establish a competitive framework. Therefore, in order to analyze the success or failure of different countries in their research policy, university systems should be assessed as a whole, and not considering each university as an individual and autonomous unit. Such approach was applied by Docampo (2011) using the Shanghai Ranking in order to analyze the university systems of the countries represented.

Despite its limitations, this study offers a glimpse of the global scenario regarding the research excellence of different countries' university systems. In Table 4.1 we show the clusters emerged from the study carried out by Docampo (2011) and the number of universities by country in different intervals according to the 2012 edition of the Shanghai Ranking. Therefore we observe a dominance of the United States and the United Kingdom which alone represent more than a third of the universities included in the ranking (37.6%), followed by Germany and Canada as the next with the highest number of universities included. However, despite the numbers, except Japan, which in this new edition includes a university in the top20, none of the others have a university positioned within this interval.In this context, the truth is that the high visibility Anglo-Saxon universities have in rankings leaves little space for others, blurring the state of other countries which are working towards a successful university model. In fact, it clearly shows the incapability of the ranking to represent national university systems with exhaustiveness.

Thus, these rankings do not offer a complete view of national higher education systems, preventing research managers and decision makers to have an accurate picture of the state of each country's university system. Hence the need for developing tools with higher levels of granularity in the information provided by rankings (Bornmann, Mutz & Daniel, 2013). For this reason, in 2010 members from the EC3 and Soft Computing research groups developed the Rankings I-UGR of Spanish Universities according to Fields and Scientific Disciplines (henceforth I-UGR Rankings) available at http://rankinguniversidades.es.It was originally named ISI Rankings but changed to its current name in its 2012 edition. This website offers 49 rankings for Spanish universities divided in 12 fields and 37 disciplines, according to their international research performance. Spain is a good example of a misrepresented higher education system. For instance, in the 2013 edition of the Shanghai Ranking only 10 universities out of 74 met the criteria for inclusion in the global ranking. In fact, none made it to the top 100 and only four were included in the 201-300 interval. Also, as it occurs with other countries such as Italy (Abramo, Cicero & D'Angelo, 2011), it is a non-competitive higher education system, which means that universities do not act as individual units but within a national framework, therefore decisions should not be taken relying on such a poor sample.

Table 4.1 University systems by country considering the results in Docampo (2011) and the 2012 Shanghai Ranking edition. Leaders, Fast followers and followers

	Countries	Nr of Universities Top20	Nr of Universities Top100	Nr of Universities Top300	Nr of Universities Top500
Leaders	United States	17	53	109	150
	United Kingdom	2	9	30	38
	Switzerland	---	4	7	7
Fast followers	Australia	---	5	9	19
	Canada	---	4	17	22
	Sweden	---	3	7	11
	Israel	---	3	4	6
	Netherlands	---	2	10	13
	Denmark	---	2	4	4
Followers	Germany	---	4	24	37
	France	---	3	13	20
	Belgium	---	1	6	7
	Norway	---	1	3	4
	Finland	---	1	1	5

The main goal of the present paper is analyze if national rankings are necessary complements to international rankings. This paper is focused at the potential use of the information provided by national and international rankings by research managers and intends to explore if the information provided by both types of rankings is complementary and useful from a research management perspective. For this we will use the I-UGR Rankings, in order to:

1) Analyze if national ranking are necessary complementing the information provided by international rankings, as the latter do not represent well national university systems.

2) Analyze the levels of agreement between national and international rankings regarding the following aspects:

a. Are the top Spanish universities the ones visible in international rankings?

b. Disciplinary concordance: Do the different classifications by fields and subjects allow an analysis by areas?

To develop this study we select the Shanghai Ranking, the Times Higher Education World University Ranking (henceforth THE Ranking), the QS Ranking, the Leiden Ranking and the National Taiwan University Ranking (henceforth NTU Ranking). The first one to include disciplinary-oriented league tables was the Shanghai Ranking, launching in 2007 rankings by five broad fields and in 2009 five more rankings in specific disciplines, followed by the THE Ranking, the QS Rankings and the NTU Ranking. The Leiden Ranking has been the last one to follow this trend and now includes in its last edition rankings by five broad areas.

The paper is structured as follows. In section 4.2 we present the Spanish case analyzing its current state and we introduce the I-UGR Rankings, we contextualize its creation and we describe the methodology employed for their development. In section 4.3 we address the main issue of this paper: we compare the results of the main international rankings and the I-UGR Rankings for Spanish universities. Finally, in Section 4.4 we resume our main findings and their consequences in a research policy scenario.

4.2. Spain as a case study: introduction to the I-UGR Rankings

The Spanish university system is formed by 74 universities: 48 public and 26 private. However in the 2013 edition of the Shanghai Ranking only 10 met the minimum requirements to be included. It is a country poorly represented in the main international rankings due to the scarce number of universities considered as World-Class universities. But the impact these rankings have in research policy threatens a good governance and sensible decision making as they do not offer a complete picture of the university system (Docampo, 2011). In fact, as observed in Table 4.2, only 19 universities (18 public and 1 private universities) are included in four of the most important rankings; that is, 25.68% of the whole system. For this reason, other tools are needed in order to complete this fragmented picture of the Spanish higher education scenario.

The first edition of the I-UGR Rankings was launched in 2010. Its development was motivated by the scarce visibility Spanish universities have in international rankings, which leads to a fragmented picture of the Spanish university system. Though other national rankings had already been developed, these were considered insufficient due to the limitations they presented which made them unsuitable as research policy tools. Among other limitations we address the following: lack of continuity over time, exclusion of private institutions, disregard of disciplinary focus, use of rudimentary bibliometric indicators, selection of unsuitable time periods or election of databases with dubious selection criteria of sources (Torres-Salinas et al., 2011a).

Data is retrieved from the Thomson Reuters Science Citation Index and Social Science Citation Index (SCI and SSCI).The reason for using such source database relies not only on its importance as a bibliometric database containing the main international scientific literature, but also due to its importance in the Spanish research evaluation system (Cabezas-Clavijo et al., 2013). In its first edition 12 rankings were offered for 12 broad fields. These fields were later expanded with 19 subfields or disciplines in the second edition (Torres-Salinas et al., 2011b) and finally, 37 disciplines in the 2012 edition. The fields and disciplines were constructed by aggregating the subject categories to which records from the Journal Citation Reports are assigned. Aggregating subject categories is a classical perspective followed in many bibliometric studies when adopting a macro-level approach (e.g., Moed, 2005; Leydesdorff & Rafols, 2009). For further information on the coverage of the I-UGR Rankings and the development of the fields and subfields the reader is referred to the following document in which methodology of the indicator for ranking universities as well as the construction of fields are defined[1].

[1]http://sci2s.ugr.es/rankinguniversidades/downloads/rankingsI-UGR_Methodology_EV.pdf

Table 4.2 Spanish universities represented within the top 500 universities in the 2013 edition of the Shanghai Ranking, the QS Ranking, the Leiden Ranking and the NTU Ranking

Shanghai Ranking		**Leiden Ranking**		**QS Ranking**		**NTU Ranking**	
Barcelona	201-300	Barcelona	259	Aut Barcelona	177	Barcelona	89
Aut Madrid	201-300	Pol Valencia	282	Barcelona	178	Aut Barcelona	169
Aut Barcelona	201-300	Santiago	317	Aut Madrid	195	Aut Madrid	214
Complutense	201-400	Aut Barcelona	333	Complutense	216	Valencia	224
Pol Valencia	301-400	Valencia	336	Pompeu Fabra	281	Complutense	259
Valencia	301-400	Aut Madrid	356	Navarra	315	Granada	267
Pompeu Fabra	301-400	Zaragoza	366	Carlos III	317	Oviedo	369
Granada	301-400	Granada	375	Pol Cataluña	345	Santiago	378
Zaragoza	401-500	Pol Cataluña	396	Pol Valencia	383	Zaragoza	392
País Vasco	401-500	Sevilla	402	Pol Madrid	389	País Vasco	421
		Complutense	406	Salamanca	441-450	Sevilla	434
		Murcia	408	Valencia	471-480	Pol Valencia	446
		Oviedo	409	Zaragoza	481-490	Pompeu Fabra	463
		País Vasco	411				
		Pol Madrid	434				

List of abbreviations used: Aut: Autónoma; Pol: Politécnica

Once the data is compiled into a relational database, the indicators defined in Table 4.3 are computed, and the index for rating each university is calculated. To rank universities we use the IFQ^{2}A Index (Torres-Salinas et al., 2011c). This indicator measures the quantitative and qualitative dimensions of the research outcome of a group of institutions in a given field. It is based on six primary bibliometric indicators, three focused on the quantitative dimension (QNIF) and the other three focused on the qualitative dimension (QLIF). These two dimensions represent two different aspects of the research activity, impact and visibility of universities. While the QNIF is based on size-dependent measures, the QLIF relies on relative measures of impact (as defined by the citations received) and visibility (as defined by the quartile to which a journal belongs according to its Impact Factor and the top papers among the 10% most cited papers). QLIF is a no size-dependent measure. In Table 3 we summarize the methodology employed for calculating the IFQ^{2}A Index. More information about the IFQ^{2}A Index may be found in Torres-Salinas et al. (2011c).

Table 4.3 Calculation of the IFQ2A Index and definition of indicators.

$QNIF = \sqrt[3]{NDOC \times NCIT \times H}$		$QLIF = \sqrt[3]{\%1Q \times ACIT \times TOPCIT}$	
NDOC	Number of citable papers published in scientific journals	%1Q	Ratio of papers published in journals in the top JCR quartile
NCIT	Number of citations received by all citable papers	ACIT	Average number of citations received by all citable papers
H	H-Index as proposed by Hirsch (2005), over all the publications of the institution	TOPCIT	Ratio of papers belonging to the top 10% most cited papers calculated within all institutions
$IFQ^2A = QNIF \times QLIF$			

The selection of the indicators as well as the conceptualization of the index, are based on the following criteria:

1) The indicators chosen must not be restrictive. That is, they should be applied to all institutions. For instance, the Shanghai Ranking uses the number of Nobel Prizes as an indicator to measure research excellence. In the Spanish case only one university is affected by it (Complutense de Madrid).

2) Rankings must be size-independent, however if the numbers are too small they may distort the ranking and introduce a certain degree of instability. This leads to the use of a bidimensional index which takes into account raw counts of papers and citations as well as relative measures which benefit small

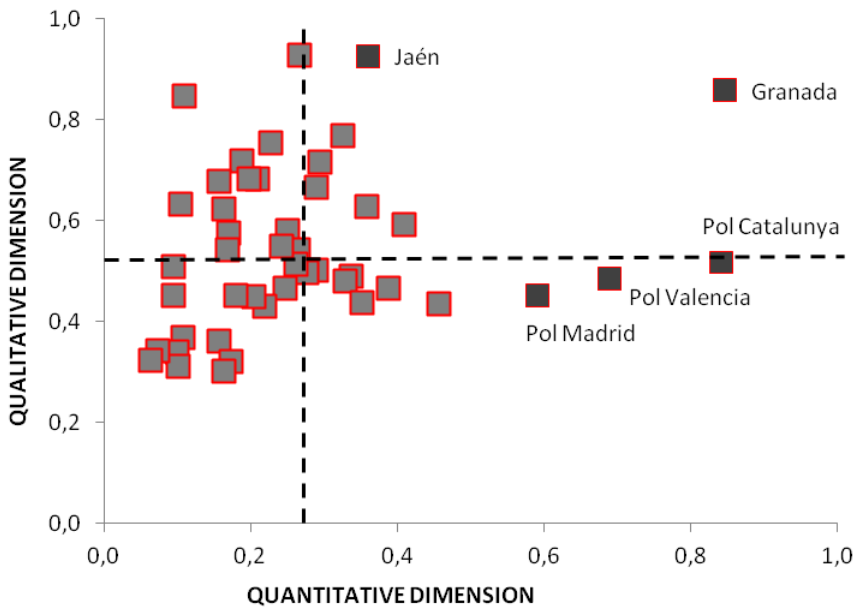
institutions which produce high quality papers (as defined by bibliometric indicators).

3) Rankings must take into account the disciplinary focus of universities. For this, a unique list cannot be provided. Contrarily, one must offer rankings by field of specialization in order to provide useful tools for research managers.

4) Seniority must not be rewarded. For this, fixed time periods must be used. Also, when calculating the H-Index, this must be considering the time frame used. In this sense, the I-UGR Rankings offer a five-year window and a ten-year window.

5) Stability must be assured. This means that the fixed time frame must be wide enough to offer stable results. A five-year time frame allows results to be consistent and significant.

Figure 4.1 Distribution of universities according to their qualitative and quantitative dimensions in the field of Computer Science. 2008-2012. Top 5 institutions according to the IFQ^{2}A Index are highlighted and labeled.



In Figure 4.1 we show the distribution of universities according to the QNIF and QLIF in the field of Computer Science for the 2008-2012 time period. The dashed lines show the average values of each dimension. Universities positioned at the top right hand of the figure are those which outstand in both dimensions. Those positioned on the bottom right outstand on the quantitative dimension but not on the qualitative dimension. At the top left, we observe university with small research output but high quality research. Lastly, in the bottom left, universities which do not outstand in any dimension are represented. As we can observe, although top universities outstand in both dimensions, many universities outstand in the qualitative dimension but do not do so in the quantitative dimension. Due to the bidimensional nature of the IFQ^2A index, these small institutions are reflected in the rankings.

4.3. Levels of agreement and disciplinary concordance between rankings: Comparison by fields of the main international rankings and the I-UGR rankings

In this section we analyze the state of the Spanish university system using international and national rankings. For this, we first establish in Section 4.3.1 a set of criteria for the selection of the rankings we will use in order to set some basic common grounds which will allow a fair comparison between them. Then, in Section 4.3.2 we match rankings by fields between the international and national rankings and finally, we analyze the level of agreement between them. For this we use two indicators. On the one hand, we calculate the Spearman's rank correlation coefficient or Spearman's rho, which will indicate to what extent are the different rankings coherent between them on the order in which Spanish universities are displayed. On the other hand, we show the level of agreement between rankings, which indicates if universities included in an international ranking coincide with those which occupy the top positions of the national ranking.

4.3.1. Selection of international rankings

The aim is to use international and national rankings as complementary tools to offer on the one hand, a global perspective of the position of Spanish universities and, on the other hand, a complete picture of the Spanish university system. For this, we first need to establish a set of criteria for choosing the most relevant rankings for our purposes. These are the following:

1) Rankings must be based on the research performance of universities, at least partially, as we are analyzing the research dimension of universities.

2)Data retrieved for the construction of the rankings must come from a reliable bibliometric database or information resource, at least partially.

3) They must offer rankings by fields, as we have considered that only this way we can provide an accurate image of universities' research performance.

Based on these criteria we selected the I-UGR Rankings as national rankings and the following international rankings. In table 4.4 we include the main characteristics of each of these rankings. For more detailed information on the methodology of each ranking, we refer the reader to its website; here we will briefly describe them:

1) Shanghai Ranking (http://www.shanghairanking.com/). It was not only the first international ranking launched (Liu & Cheng, 2005) but it is used as yardstick to measure the research excellence of universities worldwide (Docampo, 2011). It is based on six indicators, two of them (40% of the total rating) are based on data retrieved from the Web of Science (for more information on this ranking the reader is referred to Liu & Cheng, 2005; van Raan, 2005; Docampo 2011; Aguillo et al., 2010). Since 2007 it offers five rankings by field and since 2009, five rankings by subject.

2) *QS Ranking* (http://www.topuniversities.com/). The first edition of this ranking was launched in 2004. Until 2009 it was produced in partnership with the Times Higher Education, however, since then each company develops its own ranking (for more information on this ranking the reader is referred to Aguillo et al., 2010; Usher & Savino, 2007). 20% of the total rating assigned to each university is based on data retrieved from the database Scopus. It offers along with the global league table, 29 rankings by discipline classified into five major fields.

3) *NTU Ranking* (http://nturanking.lis.ntu.edu.tw). This ranking was first launched in 2007.It aims at measuring solely the quality of universities' research. It is based on 8 indicators all of them supported by bibliometric data from the Web of Science and the Thomson Reuters Essential Science Indicators(for more information on this ranking the reader is referred to e.g., Aguillo et al., 2010). Along with the global table league, it offers rankings by field and subject in a similar structure to that of the Shanghai Ranking. In this case, it offers 6 rankings by field and 14 rankings by subject.

4) *Leiden Ranking* (http://leidenranking.com). The first version of the Leiden Ranking was published in 2008[2]. However it was discontinued and, despite a 2010 edition was announced it is no longer available. In 2012 they resumed their activity and is now updated on an annual basis. Its methodology, shortcomings and potential use are discussed by Waltman et al. (2013). In its latest edition, it includes for the first time, rankings by five broad fields. These fields are constructed based on aggregations of the Web of Science subject categories. In the case a journal is assigned to several fields, its publications are assigned fractionally. The assignment of subject categories is available at the Methodology section of their website.

At this point it is important to note that the THE Rankings are not included in this study. Although they meet the criteria we do not include them for technical reasons. Only four Spanish universities are included in three of their six rankings by fields. Such a low presence does not allow its analysis and comparison with the national ranking. Also the Scimago Institutions Rankings are missing from this analysis. This is because they do not provide rankings by fields in their last edition.

4.3.2. Concordance between international and national rankings and levels of agreement

In order to establish fair comparisons and provide a global picture of the state of Spanish universities using national and international rankings, we first need to ensure that the classification of fields of national and international rankings is somehow similar and therefore, compatible. For this, we would need to analyze the way these fields are constructed for the four rankings used in this study and determine to which grade the methodology employed by each of them allows fair comparisons. As mentioned before, the I-UGR Rankings construct fields and disciplines by aggregating the Thomson Reuters subject categories. The Leiden Ranking and the NTU Ranking use the same approach, and the construction of fields and subjects is declared at their website. However, this does not occur for the other two rankings, which do not declare the methodology employed for establishing such fields. This lack of transparency is a shortcoming that must be taken into account when using these rankings for research policy.

[2]Available at http://www.cwts.nl/ranking/LeidenRankingWebSite.html

Table 4.4 Main characteristics of the Shanghai Ranking, QS Ranking, NTU Ranking and Leiden Ranking

Ranking	**Shanghai Ranking**	**QS Ranking**	**NTU Ranking**	**Leiden Ranking**
Launch year	2003	2004*	2007	2008
1st edition with fields	2007	2009	2007	2013
No. of fields	5	5	6	5
No. of subjects	5	29	14	0
Total universities	500	+701	500	500
Type of data	Bibliometric and reputational	Bibliometric, surveys and manpower	Bibliometric	Bibliometric
Bibliometric data sources	Web of Science	Scopus	Web of Science	Web of Science
Ranking focus	Research & Teaching	Research, Teaching & Innovation	Research	Research
Weight of research performance indicators2003	90%	40%	100%	100%

* It offered a joint ranking in collaboration with the Times Higher Education Suplement, since 2009 it offers an independent ranking

** Although its first edition dates back to 2008, it has not been published regularly since 2012. Since then it is published annually.

We analyzed the fields and subjects of the selected international rankings and we established the homologous field or discipline according to the I-UGR Rankings. In Tables 4.5-4.8 we show the matching of fields per ranking. In general terms, we observe that it is possible to match most of the fields between the four international rankings selected and the I-UGR Rankings, although some exceptions are noted. The areas misrepresented in the I-UGR Rankings were Mechanical Engineering (QS Ranking and NTU Ranking), Law (QS Ranking) and all of the areas considered of the Arts & Humanities fields by the QS Ranking. This is due to the way the I-UGR Rankings are constructed, as they rely on the JCR and these lack journal rankings for these fields. Also, we observe that some fields of the international rankings (i.e., the Shanghai Ranking and the field of Social Science) include more than one of the fields included in the I-UGR Rankings. Finally, the classification of fields and subfields does not always match between rankings. Although this issue has no relevance for the purposes of this analysis, we must point out that subjects considered as major areas in one ranking are considered in the other as subfields or disciplines.

The four selected rankings included a total of 33 Spanish universities dispersed in 51 different fields and subfields. In Tables 4.5-4.8 we show the levels of agreement between international and national rankings according to the assignment of areas. For each area we calculate the Spearman coefficient to analyze the consistency between both rankings and the number of universities included in international rankings which take up the top positions of the national ranking. That is, if 6 Spanish universities are included in an international ranking but only two occupy positions between 1 and 6, the coincidence will be 2/6.

The highest coincidence of universities between those present in the international rankings and top positions in the national ranking can be found in the NTU Ranking (Table 4.7), with 77.90% of the universities coinciding in both rankings. This ranking is followed by the Shanghai Ranking (Table 4.4) with 75.51% of the universities and the Leiden Ranking (Table 4.6) with 72.60%. The ranking with a lower percentage of coincidence is the QS Ranking (Table 4.5) with 56.49% of the universities present in this ranking reaching top positions in the national ranking.

Table 4.5 Matching of fields and disciplines between the Shanghai Ranking and the I-UGR Rankings

SHANGHAI RANKING	I-UGR RANKINGS	RHO	A
Natural Sciences & Mathematics	Mathematics / Physics / Chemistry	-0.866; 0; -0.866	0/3; 3/3; 2/3
Engineering/Technology & Computer Sciences	Engineering / Information & Communication Technology	*	1/3; 3/3
Life &Agricultural Sciences	Agricultural Sciences / BiologicalSciences	*	0/2; 1/2
Clinical Medicine & Pharmacy	Medicine & Pharmacy	*	2/2
Social Science	Other Social Sciences / Psychology & Education / Economics, Finance & Business	*	0/2; 0/2; 1/2
Mathematics	Mathematics	-0.817	6/9
Physics	Physics	0.179	6/7
Chemistry	Chemistry	0.523	7/9
Computer Science	Computer Science	0.677	6/9
Economics & Business	Economics, Finance & Business	0.000	2/3

Note: Rho indicates the Spearman's coefficient. A indicates the level of agreement between rankings, that is, the number of universities present in both rankings.
*Insufficient values to calculate the indicator

Analyzing the fields we find the following disciplinary concordance:

- The Shanghai Ranking is the less consistent with the I-UGR Rankings showing positive low correlation in two fields (Chemistry and Computer Science).
- The NTU Ranking shows correlations above 0.7 in 9 out of 17 fields. The three fields with the highest correlations can be found between the NTU Ranking and the I-UGR Rankings and these are Physics (0,952), Chemistry (0.945) and Biological Sciences (0,886).
- The QS Ranking shows correlations above 0.7 in 8 out of 23. The fields of Biological Sciences (0.866) and Life Sciences & Medicine (0.882 with Medicine & Pharmacy) are the fields with a higher correlation.
- The Leiden Ranking only shows a correlation above 0.7 in one field, Natural Sciences & Engineering, with the field of Chemistry in the national ranking.

If we focus on the disciplinary differences, the coincidence is especially relevant for the fields and subjects of Biomedicine, Life Sciences and Natural Sciences. This does not occur in the Social Sciences where the only exception noted is Economics.

In the case of rankings, in general terms, we can point out the following lessons learned:

- The NTU Ranking is the one which seems to be more consistent with the I-UGR Rankings. This is not surprising as it measures solely the research dimension and is fully based on the Web of Science, as it occurs with the I-UGR Rankings. Also, the confection of the fields and subfields is similar as both rankings aggregate subject categories to construct the fields, while in the other two cases this is not explained.

Table 4.6 Matching of fields and disciplines between the QS Ranking and the I-UGR Rankings

	QS RANKINGS	I-UGR RANKINGS	RHO	A
	Arts & Humanities			
	Engineering & Technology	Engineering	0.343	10/12
	LifeSciences & Medicine	Biological Sciences / Medicine & Pharmacy	0.609;0.882	9/11; 10/11
	Natural Sciences	Mathematics / Physics / Chemistry	-0.518; 0.773; 0.700	10/11; 9/11; 8/11
	Social Sciences & Management	Other Social Sciences/Psychology & Education/Economics, Finance & Business	0.545; 0.155; 0.482	8/11; 7/11; 7/11
Arts & Humanities	Philosophy			
	Modern Languages			
	Geography	Geography & City Planning	0.775	2/4
	History			
	Linguistics			
	English Language & Literature			

4.6 Table continues

	QS RANKINGS	I-UGR RANKINGS	RHO	A
Engineering & Technology	Computer Science & Information Systems	Computer Science	0.707	2/5
	Chemical Engineering	Chemical Engineering	0.463	3/6
	Civil Engineering	Civil Engineering	0.500	1/3
	Electrical Engineering	Electric & Electronic Engineering	0.154	3/6
	Mechanical Engineering			
Life Sciences & Medicine	Medicine	Medicine		1/2
	Biological Sciences	Biological Sciences	0.866	3/3
	Psychology	Psychology	0.414	4/6
	Pharmacy & Pharmacology	Pharmacy & Toxicology	0.507	5/6
	Pharmacy & Pharmacology	Pharmacy & Toxicology	0.507	5/6
	Agriculture & Forestry	Agriculture	0.461	4/10

Table 4.6 Table continues

	QS RANKINGS	I-UGR RANKINGS	RHO	A
Natural Sciences	Physics & Astronomy	Physics	0.775	3/4
	Mathematics	Mathematics	-0671	-0671
	Environmental Sciences	Earth & Environmental Sciences	0.632	2/4
	Earth & Marine Sciences	Earth & Environmental Sciences	*	1/2
	Chemistry	Chemistry	0.447	2/4
	Materials Science	Materials Science	*	2/3
Social Sciences & Management	Statistics & Operational Research	Statistics	0.612	6/10
	Sociology	Sociology	-0.289	3/5
	Politics & International Studies	Political Science	*	0/1
	Law			
	Economics & Econometrics	Economics	0.754	4/6
	Account & Finance	Business	0.775	3/4
	Communication & Media	Communication	0.289	0/5
	Education	Education	0.158	3/5

Note**:** Rho indicates the Spearman's coefficient. A indicates the level of agreement, that is, the number of universities present in both rankings.
*Insufficient values to calculate the indicator

Table 4.7 Matching of fields and disciplines between the Leiden Ranking and the I-UGR Rankings

LEIDEN RANKING	I-UGR RANKINGS	RHO	A
Biomedical & Health Sciences	Medicine & Pharmacy	0.518	10/15
Life & Earth Sciences	Biological Sciences / Earth & Environmental Sciences	0.600; 0.436	11/15; 9/15
Mathematics & Computer Science	Mathematics / Information & Communication Technology	0.307; -0.036	12/15; 9/15
Natural Sciences & Engineering	Engineering / Mathematics / Physics / Chemistry	0.350; 0.264; 0.496; 0.736	12/15; 12/15; 12/15; 11/15
Social Sciences & Humanities	Other Social Sciences / Psychology & Education / Economics...	0.121; 0.115; 0.220	8/13; 8/13; 8/13; 8/13

Note**:** Rho indicates the Spearman's coefficient. A indicates the level of agreement.*Insufficient values to calculate the indicator

Table 4.8 Matching of fields and disciplines between the NTU Ranking and the I-UGR Rankings

NTU RANKING	I-UGR RANKINGS	RHO	A
Agriculture	Agriculture	0.406	5/10
Clinical Medicine	Medicine	*	2/2
Engineering	Engineering	0.418	9/11
Life Sciences	Biological Sciences	0.886	5/6
Natural Sciences	Mathematics / Physics / Chemistry & Chemical Engineering	0.127; 0.879; 0.588	6/10; 8/10; 8/10
Social Sciences	Other Social Sciences / Psychology & Education / Economics…	0.600; 0.000; 0.400	2/4; 2/4; 2/4
Agricultural Sciences	Agricultural Sciences	0.440	18/23
Environment/Ecology	Earth & Environmental Sciences	*	0/0
Plant & Animal Science	Biological Sciences	0.552	5/10
Computer Science	Computer Science	0.812	13/16
Chemical Engineering	Chemical Engineering	0.846	8/12
Civil Engineering	Civil Engineering	0.202	10/12
Electrical Engineering	Electrical & Electronic Engineering	0.755	8/11
Mechanical Engineering			
Materials Science	Materials Science	0.757	5/7
Pharmacology	Pharmacy & Toxicology	0.300	5/5
Chemistry	Chemistry	0.945	14/15
Geosciences	Geosciences	0.847	6/7
Mathematics	Mathematics	0.524	11/12
Physics	Physics	0.952	7/8

Note**:** Rho indicates the Spearman's coefficient. A indicates the level of agreement.*Insufficient values to calculate the indicator

- Another issue which affects this in the other two ranking (Shanghai Ranking and QS Ranking) has to do with the way results are presented, as they only show the intervals in which each university is positioned after they surpass certain threshold. Although the QS Ranking provides the rating of each university, allowing the user to rank universities, this does not occur with the Shanghai Ranking.

4.4. Concluding remarks and lessons learned

In this paper we explore the possibility of using national rankings to complement international rankings, as the latter usually offer a poor representation of national university systems (no more than 25% of the system in the Spanish case). We insist on the importance of rankings by fields (García et al., 2012) as these do not neglect universities' disciplinary focus and offer a complete picture of universities' research performance. This perspective follows the recent trend on evaluative bibliometrics for 'opening up' these tools in order to offer, rather than a narrow and simplistic solution, a range of different outputs that can better serve research policy makers to make the right decisions considering their specific aims and different scenarios (Rafols et al., 2012).

We use Spain as a study case and we introduce the I-UGR Rankings for Spanish universities. This ranking uses the IFQ^{2}A Index, an indicator which measures the qualitative as well as the quantitative dimension of research (Torres-Salinas et al., 2011c). From this analysis we conclude that national rankings can complement international rankings in order to provide a complete picture of university systems despite the methodological differences aroused from the comparisons by fields. However, we must stress the importance of acknowledging such methodological differences to better interpret them. Such differences are mainly derived from the construction of fields and subfields as well as the indicator selected for ranking universities.

Our conclusion is clear as to the importance and complement that represent the national rankings to address a comprehensive analysis of the university system of a country. The joint analysis of both types of rankings will provide a complete snapshot of the universities and their scientific strengths.

These results show different levels of concordance which are affected not only by methodological issues but also by the way these fields are constructed and the difficulties implied in this process which affected differently each scientific domain. Despite this, it is possible to use both (national and international rankings) and combine the information provided in a research policy context.

References

Abramo, G., Cicero, T. &D'Angelo, C.A. (2011). The dangers of performance-based research funding in non-competitive higher education systems. *Scientometrics*, 87(3), 641-654.

Aguillo, I., Bar-Ilan, J., Levene, M. & Ortega, J.L. (2010). Comparing university rankings. *Scientometrics*, 85(1), 243-256.

Bornmann, L., Mutz, R. & Daniel, H.-D. (2013). Multilevel-statistical reformulation of citation-based university rankings: The Leiden ranking 2011/2012. *Journal of the American Society for Information Science and Technology, 64*(8), 1649-1658.

Cabezas-Clavijo, Á., Robinson-García, N, Escabias, M. & Jiménez-Contreras, E. (2013). Reviewers' ratings and bibliometric indicators: Hand in hand when assessing over research proposals? *Plos One, 8*(6), e68258.

Collini, S. (2011). *What are universities for?* London, UK: Penguin Books.

Delgado López-Cózar, E. (2012). Cómo se cocinan los rankings universitarios. *Dendra Médica. Revista de Humanidades*, 11(1), 43-58.

Docampo, D. (2011). On the use of the Shanghai Ranking to assess the research performance of university systems. *Scientometrics*, 86(1), 77-92.

García, J.A., Rodríguez-Sánchez, R., Fdez-Valdivia, J., Robinson-García, N. & Torres-Salinas, D. (2012).Mapping Institutions According to their Journal Publication Profile: Spanish Universities as a Case Study. *Journal of the American Society for Information Science and Technology*, 63(11), 2328-2340.

Hazelkorn, E. (2011). *Rankings and the Reshaping of Higher Education. The Battle for World-Class Excellence*.Houndmills, UK: Palgrave MacMillan.

Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences USA*, 102(46), 16569–16572.

Leydesdorff, L. & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 63(11), 2239-2253.

Liu, N.C. & Cheng, Y. (2005). The Academic Ranking of World Universities. *Higher Education in Europe*, 30(2), 127-136.

Marginson, S. & van der Wende, M. (2007). To Rank or to be Ranked: The Impact of Global Rankings in Higher Education. *Journal of Studies in International Education*, 11(3-4), 306-329.

Moed, H.F. (2005). *Citation Analysis in Research Evaluation.* Dordrecht: Springer.

Moed, H.F., Moya-Anegón, F., López-Illescas, C. & Visser, M. (2011). Is concentration of university research associated with better research performance? *Journal of Informetrics*, 5(4), 649-658.

Orduña-Malea, E. (2012). *Propuesta de un modelo de análisis redinformétrico multinivel para el estudio sistémico de las universidades* [Doctoral Thesis]. Valencia, Spain: Universidad Politécnica de Valencia.

Rafols, I., Ciarli, T., van Zwanenberg, P. & Stirling, A. (2012). Towards indicators 'opening up' science and technology policy. In: E. Archaumbault& V. Larivière (eds) *Proceedings of 17th International Conference on Science and Technology Indicators,* 675-682.

Shin, J. C. &Toutkoushian, R.K. (2011).The Past, Present, and Future of University Rankings. In: J.C. Shin, R.K. Toutkoushian & U. Teichler (eds). *University Rankings. Theoretical Basis, Methodology and Impacts on Global Higher Education.* Dordrecht, Netherlands: Springer, pp. 1-18.

Torres-Salinas, D., Delgado López-Cózar, E., Moreno-Torres, J.G. & Herrera, F. (2011a). Rankings ISI de las Universidades Españolas según Campos Científicos: Descripción y Resultados. *El Profesional de la Información*, 20(1), 111-118.

Torres-Salinas, D., Moreno-Torres, J.G., Robinson-García, N., Delgado López-Cózar, E. & Herrera, F. (2011b). Rankings ISI de las Universidades Españolas según Campos y Disciplinas Científicas (2ª ed. 2011). *El Profesional de la Información*, 20(6), 701-709.

Torres-Salinas, D., Moreno-Torres, J.G., Delgado López-Cózar, E. & Herrera, F. (2011c). A methodology for Institution-Field ranking based on a bidimensional analysis: the IFQ^{2}A index. *Scientometrics*, 88(3), 771-786.

Toutkoushian, R.K. & Webber, K. (2011). Measuring the Research Performance of Postsecondary Institutions. In: J.C. Shin, R.K. Toutkoushian & U. Teichler (eds). *University Rankings. Theoretical Basis, Methodology and Impacts on Global Higher Education.* Dordrecht, Netherlands: Springer, pp. 123-144.

Usher, A. & Savino, M. (2007). A Global Survey of University Rankings and League Tables. *Higher Education in Europe*, 32(1), 5-15.

Van Raan, A.F.J. (2005). Fatal Attraction: Conceptual and Methodological Problems in the Ranking of Universities in the Global Scientific Communication System. *Scientometrics*, 62(1), 133-143.

Visser, M.S., Calero-Medina, C.M. & Moed, H.F. (2007). Beyond Rankings: The Role of Large Research Universities in the Global Scientific Communication System.In: *Proceedings of 11th Conference of the International Society for Scientometrics and Informetrics.* Madrid, Spain: CINDOC-CSIC, pp. 761-765.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E.C.M., Tijssen, R.J., van Eck, N.J., et al. (2012). The Leiden Ranking 2011/2012: Data Collection, Indicators and Interpretation. *Journal of the American Society for Information Science and Technology, 63*(12), 2419-2432.

5. What do university rankings by fields rank? Exploring discrepancies between the organizational structure of universities and bibliometric classifications

Paper published in Scientometrics doi:10.1007/s11192-013-1157-7 authored by N. Robinson-García and C. Calero-Medina

University rankings by fields are usually based on the research output of universities. However, research managers and rankings consumers expect to see in such fields a reflection of the structure of their own organizational institution. In this study we address such misinterpretation by developing the research profile of the organizational units of two Spanish universities: University of Granada and Pompeu Fabra University. We use two classification systems, the subject categories offered by Thomson Scientific which are commonly used on bibliometric studies, and the 37 disciplines displayed by the Spanish I-UGR Rankings which are constructed from an aggregation of the former. We also describe in detail problems encountered when working with address data from a top down approach and we show differences between universities structures derived from the interdisciplinary organizational forms of new managerialism at universities. We conclude by highlighting that rankings by fields should clearly state the methodology for the construction of such fields. We indicate that the construction of research profiles may be a good solution for universities for finding out levels of discrepancy between organizational units and subject fields.

5.1. Introduction

One of the most common approaches in bibliometrics for benchmarking multidisciplinary entities such as universities, research teams or institutes, is the use of classification-based tools and indicators (i.e., Moed, 2010; Leydesdorff & Ophtof, 2010; Waltman & van Eck, 2012b). This is the case in university rankings, which are now incorporating league tables by fields as a response to criticisms due to an over-simplistic perspective as these rankings tend to reduce the complex framework of universities' activity to a single dimension (see e.g., van Raan, 2005). The first one to include disciplinary-oriented league tables was the Shanghai Ranking, launching in 2007 rankings by five broad fields and in 2009 five more rankings in specific disciplines. Since then, many other international rankings have followed such perspective, such as the Times Higher Education World University Rankings, the QS Rankings or the National Taiwan University Rankings, for instance. The Leiden Ranking has been the last one to follow this trend, including in its 2013 edition rankings by five broad areas. Others, such as the Scimago Institutions Rankings do not show league tables by fields but include a specialization index.

This is partly because of the influence disciplinary specialization may have on research evaluation (López-Illescas, Moya-Anegón, Moed, 2011) which means that one must identify universities with similar disciplinary focuses (García, Rodriguez-Sanchez, Fdez-Valdivia et al., 2012) as an aid to interpret such comparisons. Also, global comparisons may be 'unfair' to certain types of universities as their subject profile may influence their positioning (Bornmann, Moya-Anegón & Mutz, 2013). Cheng & Liu (2006) already attempted at identifying disciplinary-oriented institutions by using clustering methods and, in a more recent study, Bornmann et al. (2013) developed a web application which maps centers of excellence according to different fields. All these evidences show the need to bypass the use of global rankings and focus on developing field-based tools.

The most commonly used classification system in bibliometrics is the one designed by Thomson Scientific (TS), which groups scientific journals following heuristic criteria based on citation data (Pudovkin & Garfield, 2002). Although it shows some limitations when used for bibliometric purposes (Glänzel & Schubert, 2003; Waltman & van Eck, 2012a), it seems to be a practical and plausible way to aggregate categories into areas when developing rankings by fields (see e.g., the 2013 edition of the Leiden Ranking which now includes rankings by five broad fields or the Spanish fields-based I-UGR Rankings described in Robinson-García et al., 2013a based on the indicator developed by Torres-Salinas et al., 2011). But this approach based on universities' output seems counter-intuitive when being read by ranking consumers, as they expect to see a *bottom-up* methodology which would determine the institutional structure of universities and hence, develop league tables according to their units (faculties, departments, etc.). Such granularity in the information provided by rankings has already been suggested elsewhere (Bornmann, Mutz & Daniel, 2013) as there are significant differences in terms of research performance between research units of the same institution. This would allow an attribution of the performance of a given university in a particular field to researchers assigned to the units related to such field. However, this is not always possible as there is no unified database containing such information, meaning that universities should have to be involved in the data collection process (van Leeuwen, 2007). Because of this, university rankings usually adopt a *top-down* approach, a reasonable solution but one which usually leads to misinterpretations by media and research policy makers.

Even so, one could suggest the use of the address data included in publications when constructing rankings by fields. De Bruin & Moed (1993) already suggested working with address data in order to develop a subject classification scheme based on an institutional structure. They departed from three basic assumptions: 1) scientific activity can be analyzed in terms of collaboration between research groups, 2) organizational units reflect to some extent the scientific scope of their members and 3) researchers indicate in their publications the organizational units in which they work. For this, they created a genealogical structure of the address data in order to identify cognitive terms from those which weren't and then applied a clustering method to isolate each sub-field. However, address data presents many problems as the authors acknowledged, seeming unfeasible to do this at a large scale due to the heterogeneity of universities' structure, the possible changes over time as a result of organizational re-shuffling, and the growing complexity of the problem.

Finally there is another aspect that should be mentioned. Although desirable, any attempt to develop more accurate and precise bibliometric rankings can bring other unexpected issues. The bibliometric field has an applied nature and empirical roots when compared with more basic sciences. This is why it does not only require detailed and full characterization of the analyzed entities, but also a certain level of assertiveness and security over its applicability in real situations; leaving a threshold of uncertainty. Hence, although a solution to a given problem may be theoretically correct, it may be wrongly interpreted or adjusted in a certain context. Such tension between the accuracy and precision demanded by any scientific tool, and the security on any statement needed when facing possible research policy applications is explicitly defined in Duhem's Law of Cognitive Complementarity. In it, Duhem highlights the inverse relation between detail and security, stating that in order to access to truth one needs certain levels of vagueness which will secure its reliability in any given situation (Rescher, 2006). Any attempt to create rankings by fields exposes itself to such dilemma, as the applicability of such fields on the organizational structure of universities may differ from one to another.

5.2. Objectives of the study

This paper highlights the difficulties university research managers and other ranking consumers may have when understanding university rankings by fields as they misinterpret them by expecting to see in those fields a reflection of the structure of their university. As university rankings producers, bibliometricians must not only be transparent on the methodology and data employed, but also ensure a reasonable interpretation of the results they offer. We examine the relation between the institutional organization of science as reflected by authors' affiliation data and their research output. Specifically, our aims can be resumed in the following research questions (RQ):

RQ1. Do rankings by fields represent the structure of universities? De Bruin & Moed (1993), suggest that address data may be useful for identifying scientific fields and domains. Is there some kind of correspondence between the fields constructed by rankings and smaller organizational units such as departments, research group, faculties, etc.?

RQ2. Can we provide research policy managers and ranking consumers with indicators that they can use in order to understand to which degree each field (based on the Web of Science subject categories) corresponds with the output of their organizational units at an institutional level?

All in all, the purpose of this study is to offer a deeper understanding on what do the classification systems used in bibliometric studies and university rankings represent according universities' organizational units and how can bibliometric indicators ease the interpretation of such fields.

5.3. Material and methods

In this paper we focus on two Spanish universities as case studies: University of Granada and Pompeu Fabra University. We focus in a single country for two reasons. Firstly, the authors' own expertise and knowledge on the Spanish university system which helps to better interpret the results obtained. Secondly, as Bornmann, Mutz & Daniel (2013) point out, national systems influence the research performance of universities. In this sense, it is of interest to identify differences between universities of a same country. These universities represent two different types of institutions. The former is a historical university with a well-established structure and present in all editions of the Shanghai Ranking and most world-class university rankings. The latter is a small and relatively new university funded in 1990 which has rapidly gained positions in many international rankings converting itself in an interesting success case as pointed out in other studies (Robinson-García et al., 2013b). They are chosen due to their dissimilarity on size, historical background and structure, as one will expect to see some differences regarding the new managerialism of universities and its effect on their institutional organization (Morris, 2002). However, they also have some common grounds as they both belong to the same higher education system. This allows a better interpretation of the results and the effects such structural differences may have on their research performance.

This section is structured as follows. Firstly, we account for the data collection process and the time period used. Secondly, we explain in detail the problems that arise when processing address data and the many limitations it may present. Finally, we describe the methodology employed for analyzing the structure of universities and the indicators proposed for presenting the research profile of each organizational unit and understanding the relation between them and the fields as defined by a given classification system.

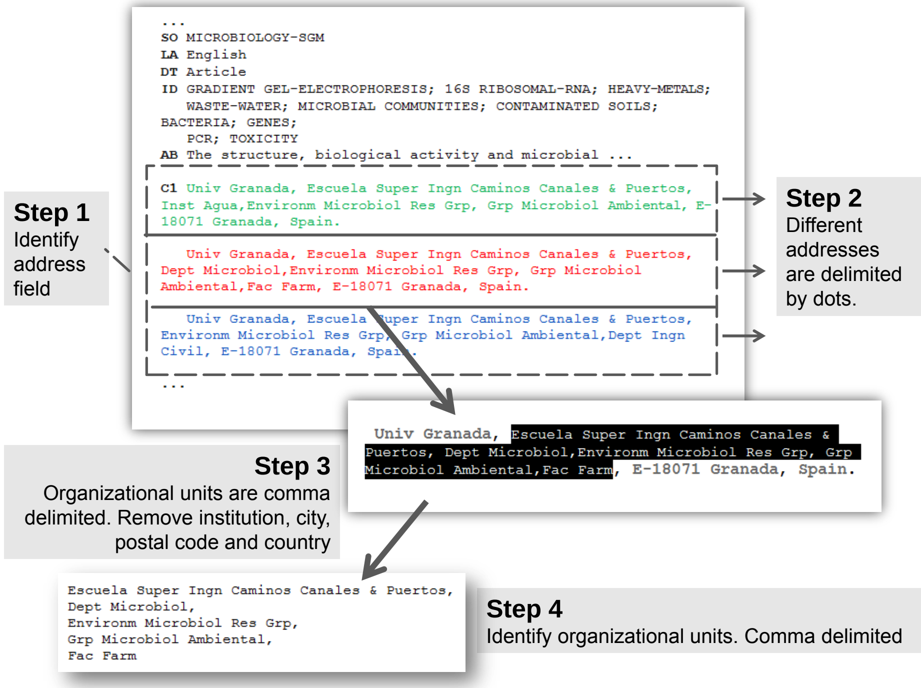
5.3.1. Data collection and processing

We used the 2006-2010 time period, a fixed four-year citation window and the TS Science Citation Index (SCI), Social Sciences Citation Index (SSCI) and Arts & Humanities Citation Index (A&HCI) as data sources. In order to gather the research output of the universities we used the in-house CWTS version of the TS Web of Science, which identifies universities' output taking into account all possible name variations.

Once the research output of a given university was identified, we delimited our focus only on addresses belonging to the institution under study. Therefore, in a paper published by Pompeu Fabra University in collaboration with London School of Economics we will only consider the affiliation data related to Pompeu Fabra University and omit the one related to the collaborating institution. This way we isolate addresses referring to the institution under study. In Figure 5.1 we show how this process was followed. Hence, on a first step we identify the address field. As observed in the figure, the record used as an example includes three different addresses delimited by dots, all of them belonging to the University of Granada. This means that none will be discarded. Next, we observe that each address is further divided by organizational units. In step 3, we notice that these are separated by commas and furthermore, that the first unit identifies the major organizational level (in our case, the university) and that the two last ones identify post code and city, and country accordingly. Thus these fields can be automatically removed. Once this has been done, we can identify the rest of the organizational units by dividing the field using commas as separators. In the example used in Figure 5.1, we identify five different organizational units in step 4. In this case, the address refers to a double affiliation including a school, a faculty, a department and two research groups. This is in fact an interesting example that shows the many inconsistencies one may find when working with address data. On the one hand, it includes two units which belong to the same hierarchical level (school and faculty) which could question if the department belongs to both organizational units or if this address should have been treated as two different addresses. Also, a closer look at the information provided for the two research groups will show that they are in fact the same research group displayed in English and Spanish language.

Although in this example organizational units belonged to four different structural levels, authors may not always indicate all units and may omit the faculty or school for instance, only reflecting departmental information or include only information regarding the highest organizational unit (school and faculty in our example). Indeed, the variability on the information provided by this field varies significantly, ranging from records which do not offer any organizational unit (meaning that they will not be retrieved in this analysis) to records which offer other information which is not always related with their position within the structure of their university. In this sense, authors may indicate as address data the funding agency which supports them, national collaboration networks, research programs or the name of the building in which they work, for instance. Also there are problems when establishing boundaries between units belonging to a university and mixed units with more than one affiliations such as hospitals, research institutes, etc. (for a further discussion the reader is referred to Praal et al. (2013) in which they address the many problems that can arise regarding the assignment of hospitals in the United Kingdom). As in this paper we focus on institutional structure, papers including only the main organizational level (university) were discarded from the analysis. In Figures 5.2 and 5.4 we show the distribution of organizational units and the proportion they represent of the total output.

Figure 5.1. Example on the procedure for identifying organizational units within bibliographic records from the Web of Science



Another relevant issue when dealing with address data has to do with the many normalization problems mainly due to misspellings, use of different languages, name changes, the use of acronyms and errors made by Web of Science; this results in a necessary manual cleaning of data. Regarding the latter problem, we have noted many inconsistencies between the information provided by the authors in publications and the one displayed in the Web of Science. In Table 5.1 we include the most common denominations out of a total of 41 name variations found for a department belonging to the University of Granada. This data cleaning process was made by checking with the institutional website in order to compare results, although in some cases we found out that certain institutions or units defined by the authors were not included in the institutional layout. Units belonging to parallel structures such as research groups, hospitals, research programs or national collaboration networks were preserved.

Table 5.1. Name variations and number of papers linked to a department from the University of Granada

	No PUB*	DEPARTMENT
CHOSEN DESIGNATION	162	**DEPT COMP SCI & ARTIFICIAL INTELLIGENCE**
VARIABLES	33	DEPT COMP SCI & AI
	18	DEPT CIENCIAS COMPUTAC & IA
	15	DEPT CIENCIAS COMPUTAC & INTELIGENCIA ARTIFICIAL
	9	DEPT COMP SCI
	4	COMP SCI & ARTIFICIAL INTELLIGENCE DEPT
	4	DEPT CIENCIAS COMP & INTELIGENCIA ARTIFICIAL
	4	DPTO CIENCIAS COMPUTAC & INTELIGENCIA ARTIFICIAL
	3	AI
	3	DEPT COMP SCI & ARTIFICAL INTELLIGENCE
	2	DECSAI
	2	DEPT CIENCIAS COMPUTAC
	2	DEPT COMPUTAT SCI & AI
	28	*OTHER VARIATIONS WITH PUBLICATION FREQUENCE 1*
TOTAL	**281**	

* A deparment may be included several times in the same paper

5.3.2. Construction of research profiles and indicators used

The goal of this study is to understand the relation between fields as constructed in rankings and bibliometric studies, and the structure of universities as defined by their organizational units and offer indicators that can explain such relation. To this aim, we developed an organizational network for each university under study which would allow a general view of its structure according to its research output. Organizational units may co-occur in a document for different reasons. Hence, the following cases may take place: several authors belonging to different departments (in-house collaboration), one author indicating different organizational units all within each hierarchical level (i.e., faculty, department, research group) or one author with double affiliation (i.e., faculty and research center). Therefore, links in our network will define organizational relations between units in its broadest sense.

Such networks are shown in Figures 5.3 and 5.5, and they allow us to identify the units which occupy a central or most 'prominent' position in the structure, that is, they have more potential power and influence due to their connections to the rest of the nodes (Borgatti, Mehra, Brass & Labianca, 2009). Here we propose the use of centrality indicators to understand the role of a given organizational unit within the rest of the network. There are different indicators which measure the centrality of nodes; in this paper we use the betweenness indicator. A node will have high betweenness centrality if it appears often in the shortest path that connects any two other nodes. In Figures 5.3 and 5.5 the betweenness centrality measure is represented by the size of the nodes. Table 5.3 and 5.4 present the research profiles of departments with more than 50 publications in the case of University of Granada and for any organizational unit with more than 50 publications in the case of Pompeu Fabra University. The methodology for the construction of research profiles is based on the work by Calero-Medina & van Leeuwen (2012) and consists on 'breaking down' the output of an organizational unit into subject fields based on a given classification system. This way one can observe the 'interdisciplinarity' of such unit. Finally, we propose to combine the betweenness centrality measure with the Gini Index, as a means to observe how well represented are organizational units by fields of a given classification system. In table 5.2 we include a list of the indicators employed along with their definition.

Table 5.2. Description of the indicators used.

Indicator	Acronym	Definition
Number of publications	P	Publications indexed in the Web of Science citation indexes (SCI, SSCI and A&CI). The considered document types were letters, articles, reviews and proceedings papers.
Betweenness Centrality	B	The Betweennes Centrality measure indicates the nodes which appear more often when connecting two other nodes in a network. A node will have high betweenness centrality if it appears often in the shortest path that connects any two other nodes.
Gini Coefficient	G	The Gini Coefficient is an inequality indicator which shows the concentration or scattering of distributions. It is commonly used in the field of Economics to analyze the distribution of wealth. In this study we use it to analyze the distribution of an organizational unit's output according to subject fields. Its value ranges from 0 to 1; 0 meaning no concentration and 1 concentration in a single subject field. In this paper we used the formula defined by Deaton (1997).
Number of subject categories	No SC	By subject categories we refer to the classification system employed by TS.
Number of disciplines	No disc	By disciplines we refer to the classification system employed by the Spanish I-UGR Rankings. Such system is based on TS' subject categories from the SCI and the SSCI, and defines a total of 37 disciplines. The construction of these disciplines is available at http://www.ugr.es/~elrobin/docs/disciplines_I-UGR_Rankings.xlsx.

Two classification systems were selected according to two possible scenarios in which institutional analyses by fields take place:

1) The TS subject categories. This is the most common classification system used in bibliometric studies. In fact, it is the one employed by Calero-Medina & van Leeuwen (2012), who use it in order to construct inverse research profiles (that is, a breakdown of subject categories into organizational units) as a means to analyze the contribution of different research programs to a given research field.

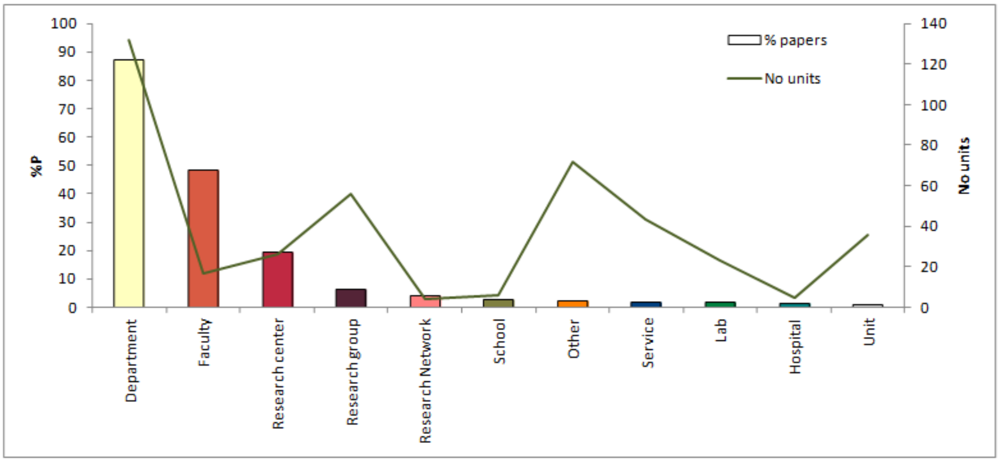
2) Aggregation of subject categories. A common methodology employed also in bibliometric analyses at a macro-level (i.e., García et al., 2012) and in university rankings by fields (i.e., the 2013 edition of the Leiden Ranking or the Spanish I-UGR Rankings of field and disciplines). In this study we will use the 37 aggregated disciplines defined in the I-UGR Rankings, as this will allow us to discuss the implications of possible discrepancies between fields and organizational units in university rankings.

5.4. Results

5.4.1. Case 1. University of Granada

The University of Granada had a total output of 6913 publications for the 2006-2010 time period of which 6337 were finally included in this study. The remaining did not include any information at the organizational unit level. In figure 5.2 we include a general overview of different types of organizational units used by authors to indicate their affiliation. As observed, the most common information included is the department, which is present in 5514 papers which represent 87.0% of the total output analyzed. Also, this organizational type is the one in which a wider number of units were found with 132 departments. In total nearly half of the total share (48.3%) included information regarding the faculty to which authors belonged, followed by far by papers including information regarding the research center to which authors were affiliated (19.6). The rest of the organizational types account each for less of 10% of the total output. The other organizational type with the largest number of different after departments is others, which is a miscellaneous group in which one may find a wide range of different units such as office, job post or errors in the database such as other universities involved, not following the rationale of the address field as described previously in Figure 5.1.

Figure 5.2. Percentage of publications and total number of units by type for the University of Granada according to organizational types. Time period 2006-2010



In Figure 5.3 we show the structure of the university according to its organizational units. As it was explained previously the links in the network will define organizational relations between units. As observed, four different components can be found, three small ones and a main component. Units are organized around faculties and departments and occasionally around research centers. These are the main organizational units. This component can be further divided in seven distinct parts. On the upper right we find organizational units related with the fields of Behavioral Sciences and Neuroscience. One of the two main clusters is formed around the Faculty of Sciences FAC SCI) which connects through the Faculty of Pharmacy (FAC PHARM) with the other main cluster representing the Biomedical Sciences and formed by the faculties of Medicine (FAC MED) and Dentistry (FAC DENT). On the upper left we have Engineering and on the lower left, Physics. On the lower right, we find units related with fields from Computer Science which connect with those related with Information Science.

Figure 5.3. Organizational network of the University of Granada according to coupling of organizational units. Map characteristics: Lines: minimum co-occurrence value >5. Isolated nodes have been removed. Colors: Types of organizational units as show in Figure 2. Size: Betweenness values

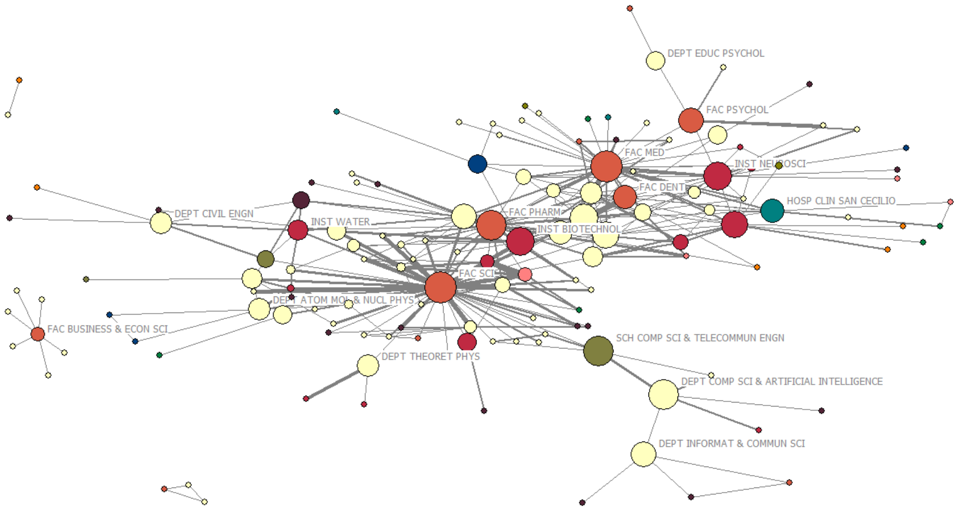


Table 5.3 includes an overview of the research profiles for all departments of the University of Granada with more than 50 publications in the 2006-2010 time period. For each department we include the total publications, betweenness centrality and their Gini Coefficient and number of subject fields according to each classification system. These three indicators offer valuable information on the research profile of each department and the capability of the classification system to isolate its output. For instance, we observe that the department of Mathematical Analysis (DEPT ANAL MATH) is the one with the highest concentration according to the TS Classification (0.86) with its output distributed among 6 subject categories and all of it included in a single discipline according to the I-UGR classification system. Regarding its importance in the rest of the network, it does not have a central position in terms of being a department that connects units that otherwise will be unconnected. This is the reason why its betweenness centrality is zero. This result goes online with the higher values for the Gini coeficients. The deparment of Mathematical Analysis is very focus on certain fields. Hence, both classification systems can accurately reflect this department's output.

Table 5.3 Bibliometric indicators and research profiles of departments from University of Granada with >50 publications for the 2006-2010 time period and their output distribution according to two classification systems: TS subject categories and I-UGR disciplines

			TS CLASSIFICATION		I-UGR CLASSIFICATION	
Department	**P**	**B**	**G**	**No SC**	**G**	**No disc**
DEPT COMP SCI & ART INTELLIGENCE	281	735.00	0.76	57	0.84*	20
DEPT APPL PHYS	266	34.02	0.63	62	0.75	21
DEPT STAT & OPERAT RES	239	682.72	0.62	75	0.62	23
DEPT ZOOL	231	0.00	0.67	36	0.74	14
DEPT INORGAN CHEM	225	0.00	0.69	34	0.82*	14
DEPT ATOM MOL & NUCL PHYS	223	250.67	0.64	26	0.67	6
DEPT THEORET PHYS	221	251.00	0.77	12	0.77	3
DEPT ANALYT CHEM	218	5.00	0.77	41	0.79	13
DEPT PHYSIOL	203	235.60	0.64	44	0.71	15
DEPT STOMATOL	180	339.98	0.80*	31	0.79	9
DEPT APPL MATH	176	136.50	0.76	37	0.67	12
DEPT MINERAL & PETROL	175	4.01	0.66	46	0.77	13
DEPT MICROBIOL	165	405.64	0.69	45	0.62	15
DEPT STRATIG & PALEONTOL	164	0.50	0.73	31	0.84*	9
DEPT GEOMETR & TOPOL	140	0.00	0.77	9	1.00*	1
DEPT MATH ANAL	136	0.00	0.86*	6	1.00*	1
DEPT EXPT PSYCHOL	124	0.00	0.68	36	0.81*	12
DEPT ORGAN CHEM	118	126.00	0.70	26	0.87*	12
DEPT CHEM ENGN	109	0.00	0.61	27	0.61	14
DEPT GEODYNAM	109	0.00	0.68	25	0.87*	8
DEPT COMP ARCHITECTURE & TECHNOL	97	0.00	0.62	35	0.78	14
DEPT PHARMACOL	97	60.63	0.58	38	0.70	14
DEPT ECOL	96	3.83	0.61	23	0.63	9
DEPT CIVIL ENGN	91	251.00	0.61	42	0.67	18
DEPT ELECTROMAGNET & PHYS MATTER	88	4.50	0.57	29	0.56	9
DEPT PHARM & PHARMACEUT TECHNOL	82	0.00	0.60	24	0.71	9
DEPT PLANT PHYSIOL	81	0.00	0.62	28	0.65	12
DEPT OPT	77	126.00	0.62	17	**0.43**	10
DEPT PERSONAL ASSESSMENT & PSYCH...	75	0.00	0.59	35	0.67	9
DEPT PHYS CHEM	75	0.00	0.64	20	0.70	8
DEPT GENET	74	7.43	0.56	25	0.65	9
DEPT NUTR & BROMATOL	74	0.00	0.67	16	0.59	10
DEPT ELECT & COMP TECHNOL	72	0.50	0.70	26	0.72	8
DEPT BIOCHEM & MOL BIOL	69	536.69	0.52	37	0.70	11
DEPT ALGEBRA	67	0.00	0.74	3	1.00*	1
DEPT BOT	66	2.45	0.54	32	0.65	14
DEPT INFORMAT & COMMUN SCI	66	374.00	0.79	13	0.80	6
DEPT ANAT PATHOL & HIST CIENCIA	58	19.22	**0.43**	33	0.64	11
DEPT LANGUAGES & COMP SYST	58	0.00	0.57	29	0.72	13
DEPT SIGN THEORY TELEMAT ...	56	0.00	0.53	29	0.61	13
DEPT BIOCHEM & MOL BIOL 2	54	194.04	0.51	25	0.62	9

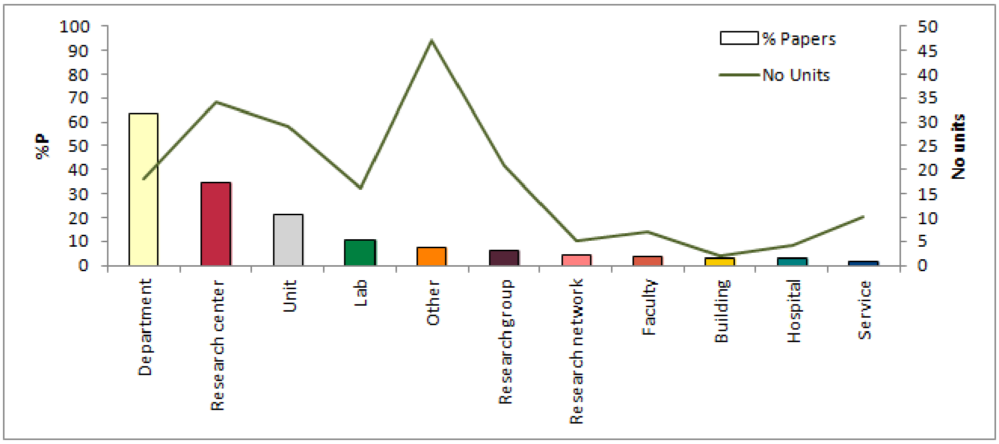
In bold when G <0.5; with an * when G>0.8

On the other extreme we find the department of Optics (DEPT OPT), which has a 0.62 Gini coefficient distributed among 17 different subject categories according to the TS classification, and a 0.43 Gini coefficient distributed among 10 different disciplines. This means that its output is neglected by subject fields as it is widely distributed. Regarding its role in the institutional structure of the university, it has betweenness value of 126.0. Another different case is the one of departments which, despite concentrating most of their output in certain subject categories, they contribute to many other fields. This occurs with the output of the department of Computer Sciences & Artificial Intelligence (DEPT COMP SCI & ARTIFICIAL INTELLIGENCE) which shows Gini values above 0.7 for both classification systems but performs in 57 different subject categories according to the TS classification and in 20 disciplines according to the I-UGR classifications. Although its output is mainly focused on Computer Sciences, it also performs in many other areas, this is also observed by its relation within the network where its betweennes centrality has a value of 735.0. Other examples of this case can be observed on the department of Statistics & Operational Research (DEPT STAT & OPERAT RES) and the department of Applied Physics (DEPT APPL PHYS). In fact, these three departments are the most productive ones leading to suggest that the more output produced by a department, the larger the contribution to other fields may be.

5.4.2. Case 2. Pompeu Fabra University

Pompeu Fabra University had a total output of 2480 publications for the 2006-2010 time period of which 1760 were finally included in this study. The remaining publications did not include any information at the organizational unit level. As it occurred with University of Granada, department is the most frequent organizational type included in its publications, although in this case its share drops to 63.5% (1117 publications). In fact, more variability is found in the organizational types adopted by authors at this university. Also the order of the most present types varies and research centers are the second most frequent choice accounting for 34.3% of the total share followed by units which represent 21.2% of the output. The faculty name is rarely used and it represents 3.5% of the total share, that is, 61 publications. Regarding the number of units, the structure of the university also differs from the first case study. There are less departments (18), and faculties (7) and the greatest number of units can be found in the miscellaneous group ‘Other’ (47), followed by research centers (34) and then units (27) and research groups (21).

Figure 5.4. Percentage of publications and total number of units by type for the Pompeu Fabra University according to organizational types. Time period 2006-2010



In figure 5.5 we show the structure of the university according to its organizational units. In this case, units are organized around departments, research centers and occasionally around units. Contrarily to Granada, and despite having seven faculties and a school, these organizational units are almost absent in the address data offered by researchers from Pompeu Fabra. The network is formed by a single component. Also, we observe that, except the lower left of the figure, most of the organizational units are related with fields from the Biomedical Sciences, displaying a highly specialized university. The organizational units displayed on the lower left are related with fields from the Social Sciences and Computer Science.

Due to the larger distribution of organizational types along with lower output figures, in this case study we have considered all organizational units with more than 50 publications and not only departments. The research profiles of each of them along with some bibliometric indicators are shown in table 5.4. In fact, only four departments are above such threshold, the most productive of them, the department of Experimental & Health Sciences (DEPT EXPT & HLTH SCI) accounting for 34.0% of the total share of Pompeu Fabra University. Also we find other institutions included which do not actually belong to this university. It is the case of the Institut Municipal D'Investigacions Mèdiques (INST MUNICIPAL MED RES IMIM) which is a mixed institution belonging Hospital del Mar but whose staff is affiliated to various institutions such as Pompeu Fabra University, Autonomous University of Barcelona or the Centre of Genomic Regulation. This institution along with the latter (CTR GENOM REGULAT CRG), are both mixed research centers with staff from different

Catalan universities. They all belong to the Barcelona Biomedical Research Park (PARC RES BIOMED BARCELONA), where along with others, these research centers are located.

Figure 5.5. Organizational network of Pompeu Fabra University according to coupling of organizational units. Map characteristics: Lines: minimum co-occurrence value >3. Isolated nodes have been removed. Colors: Types of organizational units as show in Figure 2. Size: Betweenness values

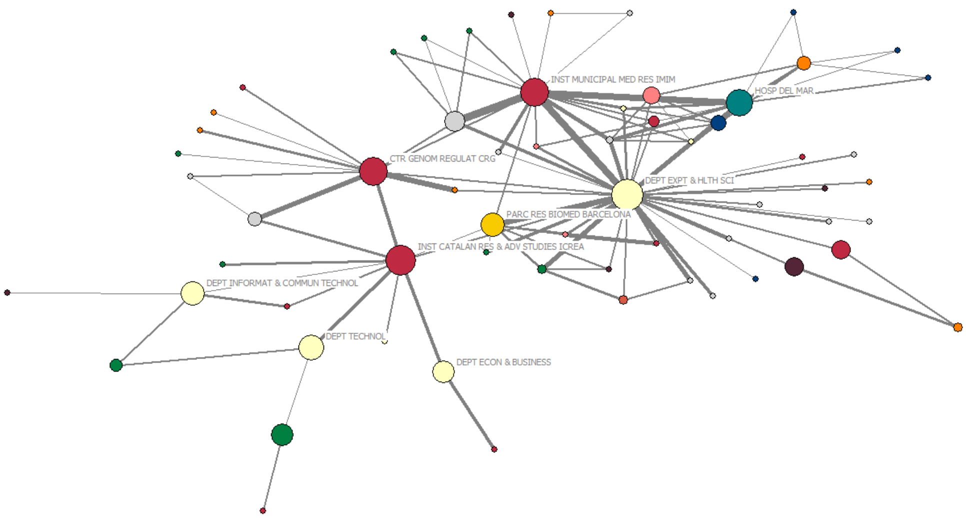


Table 5.4. Bibliometric indicators and research profiles of organizational from Pompeu Fabra University with >50 publications for the 2006-2010 time period and their output distribution according to two classification systems: TS subject categories and I-UGR disciplines

			TS CLASSIFICATION		I-UGR CLASSIFICATION	
Department	**P**	**B**	**G**	**No SC**	**G**	**No disc**
DEPT EXPT & HLTH SCI	599	1257.28	0.67	88	0.70	21
DEPT ECON & BUSINESS	206	61.00	0.67	53	0.71	17
INST MUNICIPAL MED RES IMIM	166	362.31	0.59	55	0.64	14
CTR GENOM REGULAT CRG	155	361.48	0.67	33	0.62	10
INST CATALAN RES & ADV STUDIES ICREA	122	618.83	**0.45**	51	**0.34**	14
DEPT TECHNOL	119	148.00	0.59	44	0.70	15
DEPT INFORMAT & COMMUN TECHNOL	96	90.00	0.54	43	0.66	13
UNIT BIOMED INFORMAT GRIB	88	56.90	0.60	35	0.55	10
UNIT EVOLUT BIOL	64	0.00	0.64	17	0.65	8
LAB NEUROPHARM	54	1.00	0.66	12	0.56	6
PARC RES BIOMED BARCELONA	52	73.87	0.50	35	0.57	13
HOSP DEL MAR	51	195.96	0.50	27	0.73	8

In bold when G <0.5

According to their research profile, none of these organizational units show values above 0.8 on their Gini Coefficient for any of the classifications used, distributing their research output in a wide range of subject fields. In fact, we observe that two units show Gini values under 0.5 according to the TS classification: INST CATALAN RES & ADV STUDIES ICREA (0.45) and HOSP DEL MAR (0.50). On the first case, this is quite normal as this institution is a multidisciplinary agency from the regional government focused on recruiting international researchers and integrating them in research centers and universities located within Catalonia. The second case is a hospital and it responds reasonably well. As observed, although it shows a low Gini coefficient when using the TS classification, it value raises up to 0.730 when using the I-UGR classification. Finally, we find that the betweenness centrality of the organizational present in Table 5.4 are much higher than in the previous case, with some exceptions, showcasing a much more integrating and multidisciplinary structure of university.

5.5. Discussion

In this paper we highlight the problems that may arise when interpreting university rankings by fields as these are commonly mistaken with organizational units within the structure of universities. Also, we propose the use of the Gini Index and the betweenness centrality measure as a means to understand how well are different organizational units represented by the field classification systems employed in bibliometric studies and rankings by fields. For this purpose we focus on two Spanish universities as case studies, Granada and Pompeu Fabra, which reflect two different types of institutions. Granada represents a historical university with a well-established structure while Pompeu Fabra represents a young and dynamic institution with an outstanding research performance. Then we develop a research profile for each department/organizational unit according to two different classification systems (TS subject categories and I-UGR Rankings disciplines) in order to showcase the discrepancies between the organizational units and the fields of each classification system.

Before discussing our results, we must emphasize on the implicit problems that working with addresses brings to any bibliometric analysis when adopting a *top down* approach. As it has been acknowledged elsewhere (Waltman et al., 2012), identifying institutions based on the address field of TS Web of Science means to inevitably assume some errors on the data retrieval of academic institutions. In this study we have shown that the problem may be even worse when deepening on organizational units within universities, leading to the need of manual data cleaning. Although many efforts have been done on standardizing and automatically retrieving address data (a good overview is included in Cuxac, Lamirel & Bonvallot, 2013), still the problem remains unsolved, especially at a large scale where first-hand institutional information is needed in order to verify the data provided from the database (van Leeuwen, 2007).

If solved, bibliometric analyses and rankings could greatly benefit from this kind of approach. As we see in Figures 5.3 and 5.5, no only it is possible to understand and analyze the structure of universities according to address data, but also organizational units group themselves according to fields. The structure and size of universities varies significantly due to the managerial changes that have taken place at the end of the 20th century, influenced by different socio-economic factors such as the expansion of higher education and the demand for return on investment, largely exemplified by the organizational forms defined by Gibbons et al. (1994). These changes prevent from the use of address data to construct fields as proposed by De Bruin & Moed (1993). In this sense, we find notable differences between the structure of each university, especially on the loss of importance faculties as an organizational form play in Pompeu Fabra and an increasing importance of departments along with research centers as the main joints in the university. In fact, departments in this university do not seem to be any longer the basic administrative unit and are replaced in such function by research groups, labs and units. Also, as observed in Table 5.4, we find that many organizational units behave as expected according to mode 2 and are 'based outside the university and its traditional disciplinary structure' (Morris, 2002). On the other side we find that Granada still obeys to such a disciplinary structure and in fact, higher values of concentration can be observed when developing research profiles for each department (Table 5.3).

Finally, we must emphasize the clarity with which the Gini coefficient along with the betweenness centrality value and the number of subject fields in which each organizational unit performs, reflect the levels of discrepancy between organizational units and the classification system used. Hence, we highlight the importance of university rankings by fields to provide clear instructions on the classification system used along with the necessary tools so that such profiles can be easily developed by third parties.

5.6. Concluding remarks and further research

The need for accurate and reliable data is a key issue when developing bibliometric tools and studies, and is in fact, one of the main weaknesses of university rankings (van Raan, 2005). The main problem is located in the use of bibliographic data which was not originally conceived to be used for bibliometric purposes and hence lacks of the standardization needed for this type of analyses. The rise of rankings has also raised other more fundamental questions which are still unsolved and which should be addressed before attempting at any institutional comparisons, such as what is a university? What does it mean to belong to a university? How should mixed institutions with more than on affiliation be treated? It seems that the only reasonable way to certify the accuracy and reliability of such studies is having some kind of output verification from the institutions involved. In this sense, further research is needed on analyzing the congruence between the affiliation of authors and the one they indicate in the address field, however, for this type of study to take place we would need internal information from the institution which is normally unavailable. But there is much at stake and the increasing need to offer global products which show where institutions stand at an international level leaves little choice but to assume these problems.

Rankings by fields intend to take into account the disciplinary focus of universities. But because they lack the proper information regarding the structure of universities, they are obliged to use other classification systems based on universities' research output. However, this misleads the user of rankings who expects to see on those fields a reflection of the structure of their universities. We have noted that for rankings offering a wide range of league tables by field, not only the media but also research managers tend to confound these with organizational units. Believing that a ranking on a specific field to be a ranking of faculties for instance, or not understanding why a given university can gain certain positions in fields which are not reflected on their institutional structure. In this paper we provide three measures for representing the research profiles of organizational units as a possible solution to show the levels of discrepancy of the fields offered by rankings and the structure of universities.

References

Borgatti, S.P., Mehra, A., Brass, D.J., & Labianca, G. (2009). Network analysis in the social sciences. *Science, 323*, 892-895.

Bornmann, L., Moya-Anegón, F., & Mutz, R. (2013). Do universities or research institutions with a specific profile have an advantage or a disadvantage in institutional rankings? A Latent Class Analysis with data from the SCImago Ranking. *Journal of the American Society for Information Science and Technology,* doi:10.1002/asi.22923

Bornmann, L., Mutz, R., & Daniel, H.-D. (2013). Multilevel-statistical reformulation of citation-based university rankings: The Leiden Ranking 2011/2012. *Journal of the American Society for Information Science and Technology, 64*(8), 1649-1658.

Bornmann, L., Stefaner, M., Moya-Anegón, F., & Mutz, R. (2013). Ranking and mapping of universities and research-focused institutions worldwide based on highly-cited papers: A visualization of results from multi-level models. *Online Information Review,* arXiv:1212.0304

Calero-Medina, C., & van Leeuwen, T.N. (2012). Reorganizing research with the help of bibliometric collaboration networks. Case study in a University Hospital. In: C. Calero-Medina, *Links in science: linking network and bibliometric analyses in the study of research performance* [Doctoral thesis] (pp. 45-66). Leiden: Universiteit Leiden.

Cheng, Y., & Liu, N.C. (2006). A first approach to the classification of the top 500 world universities by their disciplinary characteristics using scientometrics. *Scientometrics, 68*(1), 135-150.

Cuxac, P., Lamirel, J.-C., & Bonvallot, V. (2013). Efficient supervised and semi-supervised approaches for affiliations disambiguation. *Scientometrics.* doi: 10.1007/s11192-013-1025-5.

De Bruin, R.E., & Moed, H.F. (1993). Delimitation of scientific subfields using cognitive words from corporate addresses in scientific publications. *Scientometrics, 26*(1), 65-80.

Deaton, A. (1997). *Analysis of household surveys.* Baltimore: Johns Hopkins University Press.

García, J.A., Rodriguez-Sánchez, R., Fdez-Valdivia, J., Robinson-Garcia, N., & Torres-Salinas, D. (2012). Mapping academic institutions according to their journal publication profile: Spanish universities as a case study. *Journal of the American Society for Information Science and Technology, 63*(11), 2328-2340.

Glänzel, W., & Schubert. A. (2003). A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics, 56*(3), 357-367.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzmann, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: the dynamics of science and research in contemporary societies*. London: Sage.

Leydesdorff, L., & Opthof, T. (2010). Scopus's source normalized impact per paper (SNIP) versus a journal impact factor based on fractional counting of citations. *Journal of the American Society for Information Science and Technology, 61*(11), 2365–2369

López-Illescas, C., Moya-Anegón, F., & Moed, H.F. (2011). A ranking of universities should account for differences in their disciplinary specialization. *Scientometrics, 88*(2), 563-574.

Moed, H.F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics, 4*(3), 381-422.

Morris, N. (2002). The developing role of departments. *Research Policy, 31*(5), 817-833.

Praal, F., Koosten, J., Calero-Medina, C., & Visser, M.S. (2013). Ranking universities: The challenge of affiliated institutes. Accepted at the *18th International Conference on Science and Technology Indicators*.

Pudovkin, A.I., & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technology, 53*(13), 1113-1119.

Rescher, N. (2006). *Epistemetrics*. Cambridge: Cambridge University Press.

Robinson-García, N., Moreno-Torres, J.G., Torres-Salinas, D., Delgado López-Cózar, E., & Herrera, F. (2013a). The role of national university rankings in an international context: the case of the I-UGR Rankings of Spanish Universities. Conference Proceedings from the *14th International Society for Scientometrics and Informetrics Conference*, 2, 1550-1565. arXiv:1305.1216.

Robinson-García, N., Rodriguez-Sanchez, R., García, J.A., Torres-Salinas, D., & Fdez-Valdivia, J. (2013b). Análisis de redes de las universidades españolas de acuerdo a su perfil de publicación en revistas por áreas científicas. *Revista Española de Documentación Científica,* in press.

Torres-Salinas, D., Moreno-Torres, J.G., Delgado López-Cózar, E., & Herrera, F. (2011). A methodology for Institution-Field ranking based on a bidimensional analysis: The IFQ^{2}A-Index. *Scientometrics, 88*(3), 771-786.

Van Leeuwen, T.N. (2007). Modeling of bibliometric approaches and importance of output verification in research performance assessment. *Research Evaluation, 16*(2), 93-105.

Van Raan, A.F.J. (2005). Fatal Attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics, 62*(1), 133-143.

Waltman, L., van Eck, N.J., van Leeuwen, T.N., Visser, M.S., & van Raan, A.F.J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics, 5*(1), 37-47.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E., Tijssen, R.J.W., van Eck, N.J., et al. (2012). The Leiden Ranking 2011/2012: data collection. indicators. and interpretation. *Journal of the American Society for Information Science and Technology, 63*(12), 2419-2432.

Waltman, L., & van Eck, N.J. (2012a). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology, 63*(12), 2378-2392.

Waltman, L., & van Eck, N.J. (2012b). Source normalized indicators of citation impact: an overview of different approaches and an empirical comparison. *Scientometrics,* doi: 10.1007/s11192-012-0913-4.

6. On the use of biplot analysis for multivariate bibliometric and scientific indicators

Paper published in Journal of the American Society fo Information Science and Technology doi:10.1002/asi.22837 authored by D. Torres-Salinas, N. Robinson-García, E. Jiménez-Contreras, F. Herrera and E. Delgado López-Cózar

Bibliometric mapping and visualization techniques represent one of the main pillars in the field of scientometrics. Traditionally, the main methodologies employed for representing data are Multi-Dimensional Scaling, Principal Component Analysis or Correspondence Analysis. In this paper we aim at presenting a visualization methodology known as Biplot analysis for representing bibliometric and science and technology indicators. A Biplot is a graphical representation of multivariate data, where the elements of a data matrix are represented according to dots and vectors associated with the rows and columns of the matrix. In this paper we explore the possibilities of applying the Biplot analysis in the research policy area. More specifically we will first describe and introduce the reader to this methodology and secondly, we will analyze its strengths and weaknesses through three different study cases: countries, universities and scientific fields. For this, we use a Biplot analysis known as JK-Biplot. Finally we compare the Biplot representation with other multivariate analysis techniques. We conclude that Biplot analysis could be a useful technique in scientometrics when studying multivariate data and an easy-to-read tool for research decision makers.

6.1. Introduction

Bibliometric mapping and visualization techniques represent one of the main pillars in the field of scientometrics. Nevertheless, Derek de Solla Price, considered as the father of scientometrics, already stated his wish to "exhibit an interlocking metabolic complex of bibliometric (and scientometric) parameters in a comprehensive and integrated structure after the manner of the Nitrogen Cycle" (Price as cited by Wouters, 1999). Since this statement, this research front has greatly expanded, especially in the seventies and eighties and was revitalized again in the late nineties due to technological advancements, as a tool for research policy monitoring (Noyons, 2001). The use of science maps has long been discussed in literature, emphasizing its capability as an easy-to-read tool that enables decision makers to understand the complexity and heterogeneity of scientific systems in order to rapidly respond to their behavior (Noyons & Calero-Medina, 2009).

Visualizing bibliometric data with scientific maps allows a better understanding of the relation between disciplines, invisible colleges or research fronts, for instance. According to Klavans & Boyack (2009), scientific maps can be defined as a two-dimensional representation of a set of elements and the relationship among them. Following this line of thought, for scientific mapping two techniques must be applied: firstly, a classification methodology, and secondly, a representation technique. Traditionally, the main classifying methodologies employed for representing bibliographic data have been those based on multivariate analysis such as Multi-Dimensional Scaling (MDS), Principal Component Analysis (PCA) or Correspondence Analysis, for instance. A review on the application of these methodologies for scientific mapping can be found in Börner, Chen & Boyack (2003). However, not many representation techniques have been used; focusing especially on Pathfinder Networks (PFNet) (White, 2003), Self-organizing maps (SOM) (Moya-Anegón, Herrero-Solana & Jiménez-Contreras, 2006) or social networks (Groh & Fuchs, 2011). Drawing a low-dimensional graph implies the loss of some of the information inherent not just to the represented elements, but also to the variables that affect their similarity or disimilarity.

Regarding these techniques, in this paper we aim at presenting a visualization methodology known as Biplot analysis (Gabriel, 1971) which could introduce interesting and useful novelties in scientific maps, opening new possibilities in the field of scientometrics. A Biplot is a graphical representation of multivariate data, where the elements of a data matrix are represented according to dots and vectors associated with the rows and columns of a matrix. Contrarily to a scatter gram, the axes are not perpendicular, as they simulate the projection of an n-dimensional representation over a surface with a minimum loss of information, adding interpretative meaning to the cosine of the angles between vectors, which represents the correlation between variables. Therefore, when vectors are perpendicular, the cosine equals zero and the variables are independent. But if they are very close or represent a 180° angle, they have a highly positive or negative correlation.

In short, the Biplot analysis is a graphical representation of multivariate data that mixes variables and cases (that is the reason for the bi prefix), enabling the user, to intuitively interpret for example in a bibliometric context; indicators and cases. Not as widely expanded as other techniques such as the above mentioned, it was first proposed by Gabriel (1971) and has already been tested in its many variants and types in very different scientific fields such as: Medicine (Gabriel, 1990), Genetics (Wouters et al, 2003), Agriculture (Yan et al, 2000), Library Science (Veiga de Cabo & Martín-Rodero, 2011), Economics and Business (Galindo, Vaz & Nijkamp, 2011), Tourism (Pan, Chon & Song, 2008) or Political Science (Alcántara & Rivas, 2007). Within the field of bibliometrics, this methodology was first introduced in conference paper in which the Biplot analysis was applied in order to analyze the scientific activity in Health Sciences of a small set of Spanish universities (Arias Díaz-Faes et al, 2011).

Considering the success and expansion the Biplot methodology has had in other research areas, the main objective of this paper is to deepen into the possibilities of applying the Biplot analysis in the field of scientometrics. More specifically, we aim at firstly describe and introduce this methodology to the reader and secondly, analyze its usefulness through three different case studies, showing its easy use for understanding and reading multivariate data in a research policy context. These case studies are chosen in order to explore the methodology's strengths and weaknesses when using different contexts, types of variables and levels of analysis. Then we use the first case study in order to compare this methodology with CA, MDS and PCA. The case studies proposed are the following:

- The first case study reflects the scientific efforts of European countries and their performance considering several bibliometric and S&T indicators.

- The second study will analyze the top 25 countries in the THE Ranking according to their performance in four of the variables it uses for ranking universities.

- Thirdly, we analyze a Spanish university's research performance in different research fields according to its output in the Thomson Reuters Web of Science databases.

This paper is structured as follows. In Section 6.2 we present and describe the classic Biplot methodology. Then, we describe three case studies, for which we will apply this representation method, for this, we select the JK-Biplot type. The results of these three cases along with a comparison with other multivariate techniques are shown and discussed in Section 6.3. In Section 6.4 we conclude with some remarks on the strengths and weaknesses of this technique. Appendix 'Biplot methodology in terms of spectral decomposition' has been included at the end of the paper in order to provide a more thorough description of the Biplot methodology.

6.2. Methodology

In this section we will present the Biplot analysis and briefly introduce three case studies in which we will apply it. This section is structured as follows. Firstly we give an overview on the Biplot analysis. In subsection 6.2.2, we give the key points for interpreting a Biplot representation and we introduce the JK-Biplot based on PCA, which is the one we will use for presenting the application of this methodology in the field of scientometrics. In subsection 6.2.3. we shortly introduce the software used for developing our applications. Then, in subsection 6.2.4., we introduce the three case studies used.

6.2.1. A snapshot on the Biplot analysis

As we have previously mentioned, Biplot is a data representation technique consisting on visualizing a matrix with more than two variables in a low dimensional graph where each row represents a subject and each column a variable. This technique is usually applied after a multivariate analysis has been performed, ranging from log-ratio analysis, principal component analysis or correspondence analysis; in fact to any method based on a singular-value decomposition. Due to its simplicity, its potentiality lies on enabling to visualize not just the relation between subjects or cases considering certain variables, but also the relationship between the variables.

Gabriel originally described three types of Biplot analysis, considered as the classical ones (Cárdenas et al, 2007) depending on the quality of representation of cases and variables. Therefore, we have: the GH Biplot Analysis, which emphasizes variables' representation, the JK Biplot Analysis, focused on the represented elements, and the SQRT Biplot Analysis, which tries to balance the quality of representation of the overall matrix. Other types of Biplot analysis are HJ Biplot analysis (Galindo, 1986) and GGE Biplot analysis (Yan et al, 2000).

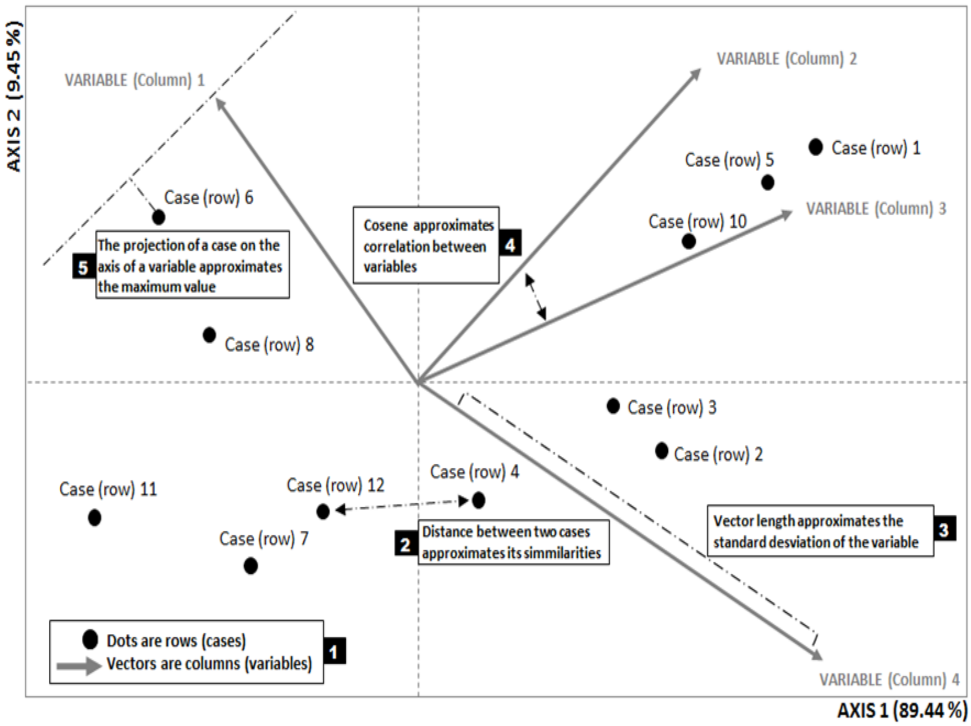
The Biplot is based on the same principles as other factorial techniques for dimensionality reduction, with the only difference that in this case, it represents the data but also the variables, obtaining a dual representation between principal components and the main coordinates. Its interpretation is based upon geometric concepts which are intuitive for the user, facilitating their understanding. In Figure 6.1 the basic ideas for interpreting a Biplot representation are explained:

- The similarity of subjects (rows) is the inverse function of the distance between them.

- The length and angles of the vectors (columns) represent variance and covariance respectively.

- The relation between rows and columns must be understood as dots products, that is, the projection of the cases over the variables.

Figure 6.1 Basic interpretation of a Biplot representation



Following this Figure we shortly introduce the 5 elements to take into consideration in the future analysis:

1. Dots are rows (cases) and vectors are columns (variables).

2. The distance between two cases approximates its similarity.

3. The vector length approximates the standard deviation of the variables.

4. The cosine between two vectors approximates the correlation between variables.

5. The projection of a case on the axis of a variable approximates the maximum value.

6.2.2. Biplot methodology

A Biplot is defined as a low-dimensional graph with a minimum loss of information of a given matrix of data $X_{(n \times p)}$, formed by markers $a_1, a_2, \ldots, a_n$ for rows and $b_1, b_2, \ldots, b_p$ for columns, chosen in such a way that each element x_{ij} , is an approximation to $x_{ij} = a_i^T b_j$ (Gabriel, 1971). In this subsection we will focus on providing clear rules for interpreting a Biplot representation. For a more exhaustive presentation of this methodology in terms of spectral decomposition, the reader is referred to Appendix Biplot methodology in terms of spectral decomposition.

The Biplot methodology offers approximate representations in a plane for data matrices with more than two dimensions that would otherwise, have to be represented in n-dimensions being n the number of variables. Variables are represented by linear axis with scales in the same way as in a normal scatter gram. Markers are located by projecting their mark perpendicularly onto the axes for variables (columns) and reading the value on the scale. These projected scale values are approximations of the true values as it is not usually possible to represent more than two variables exactly in the plane. Coordinates of markers are obtained from a PCA or a CA for instance, where the position of a marker is defined by the first two principal components. Also, the coordinates of variables are obtained with respect to the first two principal components, each weighted by the standard deviation of that component that is by the square root of the corresponding eigenvalue.

As observed in Figure 6.1, any two correlated variables are represented with their biplot axes pointing to similar directions, as markers with a high or low value for one of the correlated variables will have similar values for the other variable. On the contrary, if variables are correlated negatively, markers with a high value for one of the variables will presumably have a low value for the other variable. This means that correlation between variables can be obtained from the angle they form. Therefore, an acute angle between variables will presume a positive correlation among them; an obtuse angle will presume a negative correlation; and a right angle, no correlation between variables. These correlations are approximately represented by means of the cosines of the angles.

Another important aspect when interpreting a Biplot representation has to do with the display of the axes. Normally, these meet at the centroid which is the mark for the means of all the variables. Also, the length of the vectors (variables) is significant, as it displays the approximate value of the standard deviation of the variables. Depending on the preservation of columns or rows during the factorization we may have a Row Metric Preserving (RMP) Biplot or a Column Metric Preserving (CMP) Biplot. This two types are called a JK-Biplot and a GH-Biplot respectively and their main differences have to do with their emphasis for better representing rows than columns (JK-Biplot) and viceversa (GH-Biplot). In order to produce a symmetric Biplot we would need to balance the preservation values for columns and rows, this is what is called a SQRT Biplot.

In this paper we will use the JK-Biplot in order to explore its possibilities as it is the most common type. Its main feature is that the scalar product of the markers reproduces the matrix element. This concept is fundamental to geometrical interpretation in terms of distances, angles, orthogonal, etc.

Let consider a given set of data where the markers for rows and columns in a s dimension are:

$$A_{(s)} = J_{(s)} = U_{(s)}\Lambda_{(2)}B_{(s)} = K_{(s)} = V_{(s)}$$

This variant of Biplot analysis presents the following advantages.

Firstly, dot products with identical metric from rows of matrix X, coincide with the dot products of markers contained in J. The approximation of these dot products in a low-dimensional graph is optimal considering their minimum squares. In fact:

$$XX' = JK'KJ' = JJ'$$

Also, the spectral decomposition of the dot products matrix between rows is also the decomposition of its singular values:

$$XX' = U\Lambda^2 U$$

then, the best approximation to range s is:

$$XX' = U_{(s)}\Lambda^2_{(s)}U'_{(s)} = J_{(s)}J'_{(s)}$$

which coincides with the one obtained in the Biplot of matrix X.

Consequently, the Euclidean distance between two rows of X coincide with the Euclidean distance between markers J.

Also, markers for rows coincide with the coordinates for each case in a principal components space:

$$XV_{(s)} = U\Lambda\, V'V_{(s)} = U_{(s)}\Lambda_{(s)} = J_{(s)}$$

This means we can study similarities between cases with a minimum information loss.

Secondly, markers for rows coincide with the coordinates assigned to each case in the principal component space. In order to demonstrate this property, let consider V a matrix containing vectors from S, then coordinates over the first s components can be described as:

$$XV_s = (UDV')V_s = U_sD_s = J_s$$

This means that, when the Euclidean distance is adequate for the analysis, one can study similarities among the cases according to their markers.

Thirdly, the coordinates for columns are projections over the original axes in the principal components space. That is, coordinates of the vectors that construct the canonical base can be described as an identity matrix I_p and the projection of these over the principal components spaces can be described as:

$$I_pV_{(s)} = V_{(s)} = K_{(s)}$$

This means that coordinates for columns fix the unit for prediction scales. This property allows interpreting coordinates as the correlation between the original variables and the axes.

Finally, the last property of the JK-Biplot has to do with the quality of the representation. As mentioned above, this type of Biplot represents better rows than columns, contrarily to the GH-Biplot which emphasizes columns over rows.

6.2.3. 'MultBiplot' Software

For this study we have used the free beta version of the software 'MultBiplot' developed by Vicente-Villardón (http://Biplot.usal.es/multBiplot). This program implements the experience of the 'Applied Statistics Group' at the University of Salamanca (Spain) in working on Biplot analysis. According its authors this software is conceived not to be "another Biplot program", but to fill the gap between the static pictures and a more dynamic visual interpretation. So it is specialized on improving the visualization of Biplot diagrams. In relation to the different Biplot techniques, this program contains the Classical Biplot (JK) as well as the HJ-Biplot proposed by Galindo (1986). From the users' viewpoint, the 'MultBiplot' software does not require any kind of special training or a long learning period, being highly recommended for those who want to learn this statistical technique.

6.2.4. Data source and indicators

Considering that the aim is to present the Biplot analysis representation technique, three basic study cases were chosen, representing three different research evaluation contexts. Although this technique is usually applied to large data collections, in this paper we chose cases with a smaller size in order to ease the interpretation of the representation to the reader. We selected the JK-Biplot type which emphasizes cases representation over variables, and we used Principal Component Analysis as a classification methodology and data reduction. The three cases selected were: scientific effort and bibliometrics indicators of European countries, top Universities in the THE Ranking and the University of Granada's research performance in 12 different scientific fields. The selected data sources and the variables for each case study are displayed in Table 1. For more specific data regarding goodness of fitness and quality of representation ($QR_{overall}$, QR_{col} and QR_{row}) for each case, the reader is referred to http://www.ugr.es/~elrobin/QR_On_the_use_of_Biplot.xlsx where an excel file can be obtained with all the details.

Table 6.1 Description of the indicators used in the three different study cases

Indicator / Measure	Definition*	Acronym	Source
CASE 1: Countries			
Share of human resources in S&T	Labor force working in S&T from the total share of a country	%HR	Eurostat
R&D expenditure (Millions of €)	Total budget of countries devoted to R&D activities	MILL €	Eurostat
R&D expenditure (Percentage of GDP)	Proportion of countries' Gross Domestic Product devoted to R&D activities	GDP	Eurostat
Total Researchers	Total number of professionals devoted to activities related with R&D	RES	Eurostat
Number of Citations	Total number of citations received by publications generated by each country according to the Scopus database	CIT	SJ&CR
Number of Citable Documents	Citable documents are considered those published by journals indexed in Scopus under the following document types: articles, reviews and conference papers	DOC	SJ&CR
Citation Average	Average of citations received per citable document	CAVG	SJ&CR
Normalized Citation Average	Ratio between the average scientific impact of an institution and the world average impact of publications	NCIT	SJ&CR
CASE 2: Universities			
Research	Volume, income and reputation	RESEARCH	THE Ranking
Citation	Research influence	CITATION	THE Ranking
International Outlook	Staff, students and research	INT OUTLOOK	THE Ranking
Teaching	Learning environment	TEACHING	THE Ranking
CASE 3: Scientific Fields			
Citation Average	Average of citations received per document	ACIT	Thomson Reuters
Percentage of Top Cited Papers	Share of the total output of a university included in the top 10% of the most highly cited documents in the field according to the national output	TOPCIT	Thomson Reuters
Percentage of Fist Quartile Papers	Share of documents published in journals ranked in the top 25% according to the Thomson Reuters Journal Citation Reports	%Q1	Thomson Reuters
Number of Citations	Total number of citations received by documents published by a university in a given field	NCIT	Thomson Reuters
H-Index (Hirsch)	Number of documents (h) published by a university in a given set that has received at least h citations	H-Index	Thomson Reuters
Number of Citable Documents	Citable documents are considered those published by journals indexed in Thomson Reuters Web of Science under the following document types: articles, reviews notes and letters	NDOC	Thomson Reuters

* Definitions for variables in case 2 are displayed as stated in http://www.timeshighereducation.co.uk/story.asp?storycode=417368

6.3. Analysis and results

In the following three subsections we present the analysis and results for each case study. Finally, we briefly compare the results of one of the study cases with those given by applying other techniques (PCA, MDS, CA) in order to show the advantages of the Biplot representation in comparison with other methodologies for interpreting multivariate data with more than two variables. Usually, these techniques join together the information given by the variables, introducing two artificial variables instead and therefore, losing some information in the representation.

6.3.1. Case 1. Scientific effort and bibliometrics indicators for European Countries

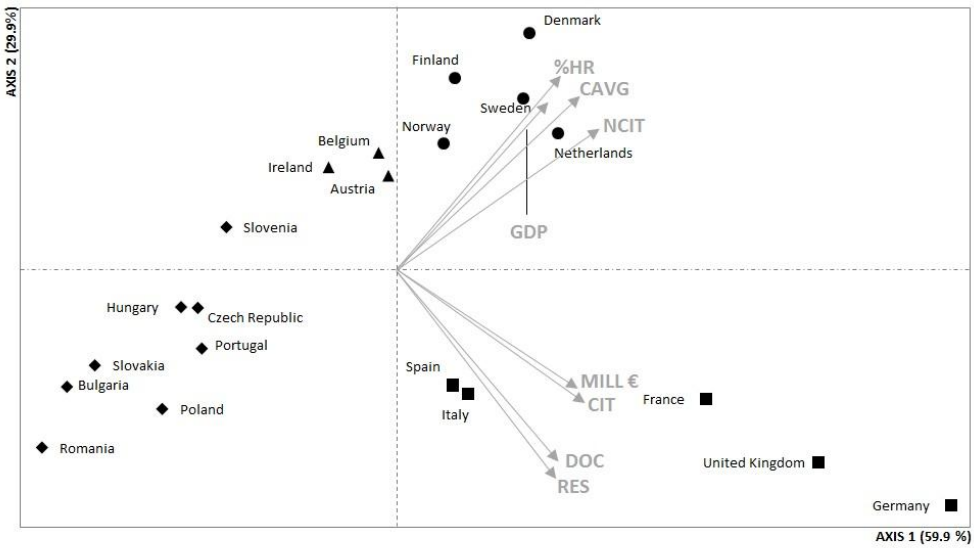
We analyze the research performance and input of a set of European countries. For this analysis we considered a 21x8 matrix where rows correspond to European countries and columns to indicators regarding R&D efforts and bibliometric indicators. The study time period used was 2009 or 2010. Data regarding R&D indicators was extracted from the EUROSTAT Portal, while bibliometric indicators were extracted and calculated from data retrieved from the Scimago Journal & Countries Rank databases. Countries and indicators are presented in table 6.2.

In Figure 6.2 we show the Biplot representation of this case. The goodness of fit is 89.9%. All variables (columns) are well represented as they all have a QR_{col} above 0.95 except GDP where it reaches 0.75. Rows are also well represented, 15 countries present a QR_{rows} above 0.90 and 6 between 0.73 and 0.86. Regarding the variables two latent variables can be clearly distinguished in the graph, indicating a high correlation between the observed variables of each of them. Therefore, the correlation between %HR and DOC is 0.198 and between CAVG and NCIT is 0.928. The first latent variables which encompasses Human resources (%HR), %GDP, average of citations (CAVG) and normalized citations (NCIT) could be defined as the qualitative axis as these measures are all normalized. The second latent variable, which is formed by variables related with raw indicators influenced by size (CIT, MILL €, DOC, RES) could be defined as one of a quantitative measure.

Table 6.2 Science & Bibliometrics for European Countries

	MILL €	GDP	RES	%HR	DOC	CIT	CAVG	NCIT
Germany	69810	2.82	484566	44.8	119216	228773	1.76	1.36
France	43633	2.26	295696	43.9	87430	148995	1.57	1.39
United Kingdom	30071	1.77	385489	45.1	123756	253482	1.81	1.42
Italy	19539	1.26	149314	33.8	67459	118043	1.6	1.23
Spain	14588	1.39	221314	39	59642	96368	1.48	1.10
Sweden	11869	3.42	72692	50.8	25257	54567	2.03	1.39
Netherlands	10769	1.83	54505	51.9	39499	96134	2.22	1.66
Austria	7890	2.76	59341	39.2	15476	31879	1.9	1.23
Denmark	7208	3.06	52568	51.9	15042	38504	2.38	1.60
Belgium	7047	1.99	55858	49.3	21978	46169	1.95	1.44
Finland	6971	3.87	55797	50.6	13308	25310	1.81	1.26
Norway	5342	1.71	44762	51.5	12755	22401	1.62	1.39
Ireland	2796	1.79	21393	45.9	9499	17728	1.73	1.24
Portugal	2747	1.59	86369	23.9	12957	16756	1.22	1.05
Poland	2607	0.74	98165	36.3	26057	23729	0.88	0.64
Czech Republic	2334	1.56	43092	37.8	13790	17005	1.18	0.77
Hungary	1126	1.16	35267	33	7542	10648	1.34	0.91
Slovenia	745	2.11	10444	40.8	4104	4697	1.1	1.05
Romania	572	0.47	30645	24.4	10897	6254	0.56	0.73
Slovakia	416	0.63	21832	33.5	4195	4043	0.93	0.72
Bulgaria	214	0.6	14699	31.6	3293	2285	0.68	0.74

Figure 6.2 JK-Biplot analysis for European Countries according to their Science & Bibliometric Indicators



In regard to the countries, we observe four distinct groups according to their scientific profile.

- There is a group formed by the Nordic countries (Norway, Sweden, Denmark and Finland) and the Netherlands (upper right), characterized as big investors in science (%HR and GDP) and with a high scientific impact (CAVG and NCIT).

- A second cluster can be observed (lower right) where countries such as Germany and United Kingdom and France perform well in all variables; effort and bibliometric indicators. A subset of this second group is formed for two Mediterranean countries; Spain and Italy, with lower values for normalized bibliometric indicators and less R&D efforts than the other members of this cluster and the first one.

- Another cluster can be found (upper left) formed by four small countries (Belgium, Ireland, Austria and Slovenia) characterized for a medium performance regarding R&D efforts and bibliometric indicators.

- Finally, we find countries (lower left), - mainly from east Europe as Bulgaria, Romania, Hungary, etc. - characterized by their low investment on R&D and their low research performance.

Consequently, we observe how this representation allows the reader to easily spot countries that are similar, not just regarding to their geographical location, but also to their scientific culture.

6.3.2. Case 2.Top Universities in the THE Ranking

We analyze 'world-class universities' performance according to the variables used in the Times Higher Education World University Ranking. We considered a 25x4 matrix where rows correspond to the top 25 universities from the 2012 and columns correspond to the different indicators and measures employed in this classification. That is: Teaching, Research, Citations and International Outlook. Industry Income was excluded for this analysis as data is not provided for all universities. A more thorough description of the methodology employed by this ranking is available at THE rankings website. Values for each university and variable are shown in Table 6.3. Figure 6.3 shows the Biplot representation.

Table 6.3 Top 25 universities according to the THE Ranking variables (data: 2012 edition)

	Teaching	International Outlook	Research	Citations
ETH Zürich -	79.1	97.5	85.8	87.2
Imperial College London	88.8	92.2	88.7	93.9
University of Oxford	89.5	91.9	96.6	97.9
University College London	77.8	91.8	84.3	89
University of British Columbia	68.6	88.7	78.6	85.2
University of Cambridge	90.5	85.3	94.2	97.3
Massachusetts Institute of Technology	92.7	79.2	87.4	100
University of Toronto	76.9	69	87.4	86.5
Columbia University	89.1	67.6	81.8	97.8
Harvard University	95.8	67.5	97.4	99.8
Georgia Institute of Technology	66.6	65	73.8	91.9
Johns Hopkins University	78.9	59.9	86.5	97.3
University of Chicago	89.4	58.8	90.8	99.4
Stanford University	94.8	57.2	98.9	99.8
California Institute of Technology	95.7	56	98.2	99.9
Yale University	92.3	55.5	91.2	96.7
Carnegie Mellon University	65.7	55	79.5	97.4
CornellUniversity	70.4	53.4	87.2	93.5
University of California Berkeley	82.8	50.4	99.4	99.4
Princeton University	91.5	49.6	99.1	100
University of Michigan	75.4	47.2	90	94.3
Duke University	62.6	46.9	77.9	97.4
University of California Los Angeles	85.9	41	92.5	97.3
University of Washington	70.8	36.9	74	98.2
University of Pennsylvania	87	34.3	86.1	97.9

The goodness of fit is 87.9%. Rows are represented with a QR_{row} above 90% for 17 universities, 80% for 3 universities and less than 75% for 5 universities. Michigan, MIT and Columbia have the lower QR_{row} as they have most of the information represented in axis 3 which is the one not covered in our biplot representation. In regard to columns, their QR_{col} is above 80% for all variables. When observing the overall representation, we must point out that, firstly, two variables do not correlate with the rest (Citations and International Outlook) and secondly, two other variables are very closely related to each other (Research and Teaching). In this last case the correlation value is 0.784. Regarding to the cases, there are four distinct clusters of universities.

- The first cluster (lower right) is formed by the universities with the highest values on Teaching and Research and which display a good performance in Citations. For instance, we see the two top British universities along with different universities from the North-American Ivy League such Harvard or

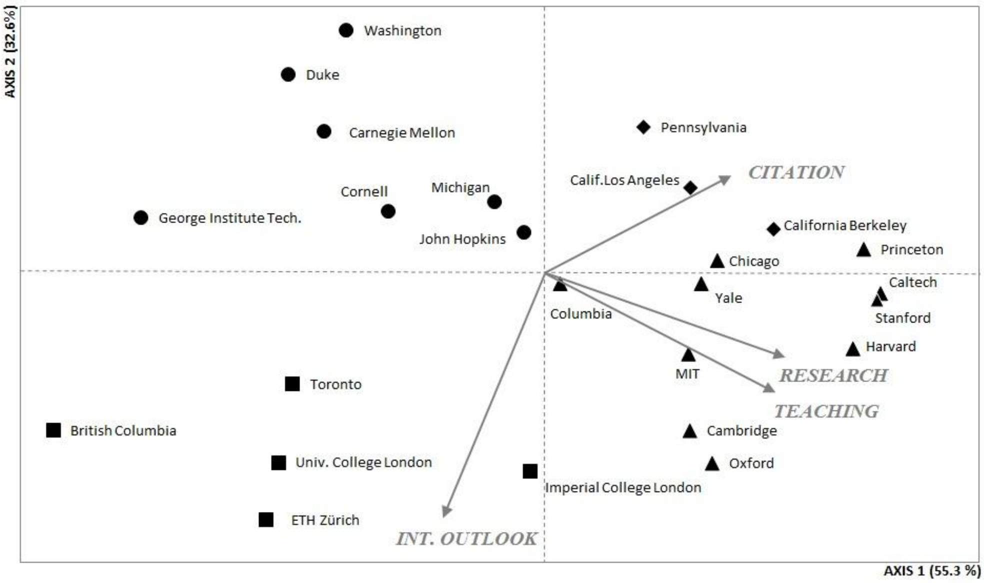
Yale, and universities from the West-Coast such California Berkeley or Caltech.

- Secondly, we find those universities which perform better in Citations but which are not in top positions in Teaching and Research, such as Pennsylvania and California Los Angeles.

- The third group (upper left) are universities that display the lowest performance in all indicators, such as Duke, Cornell or Michigan. This last group also coincides with the last top 25 universities in the THE Ranking.

- Finally the last group (lower left) is the one formed by those universities characterized mainly by their high values in International Outlook but not in the other indicators. We can distinct in this cluster the main universities from London (University College and Imperial College) and also from Canada (Toronto and British Columbia).

Figure 6.3 JK-Biplot Analysis for top 25 universities according to the THE Ranking



6.3.3. Case 3. Scientific performance of the University of Granada in 12 scientific fields

We analyze a university's research performance in 12 different scientific fields. For this, we selected the University of Granada (Spain) as a case study. We considered a 12x6 matrix where each row represents a scientific field and each column a bibliometric indicator regarding production and impact. Indicators were normalized according to all Spanish universities, meaning that the university with the best performance for a given indicator would reach a score of 1.00. We used the Thomson Reuters Web of Science databases and we selected 2006-2010 as the study time period. For more information over this data set, the reader is referred to Torres-Salinas et al. (2011a) and Torres-Salinas et al. (2011b). Indicators for each field of endeavor are shown in Table 6.4. In Figure 6.4 we illustrate the Biplot representation of this study case.

In this third case the goodness of fit is 72.2 %. It is the lower of three study cases presented. The QR_{row} is over 80% in 8 scientific fields but it is insufficient in one of the other three; Economics & Business where it is 47%. In this field, most of the information is represented the third axis, however, no variables are represented there. Therefore, no conclusion can be obtained for this field after interpreting Figure 6.4. A similar situation occurs with columns where the QR_{col} in five variables has a fit over 95% but one, %1Q, which is not well represented in axes 1 and 2. %1Q has a QR_{col} of 3%. Relating with the representation, we observe that variables/vectors are grouped into clusters according to their correlation. On the left side we find relative variables such as Top Cited Documents (TOPCIT) and Citation Average (ACIT) which are size independent. On the right side we find such as Number of Citations (NCIT), H-Index and Citable Documents (NDOC) which are related to the raw data. We find the highest correlation values between NCIT and H-Index with 0.822 and the lowest between H-Index and TOPCIT with a correlation value of -0.042.

When observing the University of Granada's behavior regarding each scientific field (cases), we must outline the following:

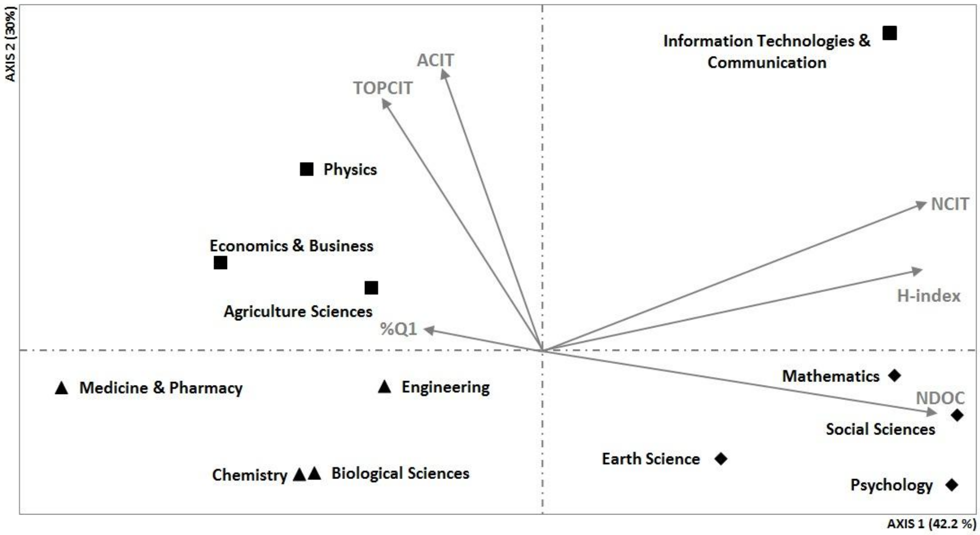
- Two latent variables emerge from the observed variables. As in case 1, we have on the one hand the qualitative axis formed by TOPCIT, ACIT and %Q1 and a quantitative axis formed by NCIT, H-index and NDOC.

- It is highly significant the position of the Information Technology & Communication field (upper right) which stands completely by itself and separate from the rest of the fields. This is due to the high values it has for indicators of both latent variables except for %Q1.

Table 6.4 Bibliometrics Indicators of the University of Granada in 12 Scientific Fields

	Bibliometrics Indicators						Normalized Indicators					
	NDOC	NCIT	H-Index	%1Q	ACIT	TOPCIT	NDOC	NCIT	H-Index	%Q1	ACIT	TOPCIT
Agricultural Sciences	174	821	14	72%	4.71	17%	0.352	0.408	0.737	0.885	0.854	0.733
Biological Sciences	958	5575	28	38%	5.81	8%	0.329	0.244	0.622	0.548	0.543	0.385
Earth Sciences	993	4567	23	54%	4.59	11%	0.729	0.577	0.742	0.891	0.658	0.579
Economics & Business	103	255	8	14%	2.46	18%	0.350	0.300	0.571	0.275	0.677	0.961
Physics	834	11763	28	62%	14.1	11%	0.374	0.577	0.560	0.793	1.000	0.662
Engineering	630	2699	22	61%	4.28	12%	0.320	0.381	0.733	0.844	0.465	0.643
Mathematics	777	1964	16	37%	2.52	10%	0.860	0.798	0.762	0.638	0.525	0.523
Medicine & Pharmacy	1412	8496	33	39%	6.01	10%	0.270	0.171	0.452	0.653	0.628	0.650
Social Sciences	263	503	11	30%	1.91	9%	0.809	0.652	0.917	0.523	0.584	0.315
Psychology	448	1477	16	23%	3.29	12%	0.911	0.652	0.800	0.376	0.456	0.335
Chemistry	1006	5595	26	58%	5.56	8%	0.376	0.262	0.591	0.813	0.534	0.379
Inf. Technology	502	2205	20	34%	4.39	19%	0.584	1.000	1.000	0.689	0.891	0.942

Figure 6.4 Biplot analysis of the University of Granada in 12 scientific fields according to bibliometric indicators



- On the lower right side we find those fields on which the University of Granada outstands at national and internal level for raw indicators such as NDOC, H-Index or NCIT, that is for the quantitative axis. For example the University Granada is the second and third most productive university in Mathematics and Earth Sciences respectively in Spain, explaining its high values for variable NDOC.

- On the upper left side we find those areas in which the university performs well for qualitative indicators. In this sense, we must emphasize Physics and Agricultural Science for two indicators; TOPCIT and ACIT. In the case of Physics, it shows the best performance for TOPCIT of all fields, as reflected in the biplot. We also find Economics along with the %Q1 variable which had been previously discussed and cannot be interpreted in this representation due to the lack of information.

- Finally, we find a fourth group of areas in which this University of Granada has the worst performance according to the indicators displayed, for instance, Chemistry or Engineering. In fact these fields are where Granada is positioned lower in national rankings.

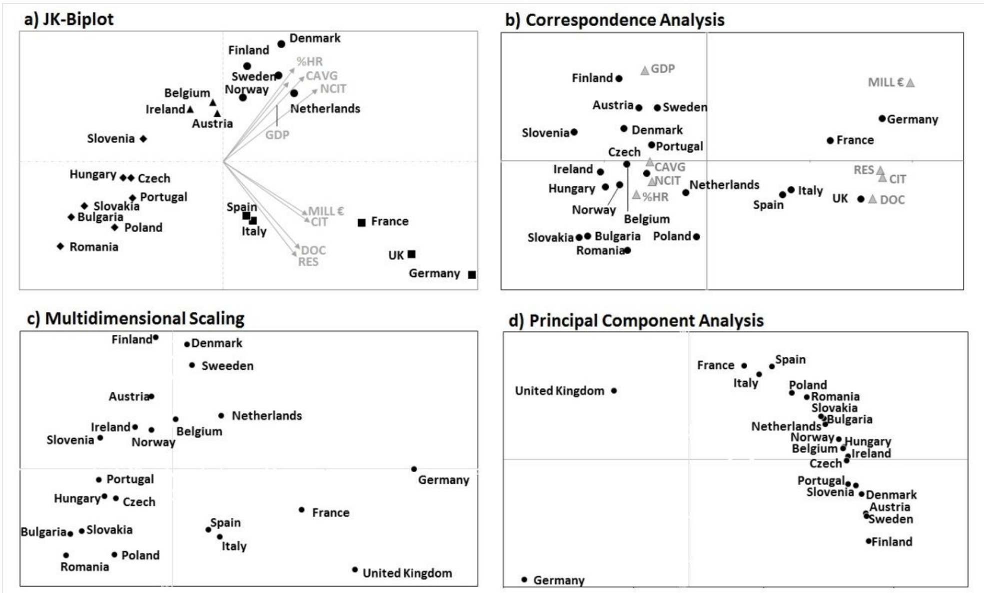
6.3.4. Comparing JK Biplot representation with other multidimensional representation techniques

Finally, in Figure 6.5 we present different visualization techniques applied to the first study case. Along with a JK Biplot representation we apply Correspondence Analysis (CA), Multidimensional Scaling (MDS) and Principal Component Analysis (PCA). We have chosen these techniques as they are the most common ones used for representing data in the field of bibliometrics. PCA is a mathematical methodology that uses orthogonal transformation converting a set of cases of possibly correlated variables in a set of values of uncorrelated variables which are known as principal components aiming at reducing the number of variables and guaranteeing that these are independent when data is jointly normally distributed. CA is a multivariate statistical methodology similar to PCA, providing the means to display and summarize a set of data in a two-dimensional graph. MDS is a visualization technique used for exploring similarities and dissimilarities in data. In the case of PCA and MDS we used the statistical software SPSS version 20.00. In the case of CA we used the statistical package XLSTAT and we used the Correspondence Factor Analysis with symmetric distances.

When comparing with MDS and PCA, Biplot representation offers a better solution, as the former are incapable of representing both, variables and cases, at the same time. However, even if it is done separately, MDS and PCA representations show similar patterns to those presented by the Biplot representation; with countries grouped in a similar way. For instance, the Biplot map and the MDS map show a very similar display of countries. Also, the PCA representation shows a similar pattern. In fact, the left corresponds with the lower right of MDS and Biplot with Germany and the UK outstanding, followed by France. The Nordic countries are displayed closely to each other as well as the pair Italy and Spain.

But if there is a method similar to the Biplot technique, that is the Correspondence Analysis (CA). This technique also represents rows and columns of a matrix, i.e. a contingency matrix, in a bidimensional graph. However, although the CA representation displayed in Figure 6.5 is similar to the Biplot map, we find it much more difficult to interpret as the relation between variables and cases is not perceived as easily as it occurs with the Biplot representation. Also, as it happened with the other two techniques, it offers a poorer representation losing much of the information, especially regarding the visualization of variables where the Biplot analysis displays their correlation between each other and their standard deviation. For these reasons many authors (Gabriel, 2002) point out the Biplot analysis as a good alternative instead of CA. We must take into account that both techniques are closely related as they both are based on the same assumption, that is, reducing the data dimensions with a minimum information loss.

Figure 6.5 Representation of the case I (countries) using different multivariate techniques



6.4. Conclusions

In this study we present a methodology for representing multivariate data in a low dimensional graph. Although many representation techniques have been applied in the field of scientometrics, emphasizing on analyzing their capability for representing with a minimum information loss multivariate data, Biplot analysis seems to be less known by this research community. We apply the JK-Biplot technique in three different case studies testing its efficiency in three different research evaluation contexts according to the aggregation levels (macro, meso and micro), different types of indicators (bibliometric and science indicators) and obtaining different results regarding the overall, row or column quality representation. We believe that, as well as it has been proved for other scientific fields, this methodology may well be an important analysis tool for bibliometric studies.

In this paper we focus on the Classical JK-Biplot analysis, however, other types of Biplot analysis should be studied in order to explore their possibilities and differences among each. We must especially mention the HJ-Biplot analysis as this type seems to overpass the limitations of the JK-Biplot analysis regarding the quality of representation for rows and columns. Although in this paper we have used small matrices for displaying the biplot analysis potential, we believe this type of analyses are of great interest and should be explored by the informetric research community, especially for studies regarding massive data sets for data mining (Theoharatos et al, 2007) and data classification patterns (Chapman et al., 2001). Finally, we must emphasize that, as well as other visual metaphors such as social networks analysis, this type of representations may be of great interest not just as research tools for analyzing variables, but also in the research policy arena as easy-to-read tools.

References

Alcántara, M., & Rivas, C. (2007). Las dimensiones de la polarización partidista en América Latina. *Política y Gobierno, 14*(2), 349-390.

Arias Díaz-Faes, A., Benito-García, N., Martín-Rodero, H., & Vicente-Villardón, J.L. (2011). Propuesta de aplicabilidad del método multivariante gráfico "BIPLOT" a los estudios bibliométricos en biomedicina. In: *XIV Jornadas Nacionales de Información y Documentación en Ciencias de la Salud, Cádiz, 13-15 de Abril de 2011* [Conference Poster]. From: http://hdl.handle.net/10760/15998

Börner, K., Chen, C., & Boyack, K.W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology, 37*, 149-255.

Cárdenas, O., Galindo, P., & Vicente-Villar don, J.L. (2007). Los métodos Biplot: evolución y aplicaciones. *Revista Venezolana de Análisis de Coyuntura, 13*(1), 279-303.

Chapman, S., Shenk, P., Kazan, K., & Manners, J. (2001). Using Biplots to interpret gene expression pattern in plants. *Bioinformatics Applications Note, 18*(1), 202-204.

Gabriel, K.R. (1971). The Biplot graphic display of matrices with application to principal component analysis. *Biometrika, 58*(3), 453-467.

Gabriel, K.R. (2002). Goodness of fit of biplots and correspondence analysis. *Biometrika, 89*(2), 423-436.

Gabriel, K.R., & Odoroff, C.L. (1990). Biplots in biomedical research. *Sta. Med., 9*, 469-485.

Galindo, P. (1986). Una alternativa de representación simultánea: HJ Biplot. *Qüestiió, 10*(1), 13-23.

Galindo, P.V., Vaz, T.D., & Nijkamp, P. (2011).Institutional capacity to dynamically innovate: an application to the Portuguese case. *Technological Forecasting and Social Change, 78*(1), 3-12.

Groh, G., & Fuchs, C. (2011).Multi-modal social networks for modeling scientific fields. *Scientometrics, 89*(2), 569-590.

Klavans, R., & Boyack, K.W. (2009). Towards a consensus map of science. *Journal of the American Society of Information Science and Technology, 60*(3), 455-476.

Moya-Anegón, F., Herrero-Solana, V., & Jiménez-Contreras, J. (2006). A connectionist and multivariate approach to science maps: the SOM, clustering and MDS applied to library and information science research. *Journal of Information Science, 32*(1), 63-77.

Noyons, E. (2001). Bibliometric mapping of science in a science policy context. *Scientometrics, 50*(1), 83-98.

Noyons, E., & Calero-Medina, C. (2009). Applying bibliometric mapping in a high level science policy context. *Scientometrics, 79*(2), 473-490.

Pan, S., Chon, K., & Song, H.Y. (2008).Visualizing tourism trends: a combination of ATLAS.ti and Biplot. *Journal of Travel Research, 46*(3), 339-348.

Theoharatos, C., Laskaris, N.A., Economu, G., & Fotopoulos, S. (2007). On the perceptual organization of image databases using cognitive discriminative biplots. *Eurasip Journal on Advances in Signal Processing*, 68165.

Torres-Salinas, D., Delgado-López-Cózar, E., Moreno-Torres, J.G., & Herrera, F. (2011a). Rankings ISI de las universidades españolas por campos científicos: Descripción y resultados. *El profesional de la información, 20*(1), 111-122.

Torres-Salinas, D., Delgado-López-Cózar, E., Moreno-Torres, J.G., & Herrera, F. (2011b). A methodology for Institution-Field ranking based on a bidimensional analysis: the IFQ2A Index. *Scientometrics, 88*(3), 771-786.

Van Eck, N.J., Waltman, L., Dekker, R., & van den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology, 61*(12), 2405-2416.

Veiga de Cabo, J., & Martín-Rodero, H. (2011). Acceso Abierto: nuevos modelos de edición científica en entornos web 2.0. *Salud Colectiva, 7*(Supl. 1), S19-S27.

White, H. (2003). Pathfinder networks and author cocitation analysis: A remapping of paradigmatic information scientists. *Journal of the American Society for Information Science and Technology, 54*(5), 423-434.

White, H., & McCain, K. (1998). Visualization of literatures. *Annual Review of Information Science of Science and Technology, 32*, 99-169.

Wouters, L., Gohlmann, H.W., Bijnens, L., Kass, S.U., Molenberghs, G., & Lewi, P.J. (2003). Graphical exploration of gene expression data: a comparative study of three multivariate methods. *Biometrics, 59*(4), 1131-1139.

Wouters, P. (1999). *The citation culture* [Doctoral Thesis]. Amsterdam: Universiteit van Amsterdam.

Yan, W. K., Hunt, L.A., Sheng, Q.L., & Szlavnics, Z. (2000).Cultivar evaluation and mega-environment investigation based on the GGE Biplot. Crop Science, 40(3), 597-605.

Appendix Biplot methodology in terms of spectral decomposition

A Biplot is defined as a low-dimensional graph with a minimum loss of information of a given matrix of data $X_{(n\times p)}$, formed by markers $a_1, a_2, \ldots, a_n$ for rows and $b_1, b_2, \ldots, b_p$ for columns, chosen in such a way that each element x_{ij}, is an approximation to $x_{ij} = a_i^T b_j$ (Gabriel, 1971).

Markers a_i for rows and markers b_i for columns are represented in a space of a dimension $s \leq \rho$ where s is the number of axes and ρ the range of X. Let $a_1, a_2, \ldots, a_n$ be markers for rows of matrix A and $b_1, b_2, \ldots, b_p$ markers for rows of matrix B, then:

$$X \cong AB'$$

where $\cong$ means that X approaches to the product from the right.

The structure of matrix X can then be visualized by representing the markers in a Euclidean space of s dimensions. When matrix X is of range 2 or 3, the representation can adjust perfectly to two or three dimensions; if not, we will need as many axes as the range of X. However, as mentioned above, a Biplot follows the same criterion as for factorial dimensional reduction techniques, therefore, only the two first axes are represented.

The markers are obtained firstly through Singular Value Decomposition (SVD) of matrix X and then, by factorizing the matrix as follows:

$A = U\Lambda^{\gamma}$ and $B = V\Lambda^{1-\gamma}$

where $0 \leq \gamma \leq 1$. Gabriel (1971) proposes different γ to which he assigns different names. Two possible factorizations are:

$$X = A^0(B^*)' = A^*(B^0)'$$

Row Metric Preserving (JK Biplot): $A^* = U\Lambda$ and $B^0 = V$

Column Metric Preserving (GH Biplot): $A^0 = U$ and $B^* = V\,\Lambda$

Then, using the two or three first columns for factorizations of matrices A and B, we obtain biplots in two or three dimensions. Row Metric Preserving (RMP) and Column Metric Preserving (CMP) refer to the preservation of rows or columns' metrics during factorization. Each factorization has a "principal factor" that emphasizes the singular values and a "standard factor" for which the singular values do not appear. In order to identifying them we use the (*) and (0) respectively.

When we use $\gamma = 1/2$ in the equations:

$A = U\Lambda^{1/2}$and $B = V\Lambda^{1/2}$

we obtain a symmetric Biplot or SQRT Biplot where $AA = I$.

One of the most important aspects one must take into account when analyzing Biplot representations are the concepts of Quality of Representation (hereafter QR) which is referred to each row and column, and the Goodness of Fitness ($QR_{overall}$), which is defined as the cumulative qualities of representation for columns. Usually, a range of representation higher than two is used. Although a Biplot representation may have a high Goodness of Fit, this does not necessarily mean that a certain marker may be represented with a low QR. Regarding goodness of fit for variables and cases, Gabriel (2002) uses a function depending on the two first eigenvalues and the Biplot classification methodology used. In his case, he uses Correspondence Analysis and shows that such function is a good indicator for SQRT and only for GH and JK when values are close to 0.95.

7. Mapping academic institutions according to their journal publication profile: Spanish universities as a case study

Paper published in Journal of the American Society for Information Science and Technology doi:10.1002/asi.22735 authored by J.A. García, R. Rodríguez-Sánchez, J. Fdez-Valdivia, N. Robinson-García and D. Torres-Salinas

We introduce a novel methodology for mapping academic institutions based on their journal publication profiles. We consider that journals in which researchers from academic institutions publish their works can be considered as useful identifiers for representing the relationships between these institutions and establishing comparisons. But, when academic journals are used for research output representation, distinctions must be introduced between them, based on their value as institution descriptors. This leads us to the use of journal weights attached to the institution identifiers. Since a journal in which researchers from a large proportion of institutions published their papers may be a bad indicator of similarity between two academic institutions, it seems reasonable to weight it in accordance with how frequently researchers from different institutions published their papers in this journal. Cluster analysis can then be applied in order to group the academic institutions, and dendrograms can be provided to illustrate groups of institutions following agglomerative hierarchical clustering. In order to test this methodology, we use a sample of Spanish universities as a case study. We first map the study sample according to institutions' overall research output, and then we use it just for two scientific fields (Information and Communication Technologies, as well as Medicine and Pharmacology) as a means to demonstrate how our methodology cannot only be applied for analyzing institutions as a whole, but also in different disciplinary contexts.

7.1. Introduction

Over the last decade a great deal of interest has been focused on scientific mapping and visualization. Although first conceived as tools for displaying the structure and dynamics of research activity, they have now been fully integrated into research evaluation (Noyons, Moed & Luwel, 1999) and combine structural and performance information that enables them as easy-to-read tools for research policy makers (Torres-Salinas, 2009). According to Klavans and Boyack (2009) a map of science can be defined as a set of elements and the existing relationships between them, considering as an element any unit of representation of science such as scientific fields, publications, or researchers. They are characterized by visualizing these elements, commonly represented in a two or three-dimensional space, and by matching pairs of elements according to their common characteristics. Science maps, also known as Atlas of Science, are commonly visualized as node-edge diagrams similar to those used in network science and they aim at analyzing the structure of science based mainly in research publications. First attempts to mapping science by applying bibliometric techniques can be traced to Henry Small and his colleagues (Griffith, Small, Stonehill & Dey, 1974; Small, 1999; Small & Garfield, 1985). These techniques vary from each other depending on the methodological choices and on the unit of analysis used.

Although first efforts were made on generating maps based on scientific papers, journals have also been used as a basic unit for mapping science for some 35 years, starting with the pioneering map by Narin, Carpenter and Berlt (1972). These maps are normally generated in two steps. Firstly, a clustering method is used for dividing journals into a number of clusters. The decision made on how these clusters are formed will determine the validity of the whole process as it will define the criteria followed for considering the elements as similar or dissimilar (Gmür, 2003). Secondly, a visualization algorithm is developed in order to generate a layout of the clusters previously formed. In a different approach, Moya-Anegón et al. (2004; 2007) introduced discipline-based maps using the Thomson Reuters subject categories system aiming at a rather ambitious goal such as representing the world's research output. Also Leydesdorff & Rafols (2009) use the Thomson Reuters subject categories for representing science in order to analyze the structure of the Science Citation Index database. Despite technological limitations at first, the emergence since the mid 1990s of new visualization tools and the availability of large amounts of data on scientific publications made possible a further development of this type of maps (Noyons, 2004). Regarding mapping institutions or universities, main efforts have been focused using research collaboration as a means for establishing networks between them (Leydesdorff & Persson, 2010; Rorissa & Yuan, 2012) or web links (Ortega, Aguillo, Cothey & Scharnhorst, 2008), but other than that no other technique has been used. This kind of techniques allow readers to rapidly learn over scientific, geographical, or social connections between different institutions, emphasizing relations that may be crucial on determinant and controversial topics such as the merging of universities (Moed, Moya-Anegón, López-Illescas & Visser, 2011), monitor collaborations and research changes over time (Rafols, Porter & Leydesdorff, 2010) or by extent, any other matter regarding research policy and management at an institutional level (Noyons, 2004).

Taking into account this background, in this paper we propose a novel methodology for representing universities according to their journal publication profile in an attempt to visually synthesize the complex relationships these institutions have with each other. We hypothesize that academic institutions which publish their research output in the same scientific journals should not only have similar research interests but also similar impact, and therefore, should have similar profiles. These last years have seen a great interest on developing measures and thresholds for monitoring and benchmarking universities. The great impact international rankings have had, has not only influenced the Higher Education scenario (Hazelkorn, 2011), but has also risen many questions and critical voices over the methodologies employed when analyzing academic institutions' research output (van Raan, 2005; Torres-Salinas, Moreno-Torres, Delgado-López-Cózar & Herrera, 2011). Universities are subject to numerous influences which differentiate them from other units of analysis such as journals or words. Not only pure research interests drive their relations: geographical and social context among other variables must also be taken into account (Gómez, Bordons, Fernández & Morillo, 2010). In this sense, the application of scientific mapping techniques may be the answer for understanding and reflecting such influences.

This study is structured as follows. In Section 7.2 we present the proposed methodology for mapping academic institutions. Section 7.3 describes the sample of 56 Spanish universities used as a case study and tests this novel methodology, applying it over the total scientific output and also focusing only on two areas (Information and Communication Technologies, as well as Medicine and Pharmacology). Section 7.4 concludes with a discussion over the obtained results.

7.2. Data and methods

The basic idea of the proposed approach is as follows. For each academic institution, we record the scientific journals in which researchers at this institution published their papers during a period of time. No distinction is made between co-authored papers and papers published in a same journal by two different institutions, as we aim at relating universities not just according to their disciplinary focus but also to other external aspects that may influence their similarities such as collaboration or geographical proximity. With the list of scientific journals we construct a journal-by-institution matrix where a given row contains the weights of the corresponding journal across the academic institutions. Here we use the inverse frequency approach (Salton & Buckley, 1988) for generating journal weights, since a journal in which researchers from a large proportion of institutions published their papers should normally be a bad indicator of similarity between two academic institutions. Following a document-document similarity approach (Ahlgren & Colliander, 2009), the behavior of the institution-institution similarity can then be inferred under two types of similarities: first-order and second-order. First-order similarities are obtained by measuring the similarity between columns in a journal-by-institution matrix. However, one may go one step further and obtain them by measuring the similarity between columns in this first-order institution-by-institution similarity matrix. This operation yields a new institution-by-institution matrix, populated with second-order similarities.

In the first-order approach, one focuses on the direct similarity between two academic institutions. The second-order approach determines that, for instance, two universities are similar by detecting that there are other academic institutions such that the two universities are both similar to each of these other institutions. Cluster analysis can then be applied to group the academic institutions in a given set, using second-order institution-institution dissimilarity values. For the cluster analysis here we follow the complete linkage method (Everitt et al., 2001).

7.2.1. Institution-institution similarities

Let $U = \{u_i\}$ be a given set of academic institutions under consideration. Here we suggest that the relationships between research output of institutions in U could be represented based on a comparison of academic journals in which researchers from the institutions in U published their manuscripts.

Let $J = \{j_m\}$ be the set of academic journals in which researchers from the institutions in U published their manuscripts during the study time period. Also, let J_{u_i} be the research output of academic institution u_i.

With the set of academic journals $J = \{j_m\}$ we construct a journal-by-institution matrix $W = \{w_{m,i}\}$ where a given row contains the weights of the corresponding journal across the academic institutions, in particular, $w_{m,i}$ denotes the weight of journal j_i for representing research output of institution u_i.

Following Salton and Buckley (1988), a formal representation of the research output of institution u_i can be obtained by including in J_{u_i} all possible academic journals in J and adding journal weight assignments to provide distinctions among the journals.

Thus if $w_{m,i}$ denotes the weight of journal j_m for representing the research output of institution u_i, and a number of M academic journals are available for research output representation, the journal vector for institution u_i can be written as follows:

$$J_{u_i} = \left(j_1, w_{1,i}; j_2, w_{2,i}; \cdots ; j_M, w_{M,i}\right)$$

(1)

In the following, the basic assumption is that $w_{m,i}$ is equal to 0 when journal j_m is not assigned to institution u_i, since researchers of u_i have not published in j_m. In order to provide a greater degree of discrimination among journals assigned for research output representation, we also assume that journal weights in decreasing journal importance order could be assigned. Hence, the journal weights $w_{m,i}$ could be allowed to vary continuously between 0 and a maximum allowed value, with the higher weight assignments (near the maximum allowed value) being used for the most important journals regarding research output identification, whereas lower weights near 0 would characterize the less important journals for identification.

Given the journal vector representations in Equation (1), an institution-institution similarity value (that is, an indicator of similarity between two academic institutions u_i and u_j in U) may be obtained by comparing the corresponding journal vectors using the vector product formula. But, the individual journal weights should depend to some extent on the weights of other journals in the same vector. To this aim, it is useful to use normalized journal weight assignments. Using a length normalized journal-weighting system, the institution-institution similarity value reduces to the cosine measure (Baeza-Yates & Ribeiro- Neto, 1999) which gives the cosine of the angle between the two vectors which represent the academic institutions u_i and u_j :

$$B(u_i, u_j) = \frac{\sum_m w_{m,i} \times w_{m,j}}{\sqrt{\sum_m (w_{m,i})^2} \sqrt{\sum_m (w_{m,j})^2}}$$

(2)

where $w_{m,i}$ $(w_{m,j})$ is the weight of journal j_m for research output of institution u_i (u_j); and sums are over all journals in the set $J = \{j_m\}$.

Of course, this is a first-order approach for measuring institution-institution similarities, but the behavior of the institution-institution similarity can be inferred under two types of similarities, first-order and second-order. First-order similarities were obtained in Equation (2) by measuring the similarity between columns in a journal-by-institution matrix $\{w_{mi}\}$, where w_{mi} denotes the weight of journal j_m for institution u_i; an operation that yields an institution-by-institution similarity matrix. However, one may go one step further and obtain the similarities by measuring the similarity between columns in this first-order institution-by-institution similarity matrix. This operation yields a new institution-by-institution similarity matrix, populated with second-order similarities. Ahlgren and Colliander (2012) observed good performance of the second-order strategy for measuring similarities in a scientometric context.

From Equation (2), a second order similarity matrix can be defined as follows (Ahlgren & Colliander, 2009):

$$S(u_i, u_j) = \frac{\sum_k B(u_k, u_i) \times B(u_k, u_j)}{\sqrt{\sum_k (B(u_k, u_i))^2} \sqrt{\sum_k \left(B(u_k, u_j)\right)^2}}$$

(3)

where sums are over all academic institutions in the set *U*.

In designing an automatic institution clustering system, two main questions must be answered. First, what appropriate research output units are to be included in the institution representations? Second, is the determination of the journal weights capable of distinguishing the important journals from those less crucial for research output identification?

Concerning the first question, that is, the choice of research output units, various possibilities may be considered. In this paper, academic journals alone were used for research output representation, given the availability of large amounts of data on scientific publications. However, sets of journals cannot provide complete identifications of research-output. But the judicious use of academic journals for institution representation is preferable when incorporating more complex entities, since the following problems would appear when producing complex identifiers (Salton & Buckley, 1988): (i) Few new identifiers are likely to become available when stringent conditions are used for the construction of complex identifiers; and (ii) many marginal institution identifiers that do not prove useful are obtained when the construction criteria for the complex entities are relaxed. Since the construction and identification of complex institution representations can be inordinately difficult, publication in academic journals was used for research output identification. In order to do so, distinctions must be introduced between individual journals, based on their value as institution descriptors. This leads to the use of journal weights attached to the institution identifiers.

In the next section we consider the generation of effective journal weighting factors.

7.2.2. Journal weighting system

A journal-weighting system should increase the effectiveness of institution descriptors. In particular, journals in which researchers from an individual institution frequently published their works appear to be useful as institution identifiers. This suggests that a journal frequency factor can be used as part of the journal-weighting system measuring the frequency of publication in academic journals for a particular institution: $freq_{mi}$ which denotes the number of papers published in journal j_m by researchers at the university u_i during the study time period.

But journal frequency factors alone cannot ensure acceptable institution representation. Specifically, if highly frequent journals are not concentrated in a few particular institutions, but they are prevalent in the whole set U, all academic institutions tend to be represented by these same high frequency-journals and it affects the representation precision. Hence a new set-dependent factor must be introduced that favors journals concentrated in a few institutions of the given set U. The well-known inverse frequency factor (Salton & Buckley, 1988) can be used to perform this function as follows.

Since a journal in which researchers from a large proportion of institutions published their papers should normally be a bad indicator of similarity between two academic institutions, it is reasonable to weight a journal j_m in accordance with how frequently researchers from different institutions in *U* published their papers in this journal, for example, by using

$$\log\left(\frac{N}{n_m}\right)$$

(4)

with *N* being the number of academic institutions in the set $U = \{u_i\}$; and n_m being the number of institutions at which researchers published their work in academic journal j_m.

To sum up, the best journals for research-output description are those able to distinguish certain individual institutions from the rest in the given set *U*. This implies that the best journals j_m for representing research output of institution u_i should have high journal frequencies, $freq_{mi}$, but low overall frequencies across institutions in *U*. Following the approach given by Salton and Buckley (1988) and Ahlgren and Colliander (2009), a reasonable measure of journal importance may then be obtained by using the product of the journal frequency and the inverse frequency factor. Let j_m be the m-th considered academic journal in *J*. We now define the weight of journal j_m for representing research output of institution u_i as:

$$w_{m,i} = freq_{mi} \times \log\left(\frac{N}{n_m}\right)$$

(5)

where $freq_{mi}$ is the number of papers published in journal j_m by researchers at the university u_i during the time period under consideration; and the inverse frequency factor $\log\left(\frac{N}{n_m}\right)$ varies inversely with the number of institutions at which researchers published their work in the same journal j_m.

7.2.3. Assigning a set of academic institutions into groups

Cluster analysis can then be applied in order to group the academic institutions in *U*. To this aim, similarity values obtained by Equation (3) are firstly converted to corresponding dissimilarity values by subtracting a given similarity value from 1. For the cluster analysis, we follow the complete linkage method

(Everitt et al., 2001). In cluster analysis, complete linkage or furthest neighbor is a method for calculating distances between clusters in agglomerative hierarchical clustering. In complete linkage, the distance between two clusters is computed as the maximum distance between a pair of objects, one in one cluster, and one in the other, (Everitt et al., 2001). Thus, the distance between two clusters of academic institutions, *C1* and *C2*, is defined as the maximum dissimilarity between two institutions *u* and *v*, where $u \in C1$ and $v \in C2$:

$$D(C1, C2) = \max_{u \in C1; v \in C2} \left(d(u, v) \right)$$

For example, complete linkage clustering, based on the generated dissimilarity matrices, can be performed following MathWorks (2012).

In agglomerative hierarchical clustering, the clusters are initially the single-member clusters. At each stage the academic institutions or groups of institutions that are closest according to the linkage criterion are joined to form a new, larger cluster. At the last stage, a single group consisting of all academic institutions is formed. This avoids the problem of determining the number of clusters which is often ambiguous, with interpretations depending on the shape and scale of the distribution of points in a data set and the desired clustering resolution of the user. The components at each iterative step are always a subset of other structures. Hence, the subsets can be represented using a tree diagram, or dendrogram. Horizontal slices of the tree at a given level indicate the clusters that exist above and below a value of the weight. Maps of academic institutions are node-edge diagrams, locating each institution in a two or three-dimensional space and with the explicit linking of pairs of institutions by virtue of the relationships between them, i.e., institution-institution similarities. In addition, dendrograms can be provided to illustrate the clustering of institutions or groups of institutions following agglomerative hierarchical clustering, (MathWorks, 2012). Table 7.1 summarizes the methodological approach for construction of maps of academic institutions and the corresponding dendrograms.

Table 7.1. Sum of the proposed methodology for mapping universities according to their journal publication profile

Algorithm 1 Methodological procedure
1. Obtain list of journals on which each institution has published for the study time period
2. Apply weights to journals for each institution according to Equation (5).
3. Construct a journal-by-institution matrix.
4. Extract values from an institution-institution matrix derived from Equation (1).
5. Apply a second-order approach to emphasize similarities among institutions.
6. Perform a complete linkage clustering method in order to set the institutions groups according to their journal publication profile.
7. Construct a dendrogram with all university groups
8. Map the universities network according to their similarity

7.2.4. Data source and processing

Considering that the aim was to visualize the relationships between universities based on their scientific production, the Thomson-Reuters Web of Science database was selected as data source. This decision is based on the great regard this database has for research policy makers, as it is considered to store the most relevant scientific literature in the world. Then, a set of academic institutions selected according to their research output and a study time period were chosen. We manually performed a search query for each university in order to download their research output data. For this, we used the 'Address' filter taking into account all possible names for each institution. Then, we downloaded all records assigned to each institution. We only considered as scientific publications those belonging to journals indexed in one of the Thomson-Reuters Journal Citation Reports (hereafter JCR). These lists of journals are divided per subject categories and contain several bibliometric indicators. One of them is the Impact Factor, which is used as a ranking indicator for ordering journals according to their impact in scientific literature. The editions of the JCR for the study time period were downloaded in

September 2011. Also, we calculated the percentage of papers indexed in fist quartile journals (hereafter Q1 journals). Despite not being necessary for reproducing the suggested methodology, we considered that introducing a color range depending on the percentage of publications in Q1 journals would enrich the maps and ease our discussion over the results when demonstrating how it does not only group universities according to their disciplinary focus but also to their capability on publishing in top journals. This should not be interpreted as assuming that certain universities publish papers of higher impact than others (García, et al, 2012a) but as a competitive advantage of its researchers in terms of visibility.

7.3. Case study: Map of Spanish universities based on institution-institution similarities

7.3.1. Global map of Spanish universities

As a means of validating and applying the proposed methodology for mapping universities (see Table 7.1), we selected a set of Spanish universities with at least 50 citable documents (articles, reviews, notes and letters) published in JCR Journals, resulting in 56 universities (see Table 7.2), and downloaded their production for the 2008-2010 time period. The timeframe chosen aims at portraying as accurately as possible the current Spanish higher education landscape regarding its research performance. For each university we retrieved all scientific journals in which researchers from each institution published their papers during the study time period. We then used the cosine measure to compute a first-order and second-order similarity between universities. The map of Spanish universities will be a node-edge diagram, locating each university in a two-dimensional space and with the explicit linking of pairs of universities by virtue of the relationships between them, i.e., university-university similarity values. For this, the software program Pajek (Networks/Pajek, 2011) was used and universities' positioning was determined in accord to the Kamada-Kawai algorithm (Kamada-Kawai, 1998), which is commonly used in this kind of representations. Next, we used the complete linkage method for clustering the 56 Spanish universities using second-order dissimilarities.

Table 7.2. Set of Spanish universities used as sample for mapping institutions according its scientific research output during de 2008-2010 time period

University	NDOC	%Q1	University	NDOC	%Q1	University	NDOC	%Q1
Barcelona	11168	56%	Alicante	2349	50%	Lleida	1124	51%
Autónoma de Barcelona	8428	56%	Córdoba	2334	57%	Almeria	1085	46%
Complutense Madrid	7629	51%	Rovira I Virgili	2302	55%	Publica De Navarra	1016	44%
Valencia	6764	54%	Valladolid	2187	43%	Palmas (Las)	1016	43%
Autónoma de Madrid	6386	56%	Laguna, La	2176	52%	Uned	929	41%
Granada	5380	49%	Malaga	2076	48%	Leon	917	48%
Politécnica de Cataluña	4992	49%	Pompeu Fabra	1972	59%	Politécnica Cartagena	908	46%
Pais Vasco	4827	52%	Cantabria	1826	51%	Huelva	748	52%
Zaragoza	4487	53%	Extremadura	1816	49%	Pablo Olavide	656	51%
Sevilla	4484	50%	Alcala de Henares	1809	46%	Burgos	478	52%
Politecnica de Valencia	4445	49%	Carlos Iii	1805	43%	Rioja (La)	446	50%
Santiago de Compostela	4400	50%	Islas Baleares	1565	56%	Ramon Llul	366	38%
Politécnica de Madrid	4065	43%	Girona	1520	53%	Europea de Madrid	190	45%
Oviedo	3232	49%	Miguel Hernandez	1519	48%	Cardenal Herrera-Ceu	189	34%
Vigo	2983	49%	Rey Juan Carlos	1512	49%	San Pablo Ceu	171	49%
Castilla La Mancha	2829	50%	Coruña, A	1439	41%	Pontificia Comillas	144	45%
Murcia	2663	45%	Jaen	1355	43%	Mondragon	80	39%
Salamanca	2510	48%	Cadiz	1261	48%	Deusto	55	22%
Navarra	2469	47%	Jaume I	1225	54%			

Indicators: NDOCS: Number of citable documents (article, review, note or letters) indexed in JCR Journals (Thomson-Reuters) **%Q1**: Number of citable documents (article, review, note or letters) indexed in Q1 JCR Journals (Thomson Reuters)

Here we have used the cosine measure to compute the first-order and second-order similarity between universities as given above (see Equations (2) and (3)). The second-order similarity matrix S contains many cells with very low similarities. From a computational point of view, it is problematic to keep all such similarities in the matrix. Moreover, to take them into account in the computations might have a negative impact on the visualization quality. We handled this problem by establishing minimum similarity values (e.g., 0.6 in Fig. 7.1).

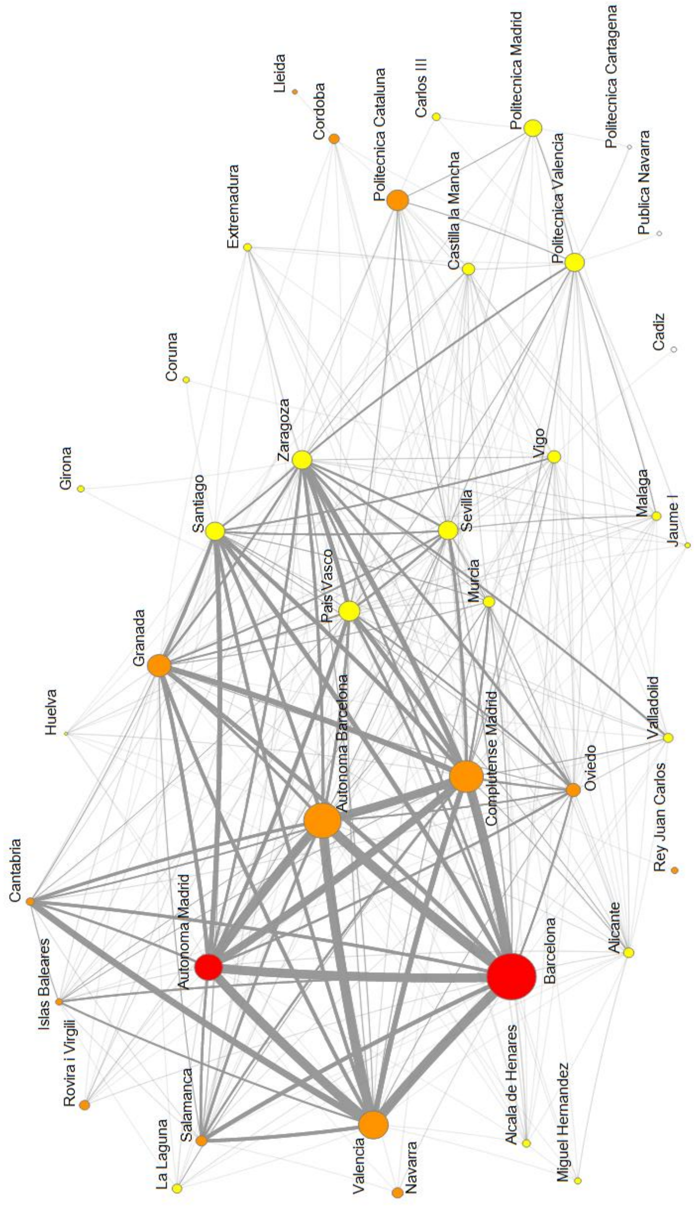
Figure 7.1 shows the resulting map for Spanish universities. Four distinct groups of universities can be inferred according to similarities in their research profile. On the first hand we have a group formed by the five universities which could be considered as the most important ones (Barcelona, Autónoma de Madrid, Autónoma de Barcelona, Valencia and Complutense Madrid) as these occupy the highest positions (for Spanish universities) in well-known international rankings such as the Shanghai Ranking (Shanghai Jiao Tong University, 2011) or the Performance Ranking of Scientific Papers for World Universities (Higher Education Evaluation & Accreditation Council of Taiwan, 2011). These universities are the ones with the highest production and more links with the rest of universities which seem to surround them. The high number of links may suggest that they are not just highly productive universities, but also generalist universities covering different disciplines. It is also noticeable that, except Valencia, all universities belong either to Madrid or Barcelona, the two main cities in Spain. They are similar universities not only in their disciplinary orientation, but also in their size and scientific impact according to its percentage of documents in Q1 journals. The second group (Granada, Santiago, Zaragoza, País Vasco, Sevilla) would be formed by a set of universities also generalist and surrounded by a dense network but of a smaller size. Funnily enough these universities usually occupy positions between 400-500 in the Shanghai Ranking; dropping out some years and appearing others, which also reinforces their similarity. However, some distinctions can be made when relating their Q1 production and their positions in the Shanghai Ranking; while Granada appears in all editions of the ranking, the others drop in some editions, maybe related to the proportion of Q1 production each university has. In this sense, it seems that this university is somewhere between these two groups.

A third group can be distinguished by less productive universities (hence, smaller universities) which have strong links only with those universities belonging to the first group, showing similarities in certain fields of endeavor.

These universities are characterized by their size. They seem to reflect the model of bigger universities and therefore their similarities with these universities. Universities belonging to this group would be Cantabria, Islas Baleares or Oviedo for instance. The fourth group is integrated by small universities with weak links to universities belonging to the first or second group. These weak links are due to a high specialization on certain fields also common to the other universities (Torres-Salinas, Delgado-López-Cózar, Moreno-Torres & Herrera; 2011). An example of this would be Navarra (Medicine and Pharmacy), Rovira i Virgili (Chemistry), or Murcia (Biological Sciences). The last group is mainly formed by the universities named as Polytechnics or Technological (Politécnica de Madrid, Politécnica de Valencia, Politécnica de Cataluña, etc.). Though these universities are linked with the rest of universities, they are also linked between them. The reason for showing such weak links is due to their high specialization on certain scientific fields belonging to the Engineering and Applied Sciences. In fact, surrounding them we also find other universities that show a tendency towards this "technological" profile, such as Zaragoza (which shares a strong link to Politécnica de Valencia), Carlos III, Pública de Navarra or Castilla La Mancha.

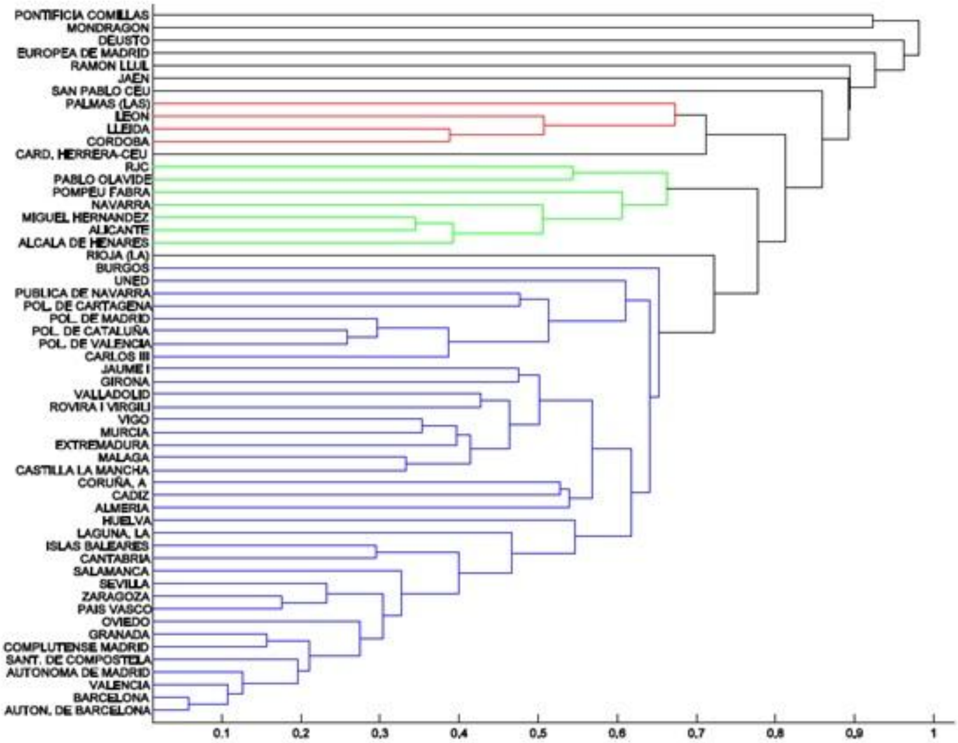
The high minimum values established in Figure 7.1, seem to eliminate most reflections of the geographical or regional relations among universities, emphasizing purely research similarities. But we can still trace this kind of relationship between three universities: Santiago de Compostela, Vigo and Coruña. In this case, the interpretation seems to be quite reasonable. The two latter universities were formed in 1990 and 1989 respectively both from campuses belonging to the former university, which is a historical university funded in the fifteenth century.

Figure 7.1. Map of main Spanish universities according to their journal publication profile.



Map Characteristic: Lines > minimum similarity value 0.60; maximum similarity value 0.98. **Isolated university nodes** have been removed. From 0.75 line-width is emphasized. **Colors:** >50% production belongs to Q1 journals; 40-50% production belongs to Q1 journals; 30-40% production belongs to Q1 journals; <30% production belongs to Q1 journals.

Figure 7.2. Dendrogram of Spanish universities according to their journal publication profile



In this map we find that one important university is missing, the University of Pompeu Fabra. This Catalan university has experienced a meteoric growth during the last years. A relatively new university (it was founded in 1990), during the last two years it has appeared in the most renown international rankings: between the 300 and 500 top class universities according to the Shanghai Ranking since 2009 or between the 150 and 200 top universities in the last two years according to The Times World Universities Ranking, for instance. Its absence in Figure 7.1 suggests that its publication patterns differ from the rest of the Spanish universities, suggesting that probably its journal publication profile may be oriented in such a way that can explain such an outburst. As we indicated before, by using common journals as a means for mapping universities, we not only group them according to their research profile, but also to their research impact (understood as the impact factor of journals in which their output is published). This university serves as a good example of this second characteristic as 59% of its production is published in Q1 journals (see Table 7.2), that is the highest proportion for the sample used. This way we can see how its absence may not have to do so much by its disciplinary profile but with the journals in which it publishes. Figure 7.2 shows a dendrogram of Spanish universities or groups of universities following agglomerative hierarchical clustering. From this figure, it follows the rapid grouping of Barcelona, Autónoma de Barcelona, Valencia, and Autónoma de Madrid, which belong to the core of the map of Spanish universities according to their journal publication profile as given in Fig. 7.1. We have also that Granada and Complutense de Madrid form a very strong grouping. Another relatively natural grouping is formed by Politécnica de Valencia, Politécnica de Cataluña and Politécnica de Madrid, all of them which are universities with a tendency towards the technological profile. From Fig. 7.2, we have that Sevilla, Zaragoza, and País Vasco belong to another group of universities according to their journal publication profile.

7.3.2. Specific maps of Spanish universities for the fields of Information and Communication Technologies, as well as Medicine and Pharmacology

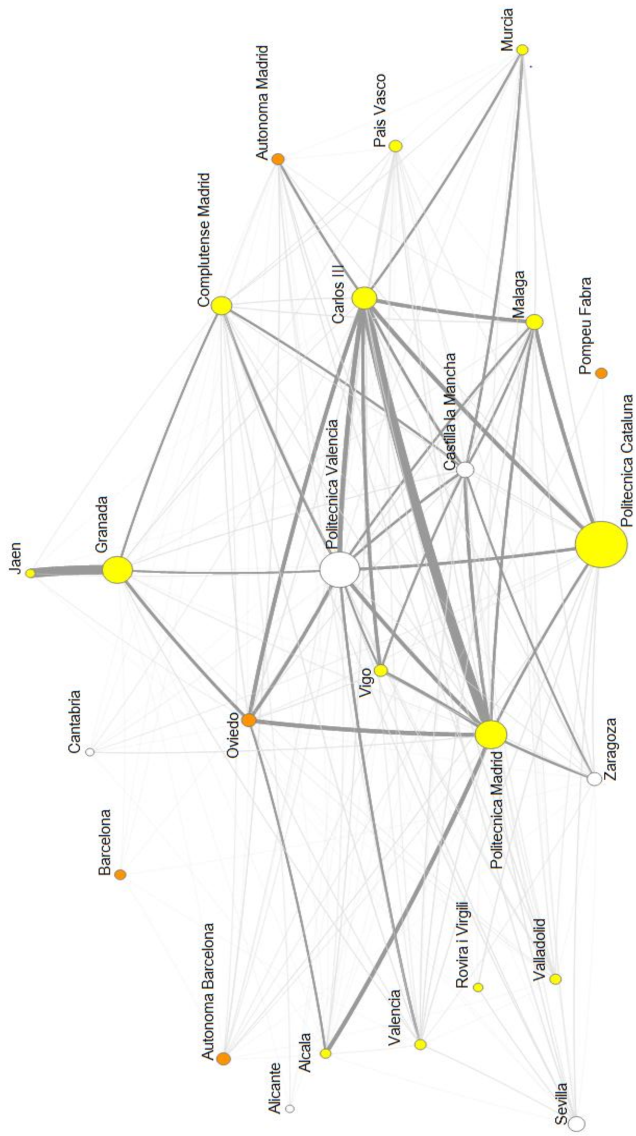
After testing our methodology for the total production of universities, we go a step further and test it for different scientific fields in the belief that in order to have a clear and more precise picture of universities' similarities, it is necessary to deepen on specific fields so that we can understand better their relations. For this, we focus in two different areas: Information and Communication Technologies (hereafter ICT) and Medicine and Pharmacology (hereafter MED). We construct these fields by aggregating thematically the Thomson Reuters subject categories, following the same criteria we did in a previous study [1] (Torres-Salinas, Moreno-Torres, Robinson-García, Delgado-López-Cózar & Herrera; 2011). We use the same set of 56 Spanish universities (Table 2) and the same study time period (2008-2010).

In Figure 7.3 we map Spanish Universities according to their journal publication profile in ICT. In this case, disciplines are crucial on shaping universities similarities. We find that Politécnica de Valencia shows a much more diversified profile in this scientific field, occupying a central place in the representation. That is, it is similar to a greater amount of universities, signifying its lesser specialization on certain disciplines. Oviedo, Politécnica de

[1] For a better understanding on how these broad scientific fields were formed the reader is referred to http://www.ugr.es/~elrobin/rankingsISI_2011.pdf where we show the correspondence followed between the ISI subject categories and 12 scientific fields including the two used in this study.

Madrid and Carlos III show greater similarities among them and also, each of them is the core for grouping other universities.

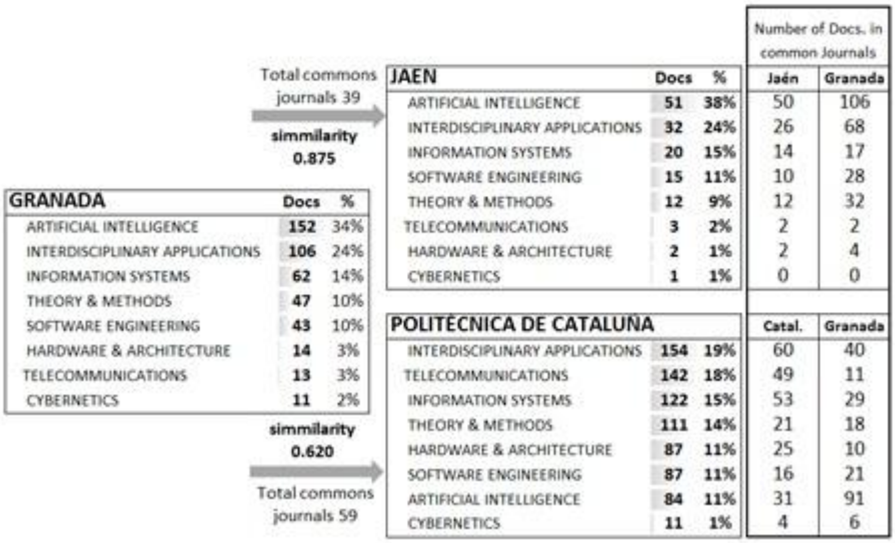
Figure 7.3. Map of Spanish universities according to their journal publication profile in ICT



Map Characteristic: Lines > minimum similarity value 0.60; maximum similarity value 0.875. **Isolated university nodes** have been removed. From 0.75 line-width is emphasized. **Colors:** ■>50% production belongs to Q1 journals; ■ 40-50% production belongs to Q1 journals; □ 30-40% production belongs to Q1 journals; □<30% production belongs to Q1 journals.

But the most interesting patterns are those followed by Granada and Politécnica de Cataluña. According to their research impact and output, these two universities are the top ones on this scientific field (Torres-Salinas, Delgado-López-Cózar, Moreno-Torres, Herrera; 2011) but they are not the core of the representation as one would have thought. Instead, they seem to follow different patterns than the rest of the universities, suggesting a highly specialized profile in both cases. While Politécnica de Cataluña shows stronger similarities with other universities such as Málaga, Carlos III, Politécnica de Madrid and Politécnica de Valencia; Granada shows a high similarity with Jaén and weaker ones with the rest. The reason for this dissimilarity could lay on a high specialization on different research lines than those followed by the rest of the universities. Also there are geographical and social factors that influence the strong similarity with Jaén among those related with research. As it occurred with Santiago de Compostela, Vigo and Coruña before, Jaén is a relatively new university (it was founded in 1993) which used to be a campus belonging to the University of Granada. This social context may explain their similarity, as there are probably still strong collaboration links between researchers in ICT belonging to both universities.

Figure 7.4. Detail of disciplinary differences in ICT between Granada, Jaén and Politécnica de Cataluña according to the Thomson Reuters subject categories



GRANADA	Docs	%
ARTIFICIAL INTELLIGENCE	152	34%
INTERDISCIPLINARY APPLICATIONS	106	24%
INFORMATION SYSTEMS	62	14%
THEORY & METHODS	47	10%
SOFTWARE ENGINEERING	43	10%
HARDWARE & ARCHITECTURE	14	3%
TELECOMMUNICATIONS	13	3%
CYBERNETICS	11	2%

Total commons journals 39 — simmilarity 0.875

JAEN	Docs	%	Number of Docs. in common Journals: Jaén	Granada
ARTIFICIAL INTELLIGENCE	51	38%	50	106
INTERDISCIPLINARY APPLICATIONS	32	24%	26	68
INFORMATION SYSTEMS	20	15%	14	17
SOFTWARE ENGINEERING	15	11%	10	28
THEORY & METHODS	12	9%	12	32
TELECOMMUNICATIONS	3	2%	2	2
HARDWARE & ARCHITECTURE	2	1%	2	4
CYBERNETICS	1	1%	0	0

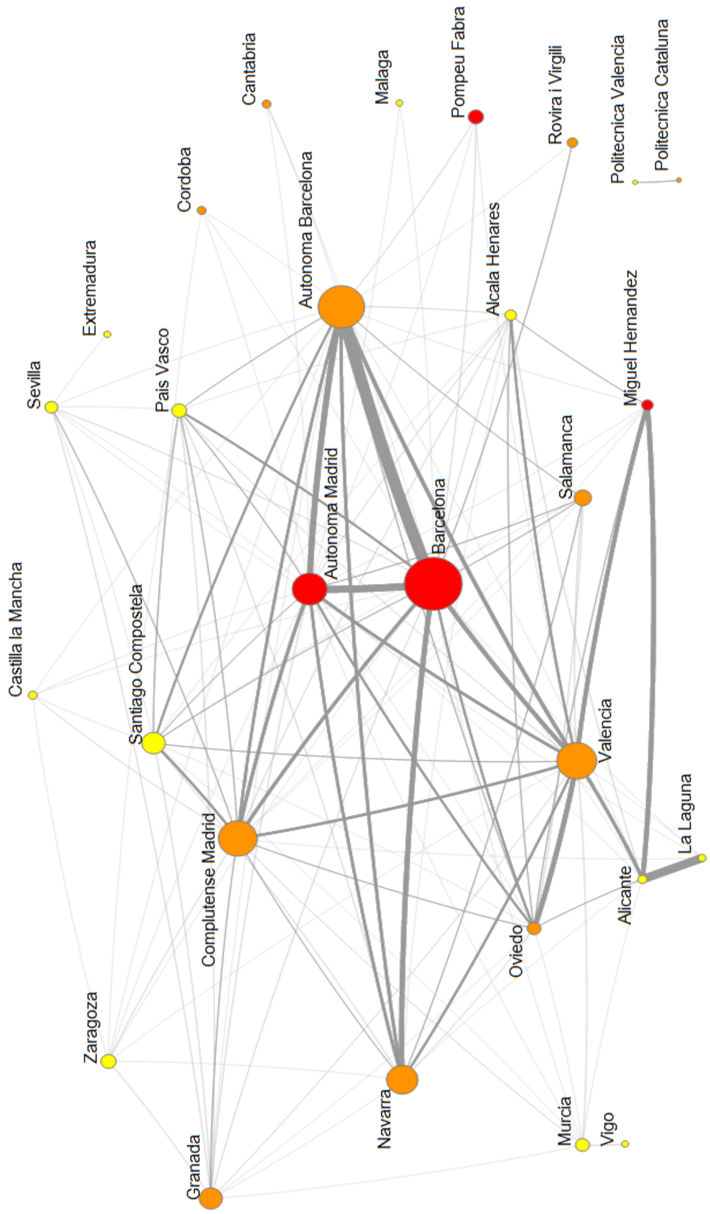
simmilarity 0.620 — Total commons journals 59

POLITECNICA DE CATALUÑA	Docs	%	Catal.	Granada
INTERDISCIPLINARY APPLICATIONS	154	19%	60	40
TELECOMMUNICATIONS	142	18%	49	11
INFORMATION SYSTEMS	122	15%	53	29
THEORY & METHODS	111	14%	21	18
HARDWARE & ARCHITECTURE	87	11%	25	10
SOFTWARE ENGINEERING	87	11%	16	21
ARTIFICIAL INTELLIGENCE	84	11%	31	91
CYBERNETICS	11	1%	4	6

This hypothesis is reinforced by Figure 7.4 in which we see the distribution of research output according to the Thomson Reuters Subject Categories for three universities: Granada, Jaén and Politécnica de Cataluña. Deepening in categories allows us to observe the similarities between the two former and dissimilarities with the latter. This way we see how high levels of similarity correspond with similar publication profiles; Jaén and Granada's research distribution per categories is very similar and much focused in two main categories (*Artificial intelligence* and *Interdisciplinary applications*) which contain more than half of their total production for both universities. On the other hand, Politécnica de Cataluña shows a more diversified profile never reaching 20% of its production in just one category. It is also interesting to see how the proposed methodology is not influenced by size. Despite having Granada more journals in common with Politécnica de Cataluña, the proportion of publications in the same journals with Jaén is higher, which explains their similar profile.

When focusing in MED a different picture emerges (Figure 7.5), signifying how necessary becomes a disciplinary approach to universities when establishing research profiles. In this case we find four distinct groups of universities. The main one is composed by Barcelona, Autónoma de Barcelona and Autónoma de Madrid, which have strong similarities among them. They are characterized by their large production and by publishing in Q1 journals (only Autónoma de Barcelona has less than half of its output published in Q1 journals). They are also the most generalist universities in this field of endeavor as they represent the core of the map. Then, we find a second group of universities with high outputs which surround this core (Complutense de Madrid, Navarra, Valencia). In the case of Navarra and comparing with Figure 7.1, it is plausible to suggest that it is a highly specialized University in MED with a very similar profile to Autónoma de Madrid, Barcelona, Autónoma de Barcelona and Valencia. The third group is formed by universities with weak links with universities belonging to the other two groups, for instance, Alcalá de Henares, Granada or País Vasco.

Figure 7.5. Map of Spanish universities according to their journal publication profile in MED



Map Characteristic: Lines > minimum similarity value 0.60; maximum similarity value 0.93. **Isolated university nodes** have been removed. From 0.75 line-width is emphasized. **Colors:** ■ >50% production belongs to Q1 journals; ■ 40-50% production belongs to Q1 journals; ■ 30-40% production belongs to Q1 journals; □ <30% production belongs to Q1 journals.

It is worth mentioning a fourth group formed by just two universities and completely separated from the rest. This is the one formed by Politécnica de Valencia and Politécnica de Cataluña. As it can be drawn through all this section, Polytechnics are very similar in their research profile. In this case, this similarity between them on the one hand, and dissimilarity from the rest of the universities on the other, is due to a research interest focused on the *Engineering, Biomedical* Thomson Reuters JCR subject category which would explain why there is no connection with the other universities. In fact, their production in this category represent 30% of their total output in MED, that is, 61 documents published by Politécnica de Cataluña and 66 documents published by Politécnica de Valencia.

In Figure 7.6 we emphasize as we did with ICT (Figure 7.4), the capability of the proposed methodology for grouping similar universities and separating dissimilar universities according to their journal publication profile in MED. In this case, we compare the distribution of research output according to the Thomson Reuters Subject Categories of Autónoma de Barcelona with Barcelona and Alcalá de Henares. That is, with its most similar university and a lesser similar one. In the first case, we observe a similarity of 0.930, which stresses how alike the profile of these two universities is in this scientific field. In fact, the eight categories in which they produce more documents are the exact same for both institutions. On the other hand, when comparing Autónoma de Barcelona with Alcalá de Henares we see that, despite publishing an important proportion of their total output in the same four categories, - mainly those related with Neurosciences, - they also present a special focus on different specialties that make them quite different (in the case of Alcalá de Henares for instance, *Ophthalmology*, *Oncology* or *Surgery*). Thereby we can witness once more how the methodology employed groups universities according to their research and publication similarities.

Figure 7.6. Detail of disciplinary differences in MED between Autónoma de Barcelona, Barcelona and Alcalá de Henares according to the Thomson Reuters subject categories

AUTONOMA BARCELONA	Docs	%
NEUROSCIENCES	258	8%
PHARMACOLOGY & PHARMACY	221	7%
CLINICAL NEUROLOGY	211	7%
PSYCHIATRY	192	6%
INFECTIOUS DISEASES	169	5%
INMUNOLOGY	168	5%
SURGERY	115	4%
ENDOCRINOLOGY & METABOLISM	114	4%

Total commons journals 613 — simmilarity 0.930

BARCELONA	Docs	%	Number of Docs. in common Journals: Barc.	Aut Barc.
PSYCHIATRY	440	9%	242	125
NEUROSCIENCES	431	9%	334	241
CLINICAL NEUROLOGY	320	7%	261	193
PHARMACOLOGY & PHARMACY	316	6%	240	198
INMUNOLOGY	246	5%	222	156
INFECTIOUS DISEASES	240	5%	227	164
SURGERY	194	4%	143	100
ENDOCRINOLOGY & METABOLISM	189	4%	146	103

simmilarity 0.690 — Total commons journals 179

ALCALA HENARES	Docs	%	Alcalá	Aut Barc
PHARMACOLOGY & PHARMACY	53	9%	31	99
NEUROSCIENCES	52	9%	39	83
OPHTHALMOLOGY	48	8%	35	29
PSYCHIATRY	41	7%	28	48
CLINICAL NEUROLOGY	40	7%	28	70
ONCOLOGY	32	5%	18	34
SURGERY	29	5%	22	30
DERMATOLOGY	26	4%	22	24

7.4. Discussion and concluding remarks

The present study aims at proposing a novel methodology for mapping academic institutions according to their research profile. Based on the presumption that similar universities should publish in the same scientific journals, we present an algorithm for measuring similarities between universities and their journal publication profile and we represent them in a dendrogram and a network map. In order to test this methodology we set a sample of 56 Spanish universities and a three-year study time period (2008-2010). Then, we apply this methodology in three different scenarios: a representation of universities according to their total output, a representation according to their output in ICT, and a representation according to their output in MED.

This way we first analyze its potential for grouping institutions in a competitive context deeply influenced by table leagues and rankings in which it has repeatedly been noted that only similar institutions can be compared in order to proceed properly when ranking (van der Wende & Westerheijden, 2009). This can be seen in Figure 7.1 where we observe how the proportion of publications in Q1 journals for universities is similar for each of the previously discussed groups. Although some attempts have been done when classifying universities according to their research performance (Shin, 2009), this approach focuses on mapping universities according to their journal publication profiles, in the belief that this perspective ends with limitations derived from a rigid classification system subjected to a fixed set of criteria. Also, it allows grouping universities taking into account their disciplinary similarities (Lopez-Illescas, Moya-Anegón & Moed, 2011) and their research impact or quality (considering as such publications in Q1 journals). This way we address not only to vertical diversity between universities, which is the one rankings emphasize, but also horizontal diversity.

In this vein go the other two tests presented. When analyzing the methodology in two different scientific fields, we intend to demonstrate how our approach can, not just group similar universities, but also detect similarities between institutions that are centered in the same disciplines and specialties. Also, we have noted that, having a previous knowledge over a determined higher education system over which the procedure is performed, we can also discover geographical, social and/or historical relationships between academic institutions, as we have previously seen in the case of Santiago de Compostela, Vigo and Coruña in Figure 7.1 or Granada and Jaén in Figure 7.3.

To validate the results illustrated in Figure 7.1, a different method with similar results needs to be presented. We used García et al (2012b) where a summary measure of multidimensional prestige of influential fields was introduced to assess the comparative performance of Spanish Universities during the period 2006-2010.

To this aim, a field of study at a given university is considered as having dimension specific prestige when its score based on a given ranking model (e.g., %Q1) exceeds a threshold value. Then, it can be defined which fields at a given university are considered to be prestigious in a multidimensional setting. Thus, a field of study at this university has multidimensional prestige only if it is an influential field with respect to a number of dimensions. Finally, after having identified the multidimensional influential fields at a particular university, their prestige scores are aggregated to a summary measure of multidimensional prestige. The summary measure is not only sensitive to the number of dimensions but also takes into account changes in the ranking scores of influential fields of study at the university.

García et al (2012b) shows the ranking of research output of Spanish universities during the period 2006-2010 (see Table 7.5). To this aim it was computed the multidimensional prestige of influential fields of study at each institution using a multivariate indicator space. Six variables were used in this analysis: (1) Raw number of citable papers (articles, reviews, notes or letters) published in scientific journals (NDOC); (2) Number of citations received by all citable papers (NCIT); (3) H-Index (H); (4) Ratio of papers published in journals in the top JCR quartile (%Q1); (5) Average number of citations received by all citable papers (ACIT); and (6) Ratio of papers that belong to the top 10% most cited (TOPCIT). The data are available at http://www.rankinguniversidades.es. Fifty-six main universities in Spain are considered in this experiment.

From the results showen in García et al (2012b), the top 8 Spanish universities during the period 2006-2010 were: (1) Barcelona; (2) Autónoma de Barcelona; (3) Autónoma de Madrid; (4) Valencia; (5) Complutense de Madrid; (6) Granada; (7) Santiago de Compostela; and (8) Zaragoza. Also it follows that País Vasco and Sevilla are very similar according to their multidimensional prestige of influential fields. This also happens to two other technological universities: Politécnica de Valencia and Politécnica de Cataluña; which are similar according to their multidimensional prestige (see Table 5 in that paper).

The interesting point is that all these results are congruent with those from the present study (as given by Figure 7.1 and Figure 7.2) where we analyze the main Spanish universities according to their journal publication profile.

This type of representation offers a new model for visualizing universities' relationships that can show more clearly than other types of mapping (such as collaboration or web-links maps) the multidimensional similarities and dissimilarities between academic institutions. Likewise, this tool serves as a perfect complement for interpreting universities' performance in rankings as a means for understanding them not as isolated entities, but as interrelated elements of a national higher education system. At a research policy level, this mapping technique may be of use when identifying and selecting universities with similar profiles, as it helps us to identify which universities can be compared and which not, not just at a national level, as has been described through all the paper, but also to compare universities at a transnational or international level. Finally, in the national context it may be of special interest for research policy managers when analyzing potential merging of universities or concentration of research. This last idea goes in consonance with recent developments in Spain regarding its research policy and the 'International Excellence Campus' [Campus de Excelencia Internacional] program which aims at encouraging universities' collaboration.

However, some limitations have also been noted. Using the journal publication approach we find too many links between universities, which makes it difficult to visualize universities under certain levels of similarity, blurring similarities between low performance universities. This limits the analysis when mapping a whole national higher education system as some universities have to inevitably, drop out. In this sense, it also understandable that applying this type of methodologies under a certain threshold is not advisable. Also it would be of interest to introduce other document types (monographs for instance) that could permit a better coverage of certain fields such as social sciences and humanities, and develop methodologies that would adjust to these document types.

References

Ahlgren, P., & Colliander, C. (2009). Document-document similarity approaches and science mapping: experimental comparison of five approaches. *Journal of Informetrics, 3*(1), 49-63.

Ahlgren, P., & Colliander, C. (2012). Experimental comparison of first and second-order similarities in a scientometric context. *Scientometrics, 90*(2), 675-685.

Baeza-Yates, R. A., & Ribeiro-Neto, B. A. (1999). *Modern information retrieval* (Chapter 2). Harlow: Addison-Wesley.

Everitt, B., Landau, S., & Leese, M. (2001). *Cluster analysis (4th ed.)*. London: Arnold. Hierarchical Clustering.

García, J. A., Rodríguez-Sánchez, R., Fdez-Valdivia, J. & Martínez-Baena, J. (2012a). On first quartile journals which are not of highest impact. *Scientometrics, 90*(3), 925-943.

García, J. A., Rodríguez-Sánchez, R., Fdez-Valdivia, J., Torres-Salinas, D. & Herrera, F. (2012b). Ranking of research output of universities of universities on the basis of the multidimensional prestige of influential fields: Spanish universities as a case study. *Scientometrics, 93*(3), 1081-1099.

Gmür, M. (2003). Co-citation analysis and the search for invisible colleges: a methodological evaluation. *Scientometrics, 57*(1), 27-57.

Gómez, I., Bordons, M., Teresa-Fernández, M. & Morillo, F. (2009). Structure and research performance of Spanish universities. *Scientometrics, 79*(1), 131-146.

Griffith, B.C., Small, H.G., Stonehill, J.A., & Dey, S. (1974). Structure of scientific literatures. 2. Toward a macrostructure and microstructure for science. *Science Studies, 4*, 339-365.

Hazelkorn, E. (2011). *Rankings and the reshaping of higher education: the battle for world-class excellence*. Basingstoke, New York: Palgrave MacMillan.

Higher Education Evaluation & Accreditation Council of Taiwan (2011). *Performance Ranking of Scientific Papers for World Universities*. http://ranking.heeact.edu.tw/en-us/2011/Page/Methodology Accessed Jan 2012.

Kamada-Kawai, S (1998). A simple method for computing general position in displaying three-dimensional objects. *Computer Vision, Graphics, and Image Processing, 41*(1), 43-56

Klavans, R., & Boyack, K.W., (2009). Toward a consensus map of science. *Journal of the American Society for Information Science and Technology, 60*(3), 455-476.

Leydesdorff, L., & Persson, O. (2010). Mapping the geography of science: distribution patterns and networks of relations among cities and institutions. *Journal of the American Society for Information Science and Technology, 61*(8), 1622-1634.

Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology, 60*(2), 348-369.

López-Illescas, C., Moya-Anegón, F., & Moed, H.F. (2011). A ranking of universities should account for differences in their disciplinary specialization. *Scientometrics, 88*(2), 563-574.

MathWorks, (2012). *Statistics toolbox*. From http://www.mathworks.es/help/toolbox/stats/linkage.html (visited January 27, 2012).

Moed, H.F., Moya-Anegón, F., López-Illescas, C., & Visser, M. (2011). Is concentration of university research associated with better research performance? *Journal of Informetrics, 5*(4), 649-658.

Moya-Anegón, F., Vargas-Quesada, B., Herrero-Solana, V., Chinchilla-Rodríguez, Z., Corera-Alvarez, E., & Muñoz-Fernández, F.J. (2004). A new technique for building maps of large scientific domains based on the co-citation of classes and categories. *Scientometrics, 61*(1), 129-145.

Moya-Anegón, F., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., Corera-Alvarez, E., Muñoz-Fernández, F.J., & Herrero-Solana, V. (2007). Visualizing the marrow of science. *Journal of the American Society for Information Science and Technology, 58*(14), 2167-2179.

Narin, F., Carpenter, M., & Berlt, N.C. (1972). Interrelationships of scientific journals. *Journal of the American Society for Information Science, 23*(5), 323-331.

Networks/Pajek (2011). http://vlado.fmf.uni-lj.si/pub/networks/pajek/ Accessed Jan 2012.

Noyons, C., (2004). Science Maps Within a Science Policy Context. H.F. Moed et al. (eds). *Handbook of Quantitative Science and Technology Research*. Dordrecht: Kluwer Academic Publishers, pp. 237-255.

Noyons, E.C.M., Moed, H.F., & Luwel, M. (1999). Combining mapping and citation analysis for evaluative bibliometric purposes: a bibliometric study. *Journal of the American Society for Information Science and Technology, 50*(2), 115-131.

Ortega, J.L., Aguillo, I., Cothey, V., & Scharnhorst, A. (2008). Maps of the academic web in the European Higher Education Area – an exploration of visual web indicators. *Scientometrics, 74*(2), 295-308.

Rafols, I., Porter, A.L., & Leydesdorff, L. (2010). Science overlay maps: a new tool for research policy and library management. *Journal of the American Society for Information Science and Technology, 61*(9), 1871-1887.

Rorissa, A. & Yuan, X. (2012). Visualizing and mapping the intellectual structure of information retrieval. *Information Processing and Management, 48*(1), 120-135.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management, 24*(5), 513-523.

Shanghai Jiao Tong University, (2011). *Academic ranking of world universities (ARWU)*. http://www. arwu.org/index.js. Accessed Dec 2011.

Shin, J.C. (2009). Classifying higher education institutions in Korea: a performance-based approach. *Higher Education, 57*(2), 247-266.

Small, H. (1999). Visualizing science by citation mapping. *Journal of the American Society for Information Science, 50*(9), 799-813.

Small, H., & Garfield, E. (1985). The geography of science: Disciplinary and national mappings. *Journal of Information Science, 11*(4), 147-159.

Torres-Salinas, D., Delgado-López-Cózar, E., Moreno-Torres, J.G., & Herrera, F. (2011). Rankings ISI de las universidades españolas por campos científicos: descripción y resultados. *El profesional de la información, 20*(1), 111-122.

Torres-Salinas, D. (2009). Evaluación bibliométrica de universidades con Scival de Elsevier. *El Profesional de la Información, 18*(6), 669-674.

Torres-Salinas, D., Moreno-Torres, J.G., Delgado-López-Cózar, E., & Herrera, F. (2011). A methodology for Institution-Field ranking based on a bidimensional analysis: the IFQ2A index. *Scientometrics, 88*(3), 771-786.

Torres-Salinas, D., Moreno-Torres, J.G., Robinson-García, N., Delgado-López-Cózar, E., & Herrera, F. (2011). Rankings ISI de las universidades españolas según campos y disciplinas científicas (2 ed. 2011). *El Profesional de la Información, 20*(6), 701-711.

Van der Wende, M., & Westerheijden, D. (2009). Rankings and classifications: the need for a multidimensional approach. Van Vught (ed), *Mapping the higher education landscape*. Springer, pp. 71-86.

Van Raan, A.F.J. (2005). Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics, 62*(1), 133-143.

8. Análisis de redes de las universidades españolas de acuerdo a su perfil de publicación en revistas por áreas científicas

Paper published in Revista Española de Documentación Científica doi:10.3989/redc.2013.1042 authored by N. Robinson-García, R. Rodríguez-Sánchez, J.A. García, D. Torres-Salinas and J. Fdez-Valdivia

Este trabajo presenta un análisis descriptivo de las universidades españolas de acuerdo a su perfil de publicación en revistas científicas en cinco áreas de conocimiento para el periodo 2007-2011. Dos universidades tienen un perfil de publicación en revistas similar cuando publican en un alto número de revistas comunes. Siguiendo este principio es posible crear mapas de universidades que ofrezcan una visión enriquecedora del sistema universitario español. Para ello, analizamos las áreas de Ciencias Sociales, Ciencias Exactas, Ingeniería y Tecnología, Ciencias de la Vida y Ciencias de la Salud. Además, utilizamos el indicador de centralidad del análisis de redes sociales para identificar aquellas universidades que muestran un rol más destacado en cada área al tener un mayor número de conexiones directas con el resto de universidades. Finalmente, discutimos la aplicación de esta metodología en un contexto de política científica de cara a la búsqueda de colaboraciones científicas potenciales.

8.1. Introducción

La hegemonía de las universidades estadounidenses en las primeras posiciones de los rankings internacionales ha tenido un gran impacto en la agenda científica de los países europeos que han visto en dicho predominio, el fracaso del modelo universitario europeo frente al modelo estadounidense. Un modelo caracterizado por incentivar la democratización de la enseñanza superior y el equilibrio entre la docencia y la investigación (Schimank y Winnes, 2000) frente a un modelo de universidades altamente competitivo (Visser, Calero Medina y Moed, 2007). A pesar de los esfuerzos de la Unión Europea por establecer un marco metodológico que permitiera mantener las bondades de las universidades europeas (van der Wende y Westerheijden, 2009) y al mismo tiempo competir por la excelencia, lo cierto es que el éxito de las universidades americanas ha precipitado a los distintos gobiernos a la búsqueda de nuevas fórmulas que permitan alcanzar altas posiciones en dichos rankings para ganar prestigio así como los recursos económicos y humanos que éste trae consigo (Hazelkorn, 2009). En este sentido, el conocido Ranking de Shanghai (Liu y Chen, 2005) cuya primera edición data de 2003, ha logrado posicionarse como referente mundial de gobiernos y universidades, que lo emplean para conocer su situación y establecer políticas encaminadas a mejorar su posición (Docampo y otros, 2012; Hazelkorn, 2011). No obstante, las graves limitaciones de estos productos (Van Raan, 2005; Billaut, Boyssou y Vincke, 2010), así como una interpretación errónea y a veces simplista de los resultados pueden tener efectos nocivos en el sistema universitario (Docampo y otros, 2012; Macilwain, 2012).

De manera paralela, se han ido desarrollando una serie de metodologías para el análisis de las universidades que permitieran sobrepasar muchas de estas limitaciones ofreciendo una visión más ajustada de la realidad universitaria. En efecto, además de las limitaciones técnicas que resultan de la tarea de realizar un ranking internacional de universidades, los principales problemas se derivan de las siguientes razones: (1) los criterios de selección de los indicadores no están científicamente argumentados (van Raan, 2005), (2) el carácter multidimensional de las universidades (Orduña-Malea, 2011) resulta en instituciones muy heterogéneas entre sí (Collini, 2011) y (3) la apuesta por rankings globales no considera el perfil disciplinario de las universidades (Visser y otros, 2007; Torres-Salinas y otros, 2011a). Un buen ejemplo de la complejidad que supone evaluar estas instituciones se evidencia en el trabajo de Bordons y otros (2010), en el que ponen de relieve los diferentes factores estructurales y contextuales que pueden incidir en la capacidad investigadora de las universidades. Tal complejidad hace poco recomendable el análisis global de las universidades a favor de análisis específicos de cada una de sus actividades. En este sentido, surgen nuevos rankings y metodologías dirigidos al análisis de alguna de las facetas de estas instituciones. Palomares-Montero, García-Aracil y Castro-Martínez (2008) ofrecen una interesante revisión bibliográfica de indicadores para el análisis de universidades.

Entre estas metodologías complementarias se encuentran los mapas de la ciencia que han demostrado ser herramientas muy útiles para la toma de decisiones en política científica (Noyons, 2004). Una de las propuestas más recientes en este sentido es el mapeo de universidades a través del Journal Publication Profile (JPP de ahora en adelante) presentado (García y otros, 2012). Este nuevo método de visualización propone un modelo de representación reticular de agentes científicos (universidades, países, etc.) basado en su similaridad a partir de su perfil de publicación en revistas científicas indexadas en los Journal Citation Reports (JCR), identificando el parecido de los agentes bajo estudio en función de las revistas comunes en las que éstos publican. Es decir, dos universidades se parecerán más entre sí si publican sus trabajos en las mismas revistas científicas, asemejándose así tanto en impacto como temáticamente. Como método final de visualización de los valores de similaridad los autores emplean el análisis de redes sociales, una metodología ampliamente utilizada en la literatura bibliométrica para estudiar las redes que se generan a partir de la cocitación (Small, 1973), la coautoría (Beaver, 2001) o la participación en tribunales de tesis (Delgado López-Cózar y otros, 2006) entre otros.

Teniendo en cuenta por un lado las limitaciones de los rankings para representar los sistemas de educación superior y por otro la utilidad de las propuestas de mapeo reseñadas, el principal objetivo de este estudio es la aplicación práctica de la metodología del JPP para realizar un análisis descriptivo en profundidad de las universidades españolas de forma que se pueda obtener un retrato mucho más certero del sistema universitario español. Más específicamente podemos establecer como objetivos:

1) La realización de mapas de la ciencia para conocer la similitud temática en investigación entre universidades españolas en función de su perfil de publicación en revistas científicas JCR en los últimos cinco años (2007-2011) para cinco grandes áreas: Ciencias Sociales, Ciencias Exactas, Ingeniería y Tecnología, Ciencias de la Vida y Ciencias de la Salud.

2) Estudiar el papel que juega cada universidad dentro de cada una de las áreas analizadas a partir del indicador de centralidad que nos proporciona el análisis de redes sociales. Se identificarán las universidades que tienen un papel preponderante dentro de un área así como las áreas en las que el perfil disciplinario de las universidades es más heterogéneo.

8.2. Material y métodos

8.2.1. Procesamiento y recogida de datos

En este estudio analizamos las universidades españolas en función de su perfil de publicación en revistas científicas. Para ello, en primer lugar seleccionamos los índices de revistas de la base de datos de Thomson Reuters Web of Science (de ahora en adelante WoS) como fuente de datos. Basamos esta decisión en la importancia que dicha base de datos tiene dentro del sistema de política científica española. El periodo de estudio corresponde al quinquenio 2007-2011; la elección de dicho periodo responde a la necesidad de establecer un margen lo suficientemente amplio para que los resultados ofrecidos sean estables y al mismo tiempo muestren una imagen lo más actual posible de la situación de las universidades. A continuación realizamos una búsqueda manual en el campo *address* de la WoS de las 77 universidades españolas teniendo en cuenta todas las variantes de firma de cada institución. Descargamos toda la producción de aquellas universidades con al menos 125 documentos citables (artículos, revisiones, notas o cartas), asignándola a cada institución e introduciéndola en una base de datos relacional para el cálculo de indicadores. Esto nos permite calcular el Índice de Actividad Temática (de ahora en adelante IAT). Este compara el porcentaje de trabajos que una universidad publica en un área determinada con el porcentaje que la población analizada dedica a dicha área.

También procedimos a la descarga de las ediciones de los JCR para el periodo de tiempo analizado y finalmente, para cada universidad solo tuvimos en cuenta la producción indexadas en los JCR. Asimismo los JCR nos permitieron calcular el porcentaje de publicaciones en revistas del primer cuartil para cada universidad. Construimos las cinco áreas científicas (Ciencias Sociales, Ciencias Exactas, Ingeniería y Tecnología, Ciencias de la Vida y Ciencias de la Salud) a partir de la agregación de las diferentes categorías temáticas de los JCR, método seguido en otros estudios como por ejemplo los 'Rankings I-UGR' (Torres-Salinas y otros, 2012). Finalmente, a la hora de representar los mapas de universidades y con el fin de facilitar su lectura, hemos introducido un rango de colores para cada universidad indicando el porcentaje de publicaciones en revistas del primer cuartil. Aunque no es necesario para establecer el perfil de publicación de las universidades, de este modo se enriquece el análisis permitiendo al lector observar qué universidades consiguen publicar un mayor porcentaje de documentos en revistas de alto impacto.

8.2.2. Metodología para el mapeo de universidades de acuerdo a su perfil de publicación en revistas científicas

Para realizar un análisis descriptivo de las universidades españolas en función de su perfil de publicación en revistas científicas es necesario, en primer lugar, identificar las revistas en las que publican los investigadores de una institución durante cierto periodo de tiempo. A partir de esta lista de revistas científicas se crea una matriz revista-universidad donde las filas contienen los pesos correspondientes a cada revista para todas las universidades. La idea básica de la que parte esta metodología es que dos universidades que publican en las mismas revistas tendrán un perfil más similar que otras dos que no lo hacen. Para generar los pesos de las revistas utilizamos un enfoque de frecuencia inversa (Salton y Buckley, 1988), ya que normalmente una revista en la que muchos investigadores de diferentes instituciones publican sus artículos será un mal indicador de la similaridad entre dos universidades. Siguiendo un enfoque de similaridad documento a documento (Ahlgren y Colliander, 2009) podremos inferir el comportamiento de la similaridad institución-institución a través de dos tipos de similaridades: de primer orden y de segundo orden. Las similaridades de primer orden se obtienen midiendo la similaridad entre columnas de la matriz revista-universidad. Sin embargo, podemos ir un paso más allá y obtenerlas midiendo la similaridad entre las columnas de esta matriz de similaridad de primer orden universidad-universidad.

Esta operación creará una nueva matriz universidad-universidad, en la que los valores indican similaridades de segundo orden. Cuando utilizamos similaridades de primer orden calculamos la similaridad directa entre dos universidades. Sin embargo, el enfoque de segundo orden determina que dos universidades son similares si la similaridad con el resto de universidades es semejante. De este modo, podemos aplicar el análisis de clúster para agrupar las universidades utilizando valores de disimilaridad de segundo orden universidad-universidad. Para realizar el análisis de clúster hemos empleado el método de vinculación completa (Everitt, Landau y Leese, 2001). Una vez agrupadas las universidades de acuerdo a su perfil de publicación en revistas, construimos los dendrogramas con dichas agrupaciones. Finalmente, creamos los mapas de universidades empleando el software Pajek versión 3.01. Sin embargo, la matriz resultante presenta problemas de visualización, al ofrecer un grafo completo con conexiones de todas a todas las universidades. Para resolver este problema se establecemos un umbral mínimo de similaridad de 0,70. La elección de un umbral con dicho valor nos permite eliminar del mapa aquellas conexiones entre universidades con menor valor de similaridad, visualizando únicamente las relaciones más fuertes que se observan en cada área, lo cual simplifica y clarifica el análisis del panorama universitario según el perfil de publicación en revistas. Para mayor detalle sobre la metodología empleada referimos al lector al trabajo de García y otros (2012).

8.2.3. Análisis de redes sociales: el concepto de centralidad

El análisis de redes sociales es una metodología importada del ámbito de la sociología y ampliamente utilizada en el ámbito de la bibliometría (Glänzel y Schubert, 2004; Wagner y Leydesdorff), que permite identificar estructuras subyacentes de la relación entre distintos actores y las situaciones de poder y subordinación que se dan entre ellos. Así pues, se consideran con más poder aquellos actores situados en posiciones 'ventajosas' o centrales, entendiendo por poder una mayor conexión o influencia con el resto de actores. Esta influencia o poder quedará definida en función del elemento de unión que se considere en el análisis, por lo que es preferible utilizar el término de 'centralidad' para referirse a dicho rol. Además, dicho análisis permite identificar ciertas propiedades y medidas que permiten caracterizar la estructura de las distintas redes. Para más información acerca de dichas propiedades y medidas referimos al lector a Sanz Menéndez (2003). Aquí destacaremos dos de ellas de cara a la interpretación de los resultados: el grado de exclusividad, que muestra las posibilidades de acceso que tienen los actores externos a una red determinada (a mayor grado, mayor dificultad); y la propiedad de transitividad, que indica la probabilidad de que dos vecinos de un nodo con una similaridad fuerte a este nodo sean a su vez similares entre sí.

En el presente trabajo, el uso del indicador de centralidad permitirá analizar en profundidad cómo se agrupan las universidades por área. En este sentido, aquella universidad con un mayor valor de centralidad en un área determinada será aquella que publica en un mayor número de trabajos en revistas comunes con el resto de universidades de la red. Para calcular la centralidad de los actores de una red existen tres indicadores: grado, cercanía e intermediación (Delgado López-Cózar y otros, 2006). En este trabajo utilizaremos la cercanía, que mide la capacidad de cada uno de los actores de la red de conectar con otros directamente y sin necesidad de intermediarios. De este modo, identificamos a aquellas universidades que guardan similaridad con un mayor número de universidades dentro de un área.

8.2.4. Material complementario On-line

Para facilitar la comprensión del artículo hemos elaborado una web (http://www.ugr.es/~elrobin/jpp.html) en la que el lector puede encontrar material complementario. Además, hemos elaborado un documento en el que se incluyen datos adicionales derivados del estudio (Robinson-García y otros, 2012). En este documento incluimos los siguientes datos para cada una de las cinco áreas analizadas. En primer lugar, los indicadores bibliométricos necesarios para el cálculo del perfil de las universidades, esto es: producción y porcentaje de documentos publicados en revistas indexadas en el primer cuartil. En segundo lugar, ofrecemos los dendrogramas derivados de los mapas de universidades de cada área. Finalmente, hemos introducido árboles de expansión mínima que sirven de apoyo visual para analizar la estructura de los mapas de universidades.

8.3. Resultados

8.3.1. Resultados generales

De las 77 universidades españolas que conforman el sistema universitario español, 57 de ellas tienen una producción citable superior a 125 documentos para el periodo 2007-2011. En la tabla 8.1 mostramos la producción (Prod) y el porcentaje de publicaciones en revistas del primer cuartil (%Prod Q1) de estas universidades. La universidad más productiva es Barcelona con 15940 documentos, seguida por la Autónoma de Barcelona (12060 documentos) y la Complutense de Madrid (11346 documentos). La universidad con una mayor proporción de sus documentos publicados en revistas del primer cuartil (Q1) es Pompeu Fabra con un 59% de su producción en revistas Q1, seguida por Barcelona (57%).

Cinco universidades copan las primeras posiciones como las más productivas en cuatro de las cinco áreas analizadas (Ciencias Sociales, Ciencias Exactas, Ciencias de la Vida y Ciencias de la Salud) con algunas excepciones (Robinson y otros, 2012). En primer lugar se posiciona Barcelona, que ocupa el primer puesto en las cuatro áreas. A continuación, la Autónoma de Barcelona, que es la segunda universidad más productiva en tres de las cuatro áreas seguida de Complutense que se sitúa siempre entre la segunda y la quinta posición, Valencia (entre la tercera y la sexta universidad más productiva por área) y la Autónoma de Madrid (ocupa entre el cuarto y el sexto puesto). En el área restante, Ingeniería y Tecnología, se ve un patrón de producción muy distinto, relegando a las universidades anteriores. Así pues, las universidades más productivas en esta área son: Politécnica de Cataluña, Politécnica de Madrid y Politécnica de Valencia. En relación al porcentaje de documentos en el primer cuartil, Barcelona es la tercera con más publicaciones en revistas Q1 (36%) en Ciencias Sociales, la segunda en Ciencias Exactas y la segunda en Ciencias de la Salud (54%). Otra de las universidades a destacar es la Pompeu Fabra, que es la universidad con un mayor % Prod Q1 en las áreas de Ciencias Sociales (43%) y Ciencias de la Salud (59%), la segunda en Ciencias de la Vida (70%) y la tercera en Ciencias Exactas (68%). En el área de Ingeniería y Tecnología destaca Córdoba que se sitúa como la segunda universidad con mayor Prod Q1 (67%), Valencia ocupa el tercer puesto (64%) y Barcelona (62%) y la Autónoma de Barcelona (61%) ocupan los puestos quinto y sexto respectivamente.

Tabla 8.1. Datos generales de producción y porcentaje de documentos en revistas del primer cuartil de las universidades españolas. Periodo 2007-2011

Universidad	Prod	% Q1	Universidad	Prod	% Q1
Barcelona	15940	57%	Carlos III Madrid	2403	41%
Autónoma Barcelona	12060	55%	Islas Baleares	2110	56%
Complutense	11346	48%	Miguel Hernández	2072	50%
Valencia	9730	52%	Girona	2005	55%
Autónoma Madrid	9228	55%	Rey Juan Carlos	1937	48%
Granada	7787	46%	A Coruña	1920	39%
Politécnica Cataluña	6841	52%	Jaén	1796	42%
País Vasco	6821	52%	Jaume I	1733	50%
Zaragoza	6236	52%	Cádiz	1580	49%
Sevilla	6178	50%	Almería	1542	43%
Santiago Compostela	6147	50%	Lleida	1470	53%
Politécnica Madrid	5779	41%	UNED	1399	35%
Politécnica Valencia	5284	51%	Las Palmas	1389	42%
Oviedo	4548	51%	Pública Navarra	1382	45%
Murcia	3900	43%	León	1338	46%
Vigo	3748	49%	Politécnica Cartagena	1131	44%
Salamanca	3690	46%	Huelva	1104	50%
Castilla La Mancha	3687	51%	Pablo Olavide	851	47%
Navarra	3384	45%	Burgos	666	56%
Alicante	3378	46%	La Rioja	641	49%
La Laguna	3183	50%	Ramón Llull	524	36%
Valladolid	3046	45%	Cardenal Herrera	312	38%
Rovira i Virgili	3018	56%	Europea de Madrid	296	39%
Córdoba	2948	56%	San Pablo CEU	261	49%
Pompeu Fabra	2911	59%	Pontificia Comillas	191	47%
Málaga	2812	45%	Católica S Antonio	179	23%
Cantabria	2537	53%	Oberta Catalunya	165	24%
Alcalá	2509	47%	Deusto	133	21%
Extremadura	2426	50%			

Tabla 8.2. Índice de actividad temática de las universidades españolas para cinco áreas científicas. Periodo 2007-2011

Universidad	CSS	CE	IT	CV	CS	Universidad	CSS	CE	IT	CV	CS
Barcelona	0,82	0,8	0,39	1	1,61	Carlos III Madrid	**2,46**	1,2	**2,77**	0,1	0,1
Autónoma Barcelona	0,98	0,8	0,6	1,1	1,42	Islas Baleares	1,08	1	0,63	1,3	0,8
Complutense	0,92	1	0,61	1	1,04	Miguel Hernández	0,92	0,8	0,61	1,2	1,7
Valencia	1,12	1,1	0,5	0,8	1,24	Girona	0,95	1	1,1	1,4	0,6
Autónoma Madrid	0,83	1,1	0,58	0,9	1,12	Rey Juan Carlos	1,05	0,9	1,23	0,9	1,3
Granada	1,51	0,9	0,84	1	0,98	A Coruña	0,98	0,9	1,11	1,2	0,9
Politécnica Cataluña	0,27	1,4	**2,72**	0,8	0,27	Jaén	1,1	1,1	1,28	1,1	0,7
País Vasco	1,02	1,4	1,09	0,7	0,67	Jaume I	**2,08**	1,3	1,54	0,5	0,3
Zaragoza	0,83	1,3	1,12	0,9	0,83	Cádiz	0,61	1	1	1,4	0,8
Sevilla	0,84	1,2	1,37	1	0,69	Almería	1,43	1,1	0,9	1,66	0,44
Santiago Compostela	0,66	1,2	0,54	1,1	1,19	Lleida	0,73	0,8	0,6	1,84	0,92
Politécnica Madrid	0,42	1	**2,44**	1,1	0,36	UNED	**3,1**	1	1,3	0,29	0,4
Politécnica Valencia	0,6	1,4	**2,39**	0,9	0,31	Las Palmas	1,24	0,5	0,8	1,4	1,33
Oviedo	1,18	1	1,25	0,9	0,9	Pública de Navarra	1,72	1,1	1,9	0,93	0,5
Murcia	1,35	0,8	0,48	1,2	1,12	León	0,68	0,4	0,4	**2,25**	0,88
Vigo	0,63	1,2	1,27	1,3	0,62	Politécnica Cartagena	1,1	1,3	2,5	1	0,13
Salamanca	1	0,7	0,42	1	1,27	Huelva	1,03	0,9	1,1	1,52	0,33
Castilla La Mancha	0,9	1	1,46	1,4	0,75	Pablo Olavide	**2,8**	0,6	0,7	1,29	0,91
Navarra	0,89	0,3	0,51	0,6	**2,09**	Burgos	1,24	1,8	1	0,69	0,33
Alicante	1,14	1,1	1,04	0,8	0,9	La Rioja	0,81	1,6	0,8	1,18	0,41
La Laguna	1,1	1,4	0,61	0,9	0,8	Ramón Llull	**3,74**	0,8	1,3	0,63	0,63
Valladolid	0,79	1,2	1,17	0,7	1	Cardenal Herrera	0,62	0,4	0,3	1,3	**2,26**
Rovira i Virgili	1,04	1,2	1,18	1	1,1	Europea de Madrid	0,88	0,2	0,5	0,43	**2,86**
Córdoba	0,32	0,9	0,79	1,8	0,9	San Pablo CEU	0,69	1	0,5	1,04	1,74
Pompeu Fabra	**2,18**	0,3	0,47	1,1	1,5	Pontificia Comillas	1,75	1	**3,1**	0,47	0,1
Málaga	1,26	0,9	1,31	1	0,7	Católica S Antonio	**2,17**	0,9	0,4	0,57	**2,12**
Cantabria	0,73	1	1,36	0,7	1,1	Oberta de Catalunya	**4,48**	0,2	1,8	0,31	0,69
Alcalá	0,83	0,7	1,11	1,1	1,4	Deusto	**5,66**	0,1	0,8	0,15	0,91
Extremadura	0,84	1	1,11	1,3	0,9						

Siguiendo el análisis de la producción de universidades por áreas en función de su índice de actividad temática (de ahora en adelante IAT) (Tabla 8.2), observamos nuevamente el perfil claramente enfocado hacia la Ingeniería y Tecnología de las universidades politécnicas. No obstante, éstas no son las únicas que muestran una mayor actividad en un área determinada. Así, Pompeu

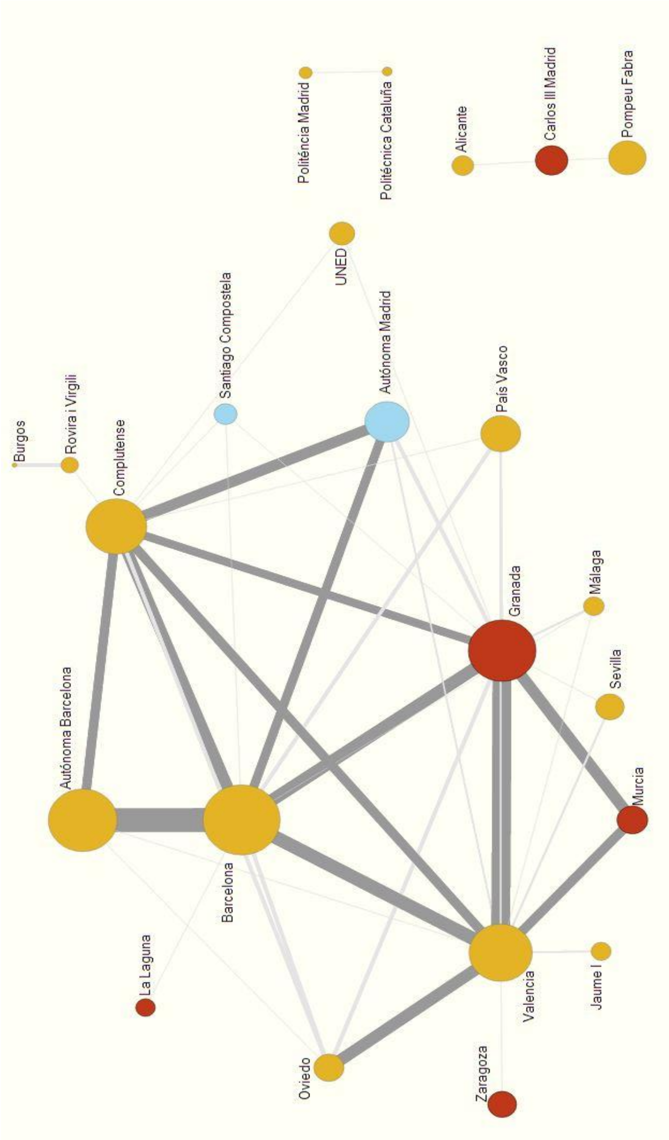
Fabra tiene un alto IAT (2,18) en Ciencias Sociales al igual que Jaume I (2,08), UNED (3,10), Pablo Olavide (2,80) y Ramón Llull (3,74). Por su parte, Carlos III de Madrid destaca en Ciencias Sociales (2,46) e Ingeniería y Tecnología (2,77), León en Ciencias de la Vida (2,25) y Navarra en Ciencias de la Salud (2,09). Por último, señalar la alta especialización de aquellas universidades con una menor producción como Cardenal Herrera, Pontificia de Comillas, Católica de San Antonio o Deusto entre otras.

8.3.2. Mapas de universidades de acuerdo a su perfil de publicación en revistas para cada área científica

A continuación mostramos los mapas de similaridad de las universidades españolas en cinco grandes áreas (Figuras 8.1-8.5). Cada nodo representa una universidad. El grosor del nodo o vértice viene dado por el porcentaje de documentos publicados en revistas Q1. Los enlaces entre los vértices representan la semejanza entre las instituciones implicadas. Un enlace más grueso responde a una mayor similaridad.

En la Figura 8.1 mostramos el mapa en el área de Ciencias Sociales. En ella se observan tres perfiles claramente diferenciados. Un grupo principal, formado por cinco universidades de gran producción (Granada, Barcelona, Valencia, Complutense y Autónoma de Barcelona) rodeadas por universidades de menor impacto que se sitúan en la periferia. Este grupo aglutina a la mayoría de las universidades y se caracteriza por tener un pequeño grupo de universidades en el centro muy similares entre sí a las cuáles conectan el resto de universidades. Alrededor de cada uno de estos nodos se aglutinan universidades de menor producción. Por otro lado, destacar el gran parecido entre el perfil de publicación de la Autónoma de Barcelona y Barcelona. El segundo perfil de universidades que se observa es el de las politécnicas (Politécnica de Madrid y Politécnica de Cataluña), demostrando un patrón de publicaciones en revista diferente al resto. Estas dos universidades se caracterizan por su poca producción. El tercer caso es el de Carlos III de Madrid, Pompeu Fabra y Alicante. Carlos III de Madrid y Pompeu Fabra son especialmente productivas en el área de Ciencias Sociales de acuerdo a su IAT (Tabla 8.2). Si comparamos estos resultados con los de la tercera edición de los Rankings I-UGR de universidades españolas (Torres-Salinas y otros, 2012), vemos que estas universidades se sitúan en las primeras posiciones del ranking de Economía para el mismo periodo de tiempo analizado (Pompeu Fabra en primer lugar, Carlos III de Madrid en cuarta posición y Alicante en sexto lugar). Este fuerte perfil en el área de Economía explica tal agrupación.

Figura 8.1. Mapa de similaridad de las universidades españolas de acuerdo a su perfil de publicación en revistas en Ciencias Sociales



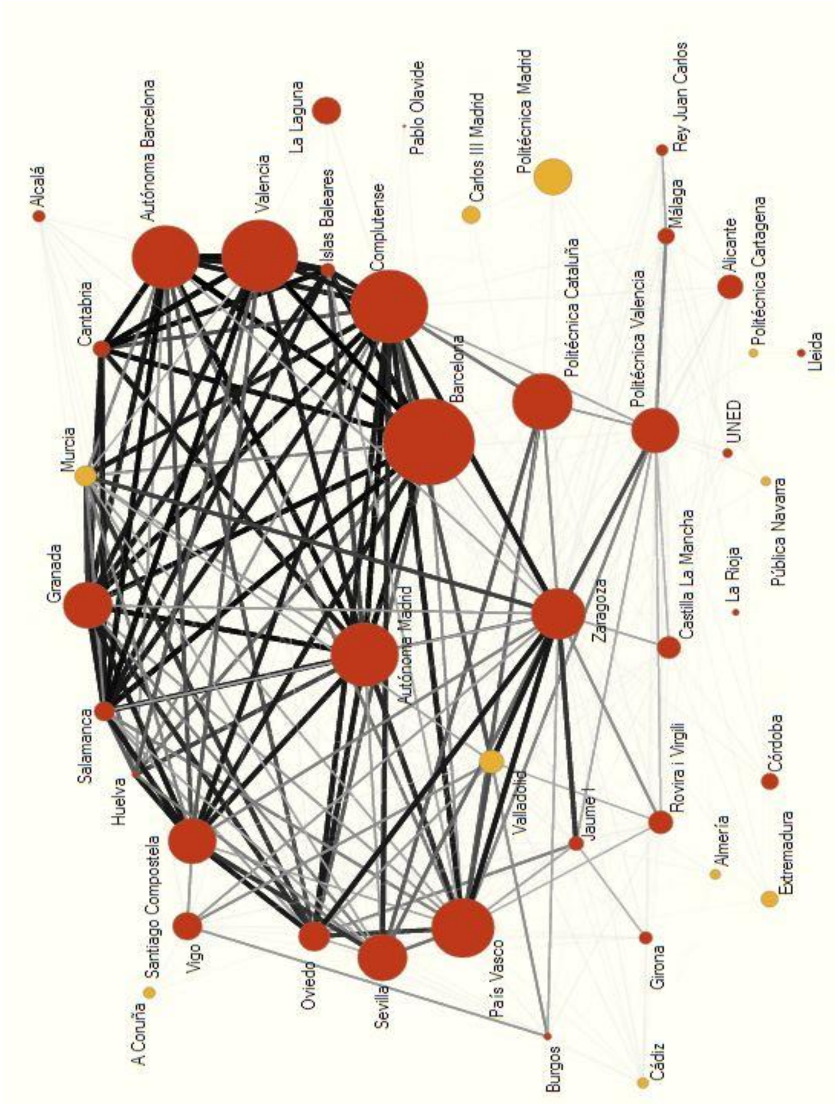
Características del mapa: Líneas > valor de similaridad mínimo 0,70; valor de similaridad máximo 0,84. Se han eliminado los nodos aislados. Colores > rojo: > 30% de la producción perteneciente a revistas Q1; amarillo: 15-30% de la producción perteneciente a revistas Q1; azul < 15% de la producción perteneciente a revistas Q1

Muy distinto es el retrato del área de Ciencias Exactas que mostramos en la Figura 8.2. Una red muy densa en la que prácticamente todas las universidades tienen patrones de similaridad muy fuertes entre sí. En este caso, la red se caracteriza por cumplir en mayor medida que en el resto de áreas la propiedad de transitividad. Esta propiedad se refiere a que existe una probabilidad alta de que dos vecinos de un nodo, ambos con una similaridad fuerte a este nodo, sean a su vez similares entre ellos con respecto a sus perfiles de publicación en revistas. El elevado nivel de agrupamiento de la red implica un alto grado de exclusividad, algo que no ocurre en el resto de áreas.

La mayoría de las universidades tienen una producción alta o media con más del 55% de su producción perteneciente a revistas Q1. Se observa un núcleo con universidades muy similares entre sí (como por ejemplo, Complutense, Autónoma de Barcelona y Valencia) con protagonismos singulares como el caso de Zaragoza que se une además a universidades de la periferia como Castilla La Mancha, Rovira i Virgili o Politécnica de Valencia. Algunas de las universidades situadas en la periferia se agrupan en función de factores geográficos (por ejemplo las universidades gallegas). Sin embargo esto no ocurre en otros casos como por ejemplo la alta similaridad existente entre la Politécnica de Valencia y Málaga, y Rey Juan Carlos y Málaga por otro lado, mostrando líneas de investigación cercanas entre los investigadores de estas instituciones. Si bien la red integra a las universidades politécnicas al contrario de lo que ocurre en otras áreas, es destacable el comportamiento de dos universidades de perfil politécnico, Politécnica de Madrid y Carlos III de Madrid (como se observa en la tabla 8.2), que se sitúan fuera del núcleo del área aún teniendo una producción relativamente elevada. Esto indica que publican en revistas distintas al resto, incluyendo a las otras universidades politécnicas.

La figura 8.3 muestra el mapa de universidades en el área de la Ingeniería y la Tecnología. En este caso vemos una red bipolar con dos núcleos claramente diferenciados con distintos perfiles de publicación unidos a través de Zaragoza que tiene gran similaridad tanto con la Complutense como con la Politécnica de Cataluña. Esta bipolaridad se constata al observar el árbol de expansión mínima (Robinson-García y otros, 2012). A la izquierda de la figura se agrupan las tres universidades politécnicas con un papel destacado de la Politécnica de Cataluña como eje central. Además, destacan Sevilla y Carlos III de Madrid. La Politécnica de Valencia muestra una gran similaridad con la Politécnica de Cataluña. El otro núcleo está formado por las grandes universidades generalistas españolas, esto es; Complutense en posición central, Barcelona, Autónoma de Madrid y Autónoma de Barcelona. A este grupo se unen Santiago de Compostela, Valencia y País Vasco. Es reseñable la posición de Vigo que aunque es similar a Santiago de Compostela, también muestra una gran similaridad con Zaragoza, Politécnica de Valencia, Politécnica de Cataluña y Carlos III de Madrid, mostrando de este modo un comportamiento más similar al de éstas últimas.

Figura 8.2. Mapa de similaridad de las universidades españolas de acuerdo a su perfil de publicación en revistas en Ciencias Exactas



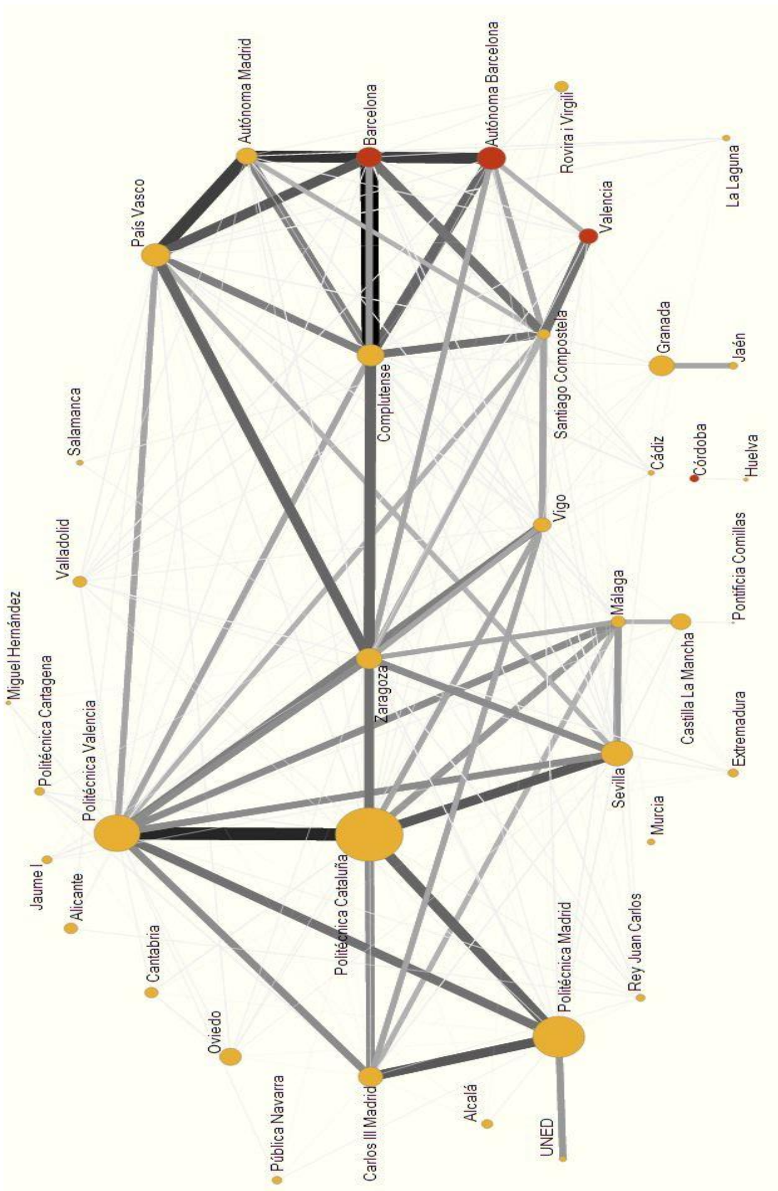
Características del mapa: Líneas > valor de similaridad mínimo 0,70; valor de similaridad máximo 0,98. Se han eliminado los nodos aislados. Colores > rojo: > 55% de la producción perteneciente a revistas Q1; amarillo: 40-55% de la producción perteneciente a revistas Q1; azul < 40% de la producción perteneciente a revistas Q1

El comportamiento de las universidades andaluzas que pone de relieve una de las principales facetas de las universidades que muestran los mapas de acuerdo al perfil de publicación en revistas: la similaridad en función del perfil disciplinario. Las siete universidades andaluzas que mostramos en este mapa se sitúan en la parte inferior de la figura, sin embargo, se observan tres subgrupos. Por un lado estarían Málaga y Sevilla con un perfil más cercano al de las universidades politécnicas. Después está Granada y Jaén, muy similares entre sí y más apegadas al grupo de las grandes universidades generalistas. Y por último, en una situación intermedia, se ubican Cádiz, Córdoba y Huelva.

En la figura 8.4 mostramos el mapa de universidades para el área de Ciencias de la Vida. Todas las universidades se agrupan en torno a la Complutense, la Autónoma de Barcelona y Barcelona, que forman el núcleo de la red. A excepción de dos universidades, Carlos III y UNED, que aparecen unidas entre sí y aisladas del resto de universidades, pero con una producción de poco tamaño e impacto. También observamos una gran similaridad de Barcelona con Valencia y País Vasco. Asimismo, destaca el aislamiento de Córdoba, Lleida, Politécnica de Valencia y Politécnica de Cartagena, que aún estando incluidas en la red, aparecen desplazadas y creando un pequeño subgrupo. Es muy relevante el papel de Pompeu Fabra, la única universidad con un porcentaje de producción en revistas Q1 superior al 60% que se muestra en el mapa y, por tanto, muestra una baja similaridad con el resto de la red.

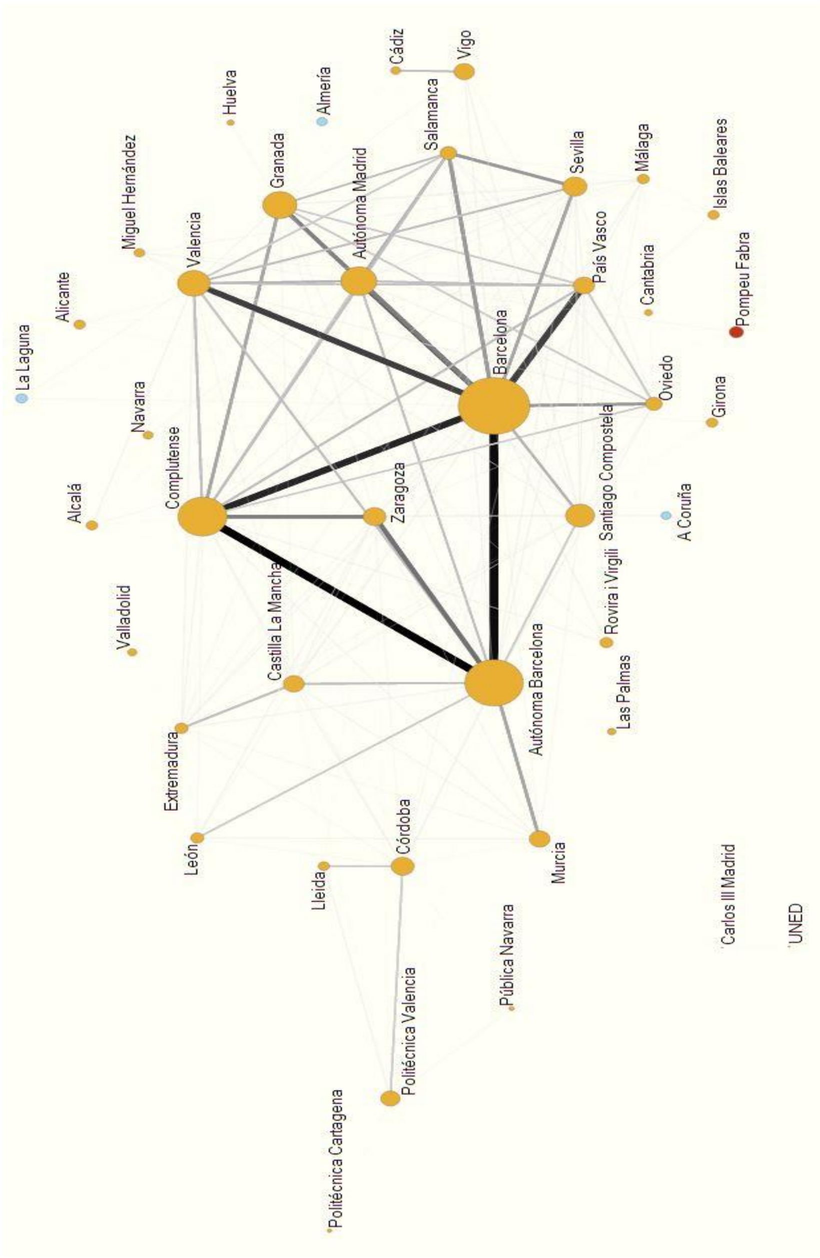
El mapa de Ciencias de la Salud (Figura 8.5) muestra cierto parecido con el anterior, en tanto en cuanto observamos en primer lugar una gran red que gira en torno a un núcleo formado por tres universidades principalmente (Autónoma Madrid, Autónoma Barcelona y Barcelona). Alrededor de estas tres universidades hay otras similares al núcleo, con un tamaño de producción medio pero también de alto impacto (más del 45% de la producción perteneciente a revistas Q1). Estas universidades son Valencia y Navarra. También Complutense se uniría a este grupo si no fuera por tener una producción con un impacto medio (su producción en revistas Q1 se sitúa entre el 25 y el 45% del total). Valencia y Complutense aglutinan a su vez, alrededor suya una red de universidades con perfiles muy similares a los suyos. Por ejemplo, Valencia y Oviedo o Complutense y Santiago de Compostela. El caso de Navarra es reseñable al dedicar un mayor esfuerzo a este área (Tabla 8.2), comportándose claramente como una institución especializada en el ámbito de Ciencias de la Salud.

Figura 8.3. Mapa de similaridad de las universidades españolas de acuerdo a su perfil de publicación en revistas en Ingeniería y Tecnología



Características del mapa: Líneas > valor de similaridad mínimo 0,70; valor de similaridad máximo 0,88. Se han eliminado los nodos aislados. Colores > rojo: > 60% de la producción perteneciente a revistas Q1; amarillo: 35-60% de la producción perteneciente a revistas Q1; azul < 35% de la producción perteneciente a revistas Q1

Figura 8.4. Mapa de similaridad de las universidades españolas de acuerdo a su perfil de publicación en revistas en Ciencias de la Vida



Características del mapa: Líneas > valor de similaridad mínimo 0,70; valor de similaridad máximo 0,90. Se han eliminado los nodos aislados. Colores > rojo: > 60% de la producción perteneciente a revistas Q1; amarillo: 45-60% de la producción perteneciente a revistas Q1; azul < 45% de la producción perteneciente a revistas Q1

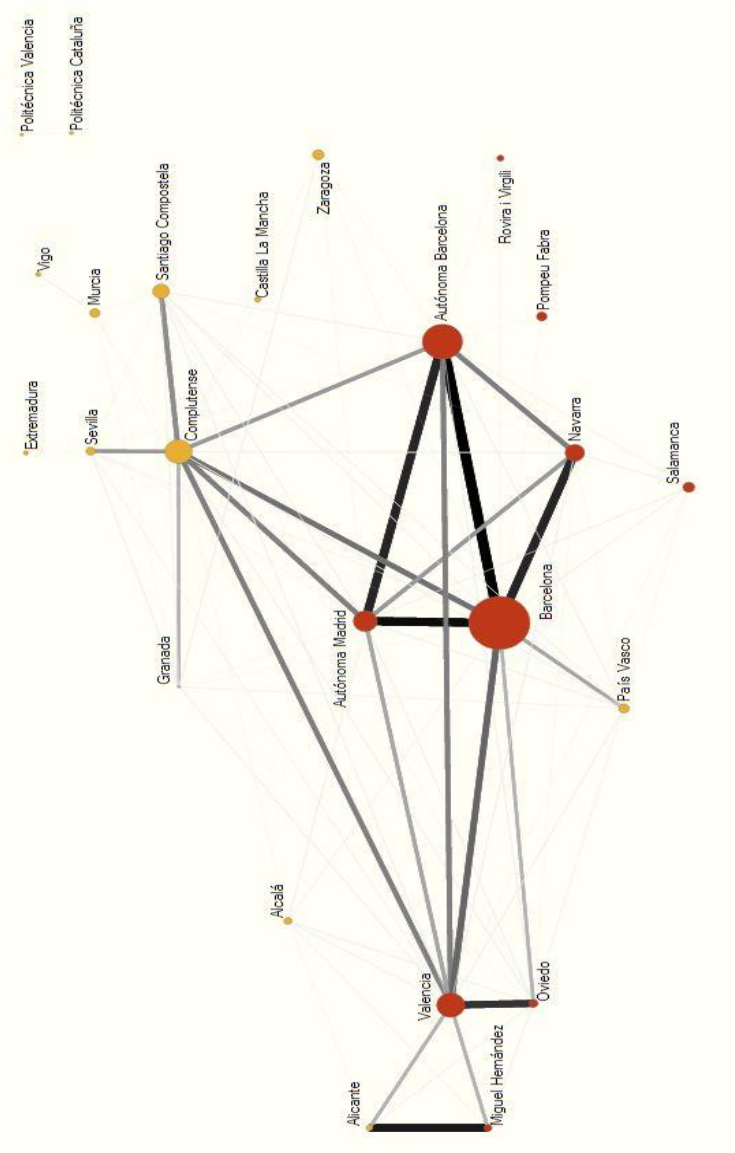
Por otra parte, Politécnica de Valencia y Politécnica de Cataluña aparecen relacionadas pero aisladas del resto de la red, mostrando nuevamente un perfil diferente al del resto de las universidades. Finalmente, indicar nuevamente el alto impacto de Pompeu Fabra que, aun estando incluida en la red, muestra unos patrones de similaridad bajos con el resto de las universidades.

En la Tabla 3 mostramos las similaridades promedio de las 57 universidades analizadas. Observamos que Ciencias Exactas es el área en el que las universidades muestran una mayor similaridad promedio (0,52), mucho más alta que en los casos de Ciencias de la Salud (0,40) y Ciencias Sociales (0,42). La similaridad promedio de Ciencias Exactas no es tan diferente al compararla con Ingeniería y Tecnología (0,51), sin embargo, son mapas con estructuras muy diferentes. En el primer caso se observa una red unipolar y en el segundo, una red bipolar. Mientras en Ciencias Exactas existe una única comunidad muy similar de acuerdo a su perfil de publicación en revistas que engloba a la mayoría de las universidades, en Ingeniería y Tecnología existen dos comunidades (universidades politécnicas y universidades generalistas) con distintos perfiles de publicación científica. Además existe también una gran diferencia entre los valores máximos de similaridad alcanzados en un caso y en otro. Concretamente el valor máximo de similaridad en Ingeniería y Tecnología es de 0,88; mientras que la máxima similaridad observada en Ciencias Exactas alcanza un valor de 0,98, por tanto muy superior. De hecho es precisamente en Ciencias Exactas donde se observan los valores más elevados de similaridad.

Tabla 8.3. Valor promedio de similaridad de acuerdo al perfil de publicación en revistas para cada área científica

Area	Máximo	Media	Desviación	Coeficiente de Variación
Ciencias Sociales	0,84	0,43	0,16	0,38
Ciencias Exactas	0,98	0,52	0,23	0,45
Ingeniería y Tecnología	0,88	0,51	0,20	0,38
Ciencias de la Vida	0,90	0,48	0,22	0,45
Ciencias de la Salud	0,95	0,40	0,20	0,50

Figura 8.5. Mapa de similaridad de las universidades españolas de acuerdo a su perfil de publicación en revistas en Ciencias de la Salud



Características del mapa: Líneas > valor de similaridad mínimo 0,70; valor de similaridad máximo 0,95. Se han eliminado los nodos aislados. Colores > rojo: > 45% de la producción perteneciente a revistas Q1; amarillo: 25-45% de la producción perteneciente a revistas; azul < 25% de la producción perteneciente a revistas

8.3.3. Análisis de la centralidad de las universidades

En la Tabla 8.4 mostramos los valores de centralidad de las universidades españolas en las cinco áreas analizadas con el objetivo de identificar cuáles son las que juegan un rol mayor de acuerdo a su perfil de publicación en revistas científicas. Para ello utilizamos como indicador de centralidad la cercanía. El orden en el que aparecen las universidades se establece en función a dos criterios: a) aparecer en un mayor número de áreas y b) un mayor valor de cercanía promedio. En este sentido, se observa claramente cómo hay cuatro grupos de universidades. El primer grupo estaría formado por aquellas universidades que tienen un perfil de publicación muy heterogéneo y que por tanto, son fácilmente comparables con el resto de universidades en cualquiera de las cinco áreas analizadas.

En segundo lugar, se encontrarían universidades que tienen un alto grado de similaridad únicamente en ciertas áreas, siendo mucho menor en el resto. En éste caso serían comparables solamente en las áreas en las que tienen un mayor grado de cercanía. Sería el caso de Vigo en Ciencias Exactas y Ciencias de la Vida o Carlos III de Madrid en Ingeniería y Tecnología. El tercer grupo de universidades es el formado por universidades con bajos valores de cercanía en todas las áreas, es decir con perfiles de publicación diferentes del resto de las universidades, pero que aún así tienen una producción media en ciertas áreas con un porcentaje considerable situado en revistas Q1. Aquí, el caso más paradigmático sería el de Pompeu Fabra, ya mencionado anteriormente y que sin embargo su grado de cercanía en ningún caso supera el 0,5. En el último grupo se encontrarían el resto de universidades, de menor producción y que no llegan a ocupar posiciones de centralidad significativas en ninguna de las áreas.

La universidad mejor ubicada en todas las áreas es la Complutense de Madrid, que se sitúa entre las tres universidades con mayor grado de cercanía en las áreas de Ciencias Sociales, Ciencias Exactas y Ciencias de la Salud. Sevilla aparece en segundo lugar, aunque no logra en ningún caso situarse entre las tres top y País Vasco, que sí que logra estar entre las tres universidades con mayor grado de centralidad en Ciencias Exactas. Barcelona también destaca al situarse entre las tres primeras posiciones en Ciencias Sociales, Ciencias de la Vida y Ciencias de la Salud. Otras universidades a destacar son Zaragoza (entre las tres primeras universidades en Ciencias Exactas e Ingeniería y Tecnología) y Valencia (en Ciencias Sociales y Ciencias de la Salud). El caso de Zaragoza es especialmente relevante ya que, según se observa, hace de nexo de unión entre el núcleo y la periferia en Ciencias Exactas (Figura 8.2) y entre los dos grupos de universidades en Ingeniería y Tecnología (Figura 8.3).

Tabla 8.4 Grado de cercanía de las universidades españolas en las cinco áreas analizadas.

Universidades	CSS	CE	IT	CV	CS	Universidades	CSS	CE	IT	CV	CS
Complutense	**0,55**	**0,86**	0,74	0,62	**0,69**	Cádiz	--	0,62	0,57	0,44	--
Sevilla	0,37	0,81	0,71	0,63	0,57	Pompeu Fabra	0,09	--	--	0,43	0,41
País Vasco	0,43	**0,82**	0,67	0,61	0,58	Córdoba	--	0,57	0,47	0,56	--
Barcelona	**0,53**	0,62	0,64	**0,74**	**0,69**	Cantabria	--	0,58	0,49	0,43	--
Zaragoza	0,34	**0,89**	**0,77**	0,62	0,47	Huelva	--	0,58	0,32	0,38	--
Granada	**0,55**	0,70	0,57	0,63	0,60	Pública Navarra	--	0,52	0,49	0,37	--
Oviedo	0,44	0,76	0,57	0,57	0,62	Politécnica Cartagena	--	0,44	0,51	0,31	--
Santiago Compostela	0,39	0,69	0,63	0,62	0,62	Girona	--	0,58	--	0,48	--
Autónoma Barcelona	0,43	0,60	0,60	**0,66**	0,66	Navarra	--	--	--	0,43	0,55
Valencia	**0,58**	0,58	0,56	0,63	**0,71**	Rey Juan Carlos	--	0,59	0,58	--	--
Autónoma Madrid	0,43	0,61	0,64	0,60	0,64	Burgos	0,25	0,60	--	--	--
Murcia	0,37	0,78	0,42	0,52	0,48	Islas Baleares	--	0,60	--	0,44	--
Rovira i Virgili	0,35	0,67	0,54	0,48	0,41	Almería	--	0,52	--	0,38	--
Alicante	0,09	0,64	0,50	0,46	0,44	A Coruña	--	0,50	--	0,38	--
Málaga	0,39	0,65	0,65	0,56	--	Lleida	--	0,31	--	0,42	--
Politécnica Cataluña	0,09	0,70	**0,77**	--	0,08	León	--	--	--	0,45	--
Castilla La Mancha	--	0,68	0,56	**0,65**	0,40	Jaén	--	0,46	--	--	--
Politécnica Valencia	--	0,75	**0,82**	0,45	0,08	Las Palmas	--	--	0,39	--	--
Vigo	--	0,66	0,73	0,48	0,32	Pontificia Comillas	--	--	0,36	--	--
Salamanca	--	0,63	0,46	0,57	0,49	Pablo Olavide	--	0,50	--	--	--
Alcalá	--	0,62	0,49	0,46	0,53	La Rioja	--	0,48	--	--	--
Extremadura	--	0,55	0,52	0,59	0,36	Cardenal Herrera	--	--	--	--	--
Carlos III Madrid	0,13	0,51	0,66	0,05	--	Católica S Antonio	--	--	--	--	--
La Laguna	0,32	0,48	0,47	0,44	--	Deusto	--	--	--	--	--
UNED	0,37	0,58	0,42	0,05	--	Europea de Madrid	--	--	--	--	--
Jaume I	0,34	0,68	0,53	--	--	Oberta de Catalunya	--	--	--	--	--
Valladolid	--	0,72	0,62	0,40	--	Ramón Llull	--	--	--	--	--
Politécnica Madrid	0,09	0,55	0,70	--	--	San Pablo CEU	--	--	--	--	--
Miguel Hernández	--	--	0,50	0,47	0,44						

8.4. Discusión

En este trabajo presentamos un análisis descriptivo de las universidades españolas según su perfil de publicación en revistas científicas durante el quinquenio 2007-2011 en cinco grandes áreas. Para ello, hemos mostrado en primer lugar los datos generales de producción de las universidades en el periodo analizado así como su IAT para cada una de las áreas analizadas. En segundo lugar, hemos mostrado los mapas de similaridad de universidades por área en los que se observan las agrupaciones resultantes. Finalmente, hemos introducido los valores de cercanía de las universidades para así complementar la información ofrecida en los mapas e identificar las universidades que juegan un mayor rol en cada área y poder tipificarlas. En las cinco áreas científicas analizadas se observa la presencia notable de universidades catalanas y madrileñas siempre situándose en posiciones centrales de las redes. Un predominio que concuerda con su posición en los distintos rankings de universidades (Torres-Salinas y otros, 2011b; Docampo y otros, 2012). Además, al relacionar la posición de estas universidades con el tamaño de su producción, muestra cómo las universidades más grandes tienden a tener valores mayores de centralidad al poder abordar distintos frentes de investigación.

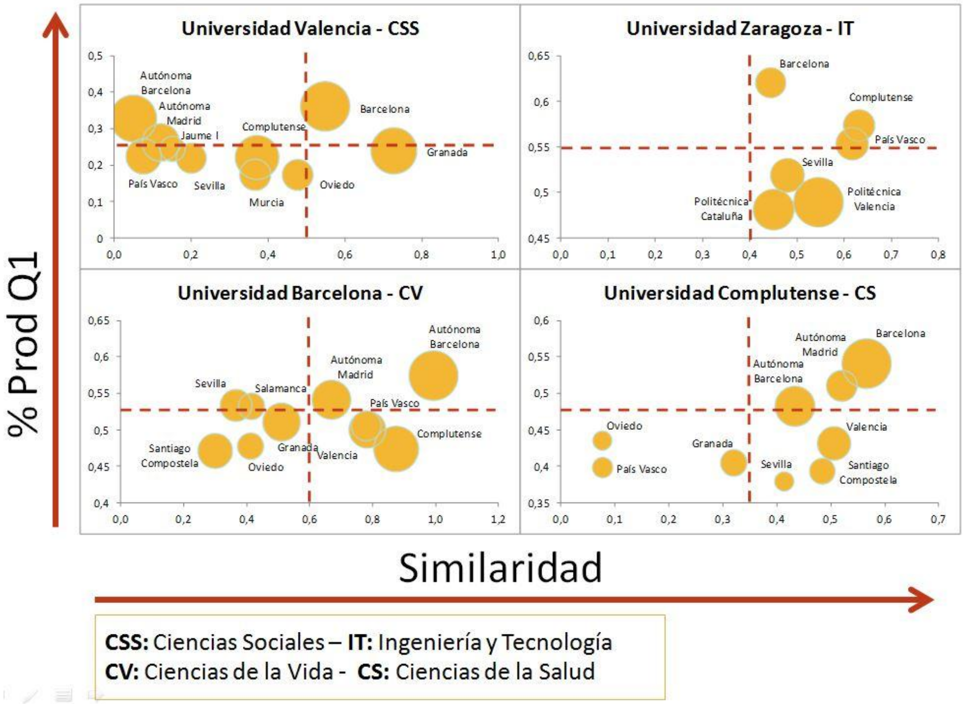
Sin embargo, la estructura de las redes difiere según el área analizada. Así pues, el área de Ingeniería y Tecnología (Figura 8.3) muestra una red bipolar distinta del resto, que tienen un único núcleo de universidades en torno al que gira el resto de los miembros de la red. Mientras que en este caso se observan dos perfiles de publicación claramente diferenciados, con las universidades politécnicas por un lado y las universidades generalistas por otro, en el resto de áreas se muestra un único perfil principal, aunque bien es cierto que las universidades politécnicas tienden a aparecer apartadas del resto de universidades. No obstante, el grado de cohesión de la red en Ciencias Exactas es mucho mayor que en el resto de las áreas, siendo ésta la única área en la que un número significativo de universidades politécnicas aparece completamente integrado en la red. Se trata del área con mayor similaridad promedio (Tabla 8.3) y en la que un mayor número de universidades tiene un grado de cercanía superior a 0,8 (Tabla 8.4), revelando cierto predominio en el área por parte de éstas. Es decir, en esta área tiene mayor sentido comparar universidades al tener todas un perfil de publicación en revistas muy similar entre sí. Esta red evidencia un perfil de publicación muy homogéneo con una elevada transitividad entre los nodos del mapa de similaridad.

Es importante señalar el caso paradigmático de Pompeu Fabra, una universidad claramente especializada en las áreas de Ciencias Sociales y Ciencias de la Salud (Tabla 8.2), publicando un alto porcentaje de su producción en revistas Q1 y que sin embargo guarda muy poca similaridad con el resto de universidades (Figuras 8.1 y 8.5). Esta universidad sigue un perfil de publicación en revistas muy distinto al resto de universidades que sin embargo, se traduce en un buen resultado en relación a su presencia en los principales rankings internacionales, (Delgado López-Cózar, Jiménez-Contreras y Robinson-García, 2012) revelando una clara política de publicación.

En general, vemos cómo el análisis de universidades mediante el mapeo de acuerdo a su perfil de publicación en revistas científicas permite profundizar en las relaciones que se establecen entre las universidades y las motivaciones de dichas relaciones. Así pues, su aplicación de cara a la política científica puede ser muy útil al ofrecer no solo mayores claves interpretativas a la hora de analizar un sistema universitario, sino también al hacerlo a través de herramientas de fácil lectura como son los mapas de la ciencia. En este sentido, al identificar universidades similares en áreas concretas a través del análisis de universidades a partir de su perfil de publicación en revistas, una universidad puede identificar posibles socios por ejemplo para establecer proyectos de colaboración a través de convocatorias como la de los Campus de Excelencia (Docampo y otros, 2012) o para reforzar colaboraciones ya existentes con universidades similares que publican en revistas de alto impacto.

La Figura 8.6 muestra una forma fácil e intuitiva de identificar universidades similares y de alto impacto en un área determinada con respecto a la universidad que se analiza. Por ejemplo, se observa cómo la universidad de más impacto y más similaridad a Valencia en el área de Ciencias Sociales es Barcelona. También se observa, por ejemplo, cómo existen universidades que guardan gran semejanza con Complutense en el área de Ciencias de la Salud pero que no logran situar en ningún caso más del 50% de sus publicaciones en revistas Q1 siendo a priori, colaboradores potenciales menos interesantes que Barcelona o Autónoma de Madrid que sí superan dicho umbral. Así, queda claro que las relaciones de colaboración no pueden plantearse en ningún caso a nivel general, sino que deben ser temáticas. Aunque en este estudio mostramos un ejemplo referido a áreas, debido a su heterogeneidad, lo ideal que este tipo de análisis se aplicaran a disciplinas. En esta línea, en la web http://www.ugr.es/~elrobin/jpp.html el lector podrá encontrar más ejemplos para las cuatro áreas reseñadas en la figura 6. Como línea futura de investigación, sería de interés comparar estas redes de universidades con la oferta docente que ofrecen a fin de contrastar la coherencia entre la dimensión investigadora y la dimensión docente de las universidades.

Figura 8.6 Identificación de colaboradores potenciales a partir del %Q1 y el perfil de publicación en revistas científicas en cuatro áreas científicas



Finalmente, concluimos señalando las cuestiones más relevantes desveladas en el presente estudio y respondiendo a los objetivos planteados al inicio del mismo:

- Los mapas de universidades basados en el perfil de publicación de revistas son una herramienta simple y útil para ver cuándo las comparaciones entre universidades tienen sentido y cuándo no lo tienen, al mostrar la semejanza temática y de impacto de las mismas. Además, pueden resultar de gran utilidad de cara a la toma de decisiones en política científica al permitir fácilmente identificar colaboraciones potenciales más o menos interesantes para una institución.

- Hemos identificado cuatro tipos de universidades dentro del sistema español (Tabla 8.5). En primer lugar un grupo caracterizado por su gran tamaño y su carácter multidisciplinar, liderado por las universidades madrileñas y catalanas como Barcelona o Complutense y que por tanto, alcanzan un mayor grado de centralidad en todas las áreas. Un segundo grupo formado por universidades especializadas en un área determinada. Este grupo está principalmente caracterizado por las universidades politécnicas, muy similares entre sí y diferentes del resto de universidades en casi todas las áreas excepto en Ciencias Exactas. Sin embargo, también se incluiría Navarra, altamente especializada en Ciencias de la Salud o Carlos III que destacan en Ingeniería y Tecnología y Ciencias Sociales. El tercer grupo estaría formado únicamente por Pompeu Fabra. Al igual que en el caso anterior, se trata de una universidad altamente especializada, aunque en este caso su perfil de publicación difiere del resto de universidades. Finalmente, en el cuarto grupo estaría el resto de universidades de menor tamaño y menor presencia investigadora.

Tabla 8.5 Tipo de universidades en el sistema español de acuerdo a su perfil de publicación en revistas.

	Características	Ejemplo
Tipo I	Gran tamaño Carácter multidisciplinar Mayor centralidad en la red	Universidad de Barcelona Universidad Complutense Autónoma de Madrid ...
Tipo II	Alta especialización en un área determinada	Politécnica de Cataluña Navarra Carlos III de Madrid ...
Tipo III	Alta especialización en un área determinada Distinto perfil de publicación al resto de universidades	Pompeu Fabra
Tipo IV	Poca productividad	Burgos León Jaén ...

Referencias

Ahlgren, P. y Colliander, C. (2009). Document-document similarity approaches and science mapping: Experimental comparison of five approaches. *Journal of Informetrics, 3*(1), 49-63.

Beaver, D. (2001). Reflections on collaboration (and its study): past, present and future. *Scientometrics, 52*(3), 365-377.

Billaout, J.-C., Bouyssou, D. y Vincke, P. (2010). Should we believe the Shanghai Ranking? An MCDM view. *Scientometrics, 84*(1), 237-263.

Bordons, M., Sancho, R., Morillo, F. y Gómez, I. (2010). Perfil de actividad científica de las universidades españolas en cuatro áreas científicas: un enfoque multifactorial. *Revista Española de Documentación Científica, 33*(1), 9-33.

Carayol, N., Filliatreau, G. y Lahatte, A. (2012). Reference classes: a tool for benchmarking universities' research. *Scientometrics, 93*(2), 351-371.

Collini, S. (2011). *What are universities for?* London, UK: Penguin Books.

Delgado López-Cózar, E., Torres-Salinas, D., Jiménez-Contreras, E. y Ruiz-Pérez, R. (2006). Análisis bibliométrico y de redes sociales aplicado a las tesis bibliométricas defendidas en España (1976-2002): temas, escuelas científicas y redes académicas. *Revista Española de Documentación Científica, 29*(4), 493-524,

Delgado López-Cózar, E., Jiménez-Contreras, E. y Robinson-García, N. (2012). Meta Ranking EC3 de las universidades españolas. Disponible en: http://ec3.ugr.es/metaranking/metaranking.html [Consultado el 12 de noviembre de 2012].

Docampo, D., Herrera, F., Luque-Martínez, T. y Torres-Salinas, D. (2012). Efecto de la agregación de universidades españolas en el Ranking de Shanghai (ARWU): caso de las comunidades autónomas y los campus de excelencia. *El profesional de la información, 21*(4), 428-432.

Everitt, B., Landau, S. y Leese, M. (2001). *Cluster analysis.* (4ª ed.) Londres; Addison-Wesley.

García, J.A., Rodríguez-Sánchez, R., Fdez-Valdivia, J., Robinson-García, N. y Torres-Salinas, D. (2012). Mapping academic institutions according to their journal publication profile: Spanish universities as a case of study. *Journal of the American Society for Information Science and Technology, 63*(11), 2328-2340.

Glänzel, W. y Schubert, A. (2004). Analysing scientific networks through co-authorship. En: Moed, H.F., Glänzel, W. y Schmochl, U. (eds). *Handbook of quantitative science and technology research*. New York, EE.UU.: Kluwer Academic Publishers.

Hazelkorn, E. (2009). Rankings and the battle for world-class excellence: Institutional strategies and policy choices. *Higher Education Management and Policy, 21*(1), 55-76.

Hazelkorn, E. (2011). *Rankings and the reshaping of higher education: The battle for world-class excellence.* Basingstoke, UK, New York; Palgrave MacMillan.

Liu, N.C. y Chen, Y. (2005). The Academic Ranking of World Universities. *Higher Education in Europe, 30*(2), 127-136.

Macilwain, C. (2012). Excellence, Ja, Elitism, Non. *Science, 338*, 596-599.

Noyons, E. (2004). Science maps within a science policy context. En: H.F. Moed, W. Glänzel y U. Schmochl (eds). *Handbook of quantitative science and technology research*. New York, EE.UU.: Kluwer Academic Publishers.

Orduña-Malea E. (2012). *Propuesta de un modelo de análisis redinformétrico multinivel para el estudio sistémico de las universidades* [Tesis doctoral]. Valencia: Universidad Politécnica de Valencia.

Palomares-Montero, D., García-Aracil, A. y Castro-Martínez, E. (2008). Evaluación de las instituciones de educación superior: revisión bibliográfica de sistemas de indicadores. *Revista Española de Documentación Científica, 31*(2), 205-229.

Robinson-García, N., Rodriguez-Sánchez, R., García, J.A., Torres-Salinas, D. y Fdez-Valdivia, J. (2012). Material complementario del estudio 'Análisis de redes de las universidades españolas de acuerdo a su perfil de publicación en revistas por áreas científicas'. Disponible en: http://hdl.handle.net/10481/22367 [Consultado el 15 de noviembre de 2012]

Salton, G. y Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing and Management, 24*(5), 513-523.

Sanz-Menéndez, L. (2003). Análisis de redes sociales: o cómo representar las estructuras sociales subyacentes. *Documento de trabajo 03-07, Unidad de Políticas Comparadas*. Disponible en: http://digital.csic.es/bitstream/10261/1569/1/dt-0307.pdf [Consultado el 11 de marzo de 2013].

Schimank, U. y Winnes, M. (2000). Beyond Humboldt? The relationship between teaching and research in European university systems. *Science and Public Policy, 27*(6), 397-408.

Small, H. (1973). Cocitation in the scientific literature: a new measure of the relationship between two documents. *Journal of the American Society for Information Science, 24*(4), 265-269.

Torres-Salinas, D., Delgado López-Cózar, E., Moreno-Torres, J.G. y Herrera, F. (2011a). Rankings ISI de las universidades españolas según campos científicos: descripción y resultados. *El profesional de la información, 20*(1), 111-118.

Torres-Salinas, D., Moreno-Torres, J.G., Robinson-García, N., Delgado López-Cózar, E. y Herrera, F. (2011b). Rankings ISI de las universidades españolas según campos y disciplinas científicas (2ª ed. 2011). *El profesional de la información, 20*(6), 701-711.

Torres-Salinas, D., Moreno-Torres, J.G., Robinson-García, N., Delgado López-Cózar y Herrera, F. (2012). Rankings I-UGR de universidades españolas según campos y disciplinas científicas (3ª edición). Disponible en: http://rankinguniversidades.es/ [Consultado el 12 de noviembre de 2012].

Van der Wende, M. y Westerheijden, D. (2009). Rankings and classifications: The need for a multidimensional approach. En: F.A. van Vught (ed). *Mapping the higher education landscape: Towards a European classification for higher education.* Springer.

Van Raan, A.F.J. (2005). Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics, 62*(1), 133-143.

Visser, M.S., Calero Medina, C.M y Moed, H.F (2007). Beyond rankings: The role of large research universities in the global scientific communication system. *Proceedings 11th International Conference of the International Society for Scientometrics and Informetrics*. Madrid, España: CINDOC-CSIC, pp. 761-765.

Wagner, C.S. y Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research Policy*, *34*(10), 1608-1618

PART 3. Conclusions

9. Discussion and concluding remarks

In this chapter we discuss how the contributions presented respond to the research questions posed in this thesis, analyzing 1) the role of national university rankings by fields and their benefits and pitfalls as tools for decision-making in research policy, and 2) science mapping techniques that can serve as complementary tools to rankings, describing the disciplinary profile of universities. For this, we structure this chapter in three sections. The first two relate to the research questions this thesis aims to answer, while the last section discusses the implications of the work presented.

The studies presented in part go further into the use of bibliometric tools and rankings for research policy-making adopting a critical perspective that aims to acknowledge its advantages but also warn on its limitations. Rather than aiming at the ultimate solution for developing university rankings based on research performance that surpass their present shortcomings, we highlight some of the issues that should be considered and we present possible solutions to complement the information they provide. We adopt a critical perspective on university rankings and detail their limitations when it comes to assessing the research performance of universities: neglecting differences between disciplines.

9.1. The use of university rankings for research evaluation purposes

Worldwide, the use of bibliometric techniques and methodologies is nowadays widespread in the assessment and analysis of research activity. Research managers have relied heavily on them for the evaluation of the research performance of academic institutions. However, sloppy evaluation exercises as well as an overconfidence in their results by decision makers has led bibliometricians to adopt a more critical view of the consequences of their studies. Bibliometric techniques allow more transparent and accountable evaluation processes than traditional evaluation systems based on peer review. Also, they are much cheaper than the latter. In this sense, they seem to be a plausible option to substitute peer review (i.e., Watkins, 1998). University rankings and their impact in higher education are a good example of the expansion of bibliometric techniques in research evaluation. Bibliometricians have produced much literature warning against the shortcomings of their methodologies (i.e., van Raan, van Leeuwen & Visser, 2011) and even suggested good practices in order to make research policy-makers understand that their analyses can be very useful in facilitating decision-making but should not be decisive (i.e., van Leeuwen, 2007). In this section we will respond to the first research question posed in chapter 3.

Are national university rankings necessary in an international context?

This question aims to solve a research policy issue when using world-class university rankings to assess the performance of a country's university system which is not well-positioned. Chapter 4 addresses the need to offer rankings at the national level. The attraction rankings exert over research managers makes their presence unavoidable for bibliometricians, hence the need to provide meaningful rankings with transparency in the data retrieval process and methodology, allowing reproducibility and highlighting their strengths and also their weaknesses. Exercises such as that performed in chapter 4 or in Robinson-García, Jiménez-Contreras and Delgado López-Cózar (2013) show that combining and analyzing critically the information provided by the main international rankings and a national ranking are a good means of providing a relatively accurate, complete image of a national university system.

To do so, one must acknowledge the methodologies, data sources and variables used by each ranking when measuring the research performance of universities. Differences in this regard will be of key importance in understanding discrepancies between universities' positions. Also, the use of global positions may be misleading as there is a bias towards certain fields and languages (Bornmann, Moya-Anegón & Mutz, 2013; van Raan, van Leeuwen & Visser, 2011). In this sense, it is advisable to turn our focus towards university rankings by fields which offer a better account of the research perfomance of a university, comparing within the same fields and hence, leaving aside possible disciplinary differences between universities.

Still, this may also be problematic, as observed in chapter 5. Ranking consumers and research managers tend to ignore how these fields have been constructed and delimited, confounding research fields based on journal classification systems with research fields based on the organizational structure of their universities. Here we are not dealing with a methodological or transparency issue but with an interpretation problem. In the study presented we show the difficulties of constructing rankings based on organizational units and their possible irrelevance especially when dealing with relatively new institutions which do not follow a traditional structure due to the introduction of New Public Management. Another important problem is the impossibility of using a *bottom-up* approach in the data retrieval process which would allow ranking producers to construct from basic organizational units to entire academic institutions by aggregating the output, as suggested elsewhere (Bornmann, Mutz & Daniel, 2013). Although this is what research managers are really looking for when analyzing university rankings by fields, due to data availability issues this implementation is not feasible

Indeed, the unavoidable difficulties one would have to face to establish a link between research papers and individuals make it impossible to relate input data with output data. An alternative solution would be to use visualization techniques for the analysis of multivariate data which would allow a better insight into how these relate. The study included in chapter 6 presents the biplot technique and offers three different case studies in which it could be a solution for research policy-makers. In the first case we analyze R&D indicators (input) with bibliometric indicators (output), comparing and analyzing the research activity of 21 European countries. However, its use as a complement to rankings is better seen in the second case study. Here, we analyzed four variables of the top 25 universities according to the THE World University Rankings. Unrelated variables such as international outlook or citations are displayed, showing different patterns which influence the position of universities and illustrate different university profiles (research-oriented, teaching-oriented, etc.)

9.2. Issues with the analysis of the disciplinary profile of universities

A further issue of concern when dealing with university rankings by fields is that they 'reassert the hierarchy of traditional knowledge' (Hazelkorn, 2011), extending the evaluation, publication and citation practices of the Basic and Biomedical Sciences to other areas in which their utility is limited (Hammarfelt, 2012). Chapters 7 and 8 deal specifically with this topic and focus on the use of science mapping techniques for identifying similarities and differences between universities that go beyond bibliometric classification systems. These two studies relate to the second research question posed in chapter 3.

How can we develop tools for research policy-makers that can allow them to identify disciplinary profiles?

In chapter 6 we already started to explore visualization techniques for analyzing disciplinary differences within a single university. However, it is with the introduction of the *journal publication profile* methodology in chapter 7 when we start to explore a means to classify Spanish universities according to their disciplinary focus. This classification is further developed in chapter 8.

Disciplinary differences occur at different levels and may be influenced by external factors such as geographical proximity, strong collaboration ties or type of institution. Mapping techniques allow us to avoid Duhem's Law of Cognitive Complementarity, which states that certain levels of vagueness are needed in order to secure reliability in any statement given when defining a specific phenomenon (Rescher, 2006). At the same time they do not necessarily limit their function at defining a particular characteristic, but go beyond showing other types of relations. The *journal publication profile* methodology considers common journals in which two institutions have published as the link that relates them. Using journals as links broadens the relation between institutions as it can be derived from different aspects of research such as collaboration or topic relation. Also, it is not dependent on any classification system. On the other hand, it is advisable to employ this methodology at a meso-level, that is, for the analysis of specific fields or areas, which means that it does not completely avoid the use of classification schemes.

From the analyses conducted using this methodology we have detected five different aspects that characterize our classification:

- **Geographical and historical ties.** Universities with a common origin and geographically near to each other are characterized by having strong collaboration links in certain fields of endeavor. An example of these ties are seen in chapter 7, where we refer to similarities between the University of Granada and the University of Jaén in the field of Computer Science (figure 7.3).

- **Disciplinary relations.** Universities specialized in common fields are easily spotted when visualizing their relations through the *journal publication* *profile* methodology. This is observed in chapter 8 when analyzing the area of the Social Sciences (figure 8.1). Here, a cluster of three universities especialized in the field of Economics (Alicante, Carlos III and Pompeu Fabra) form a distinctive separate component of the network.

- **Research orientation.** One of the main shortcomings of university rankings by fields is their inability to distinguish applied from basic sciences. This limitation is overcome with the suggested methodology as observed in Figure 7.1 (all areas) and Figure 8.3 (Engineering and Technology) where polytechnic universities tend to separate from more traditional and historical universities with a less applied nature.

- **Publication patterns.** Outliers of the system are quickly identified, allowing policy-makers to rapidly draw their attention towards these special cases and analyze them separately. This is the case of the Pompeu Fabra University which seems to show a completely different pattern to the rest of the universities in the Spanish system. Indeed, these differences have led us to consider it a separate class of university when classifying them (Table 8.5).

- **Size dependence.** There seems to be a direct relation between size and specialization. Bigger universities seem to have a more generalist profile as observed from our analyses. This may be a way for smaller universities to make a difference by concentrating their research efforts in specific scientific areas.

9.3 Implications and further research

University rankings are powerful tools which have proven to be of high regard by research managers. However, due to the many shortcomings they suffer and the many critiques they have received, ranking producers are constantly improving them and converting them into something more than rankings. Trends such as the release of university rankings by fields, league tables for specific regions or countries, turning away from global indicators, or the offer of more transparent products are just a few examples of these efforts.

However, there are still several issues which still need to be analyzed in order to provide meaningful metrics to university managers. For instance, the dependence on a classification scheme which will not acknowledge interdisciplinarity still remains a serious issue which requires further considerations. Although bibliometric classification systems are necessary not only to establish boundaries between fields but to ease interpretation, they are artificial artifacts that may not adapt adequately when performing macro analyses such as the ones rankings introduce by considering a large and heterogeneous variety of academic institutions. Also, rankings are very popular due to their simplicity. The combination of different analyses of other facets of universities' activity (such as research, education and knowledge transfer) offering a synthesized picture of the global higher education landscape while at the same time preserving some accuracy and easing interpretation remains a major concern.

Here we focus on the research dimension and on a specific aspect of it: the disciplinary focus of universities. The analysis of disciplinary differences at the institutional level is far too complex for us to be able to reduce it to league tables. Science mapping techniques are a good way to complement the information provided by rankings as they are easy-to-read tools that represent with relative accuracy the relation between universities and hence their similarities and dissimilarities. Also, they provide far richer information regarding the performance of academic institutions. In other words, one could argue that while league tables show vertical differences, science maps show horizontal stratification, allowing a better analysis not only of single institutions but of entire national higher education systems.

References

Bornmann, L., Moya-Anegón, F., Mutz, R. (2013). Do universities or research institutions with a specific subject profile have an advantage or disadvantage in institutional rankings? *Journal of the American Society for Information Science and Technology, 64*(11), 2310-2316.

Bornmann, L., Mutz, R., & Daniel, H.-D. (2013). Multilevel-statistical reformulation of citation-based university rankings: The Leiden Ranking 2011/2012. *Journal of the American Society for Information Science and Technology, 64*(8), 1649-1658.

Hammarfelt, B. (2012). *Following the footnotes: A bibliometric analysis of citation patterns in literary studies* [Doctoral thesis]. Uppsala: Uppsala University.

Hazelkorn, E. (2011). *Rankings and the reshaping of higher education: The battle for World-Class excellence.* Palgrave MacMillan.

Rescher, N. (2006). *Epistemetrics*. Cambridge: Cambridge University Press.

Robinson-García, N., Jiménez-Contreras, E., & Delgado López-Cózar, E. (2013). Las universidades españolas en los principales rankings de universidades. 2013. *EC3 Working Papers, 14*, November 20. Accessible at: http://digibug.ugr.es/bitstream/10481/29083/7/InformeMetaRankingv104-1.pdf

Van Leeuwen, T.N. (2007). Modelling of bibliometric approaches and importance of output verification in research performance assessment. *Research Evaluation, 16*(2), 93-105

Van Raan, T., van Leeuwen, T., & Visser, M. (2011). Non-English papers decrease rankings. *Nature, 469*(7328), 34.

Watkins, P.R.T. (1998). Institutional research rankings via bibliometric analysis and direct peer review: A comparative case stud with policy implications. *Scientometrics, 41*(3), 335-355.

Appendix Impact report and Curriculum Vitae

This thesis is presented following the *selection of publications* modality. In this appendix we provide proof of the approppriateness of the contributions included in terms of visibility of the journal in which the study was published and scientific impact (citations) received. I also give account of the scientific output produced since the beginning of my PhD training in October, 2010.

Visibility and impact of the contributions

This thesis includes five journal articles which analyze different aspects of the thesis related with university rankings. These have been published in international journals all of them included in the *Web of Science Social Science Citation Index*.

In table 1, I include a bibliometric overview of the impact (citations received) of the publications and visibility of the journal in which these were published (Journal Impact Factor and ranking position in the *Journal Citation Reports*). All of the papers have been published in journals belonging to the *Information Science & Library Science* subject category. Two of the publications (chapters 4 and 5) were published in *Scientometrics*, the main international journal in the field of bibliometrics. Chapters 6 and 7 were published in the *Journal of the American Society for Information Science and Technology* (JASIST), which is the main international journal for the field of Library Sciences. The last contribution (chapter 8) was published in *Revista Española de Documentación Científica* (REDC), which is one of the most important journals in the field of Library and Information Science in Spain and the Spanish journal with the highest Journal Impact Factor in the category. Both, JASIST and Scientometrics, belong to the top10 journals in their category with a Journal Impact Factor of 2.005 and 2.133 respectively, which highlights their great visibility in the field. As REDC is a national journal, it is positioned in the third quartile of the category.

Regarding the citations received so far, all of the contributions are relatively recent, hence little can be reported. However, the contribution first published (Chapter 7) has already received 4 citations according the Web of Science. Chapters 6 and 8 are already started to be cited as observed according to Scopus and Google Scholar, while no citations were reported for chapters 4 and 5 which are the most recent publications.

Table 1. General description of published contributions

	Journal	Year	Citations			JIF	Rank
			ISI	Scopus	Gscholar		
Chapter 4	Scientometrics	Online	0	0	0	2.133	7/85
Chapter 5	Scientometrics	2014	0	0	0	2.133	7/85
Chapter 6	J Am Soc Inf Sci Tec	2013	0	1	1	2.005	10/85
Chapter 7	J Am Soc Inf Sci Tec	2012	4	5	7	2.005	10/85
Chapter 8	Rev Esp Doc Cient	2013	0	0	1	0.453	51/85

Other PhD-related contributions

As mentioned in section 1.2 of chapter 1, the present thesis is motivated because of my involvement in the team that produces the I-UGR Rankings. The first edition of these rankings was released in 2010, however I started to work in the project in its 2nd edition. During the ellaboration of the rankings the team involved has been especially keen on confronting our progresses to the research community. Followingly I include a list of the publications I have co-authored derived from the I-UGR Rankings (chapter 4 is not included as it has been mentioned before):

Robinson-García, N., Moreno-Torres, J.G., Torres Salinas, D., Delgado-López Cózar, E., & Herrera, F. (2013). The role of national university rankings in an international context: the case of the I-UGR Rankings of Spanish universities. *14th International Society of Scientometrics and Informetrics Conference, 2*, 1550-1565.

Torres-Salinas, D., García-Moreno-Torres, J., Robinson-Garcia, N., Delgado-López-Cózar, E., & Herrera, F. (2011). Rankings ISI de las Universidades Españolas según campos y disciplinas científicas (2ª ed. 2011). *El Profesional de la Información, 20*(6), 701-709.

Torres-Salinas, D., Delgado López-Cózar, E., Robinson-Garcia, N., Triguero, I., Herrera, F. (2013). *Posiciones de las universidades españolas y de las comunidades autónomas en los Rankings I-UGR según campos y disciplinas científicas* [Report]. May 23, 2013. Accessible at: http://sci2s.ugr.es/rankinguniversidades/downloads/rankingsI-UGR_Posiciones%202013.pdf

Torres-Salinas, D., Delgado López-Cózar, E., Robinson-Garcia, N., Moreno-Torres, J.G., & Herrera, F. (2011). *Rankings ISI: la universidad española en la Web of Science.*

2001-2010. Accessible at: http://ec3.ugr.es/metaranking/images/rankingsISI_2011_InformeResumen.pdf

Also, in 2011 we started the *MetaRanking EC3 de las Universidades Españolas* project along with my supervisors which intended to analyze the position of Spanish universities in the main international rankings as well as the I-UGR Rankings in order to provide small profiles of the performance of each university. The latest version released in 2013 is available at http://wdb.ugr.es/~elrobin and associated to it is the following report:

Robinson-Garcia, N., Jiménez-Contreras, E., & Delgado López-Cózar, E. (2013). Las universidades españolas en los principales rankings de universidades. 2013. *EC3 Working Papers 14*, 20 de noviembre de 2013. Accessible at: http://digibug.ugr.es/bitstream/10481/29083/7/InformeMetaRankingv104-1.pdf

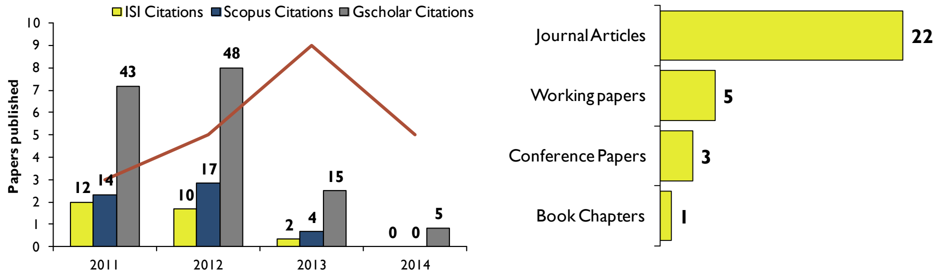
Research output for the 2010-2014 time period

During my PhD training I have collaborated in a total of 31 papers with 23 different researchers from three different countries (Spain, The Netherlands and Austria). The three resarchers with whom I have worked together the most are Daniel Torres-Salinas (24), Emilio Delgado López-Cózar (16) and Evaristo Jiménez-Contreras (8). As well as my PhD topic I have researched in other research topics such as the analysis of the Book Citation Index, Data Sharing, Google Scholar and domain-specific bibliometric analyses. An overview of my performance in terms of scientific impact, visibility and production is shown in figure 1. Of the 31 papers 22 were journal articles and 18 are included in the Web of Science database. From the PhD time period, 2013 was my most productive year with 9 journal papers published.

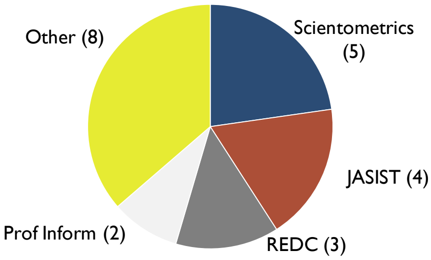
I have published in 12 different journals which include among others, Science, Journal of Informetrics, JASIST, Scientometrics or PLOS One. However, the main journals in which I have published are Scientometrics, JASIST and REDC. For further information regarding my research output the reader is referred to my personal website available at http://www.ugr.es/~elrobin.

Figure 1. Overview of research production, citations received and journal visibility for the 2011-2014 time period

PUBLICATIONS AND CITATIONS



JOURNALS AND IMPACT FACTOR



Rank	Journal	JIF 2012
1	Science	31.027
2	J Informetrics	4.153
3	PLOS One	3.733
4	Scientometrics	2.133
5	JASIST	2.005

Apéndice Discusión y conclusiones

En este capítulo comentamos las contribuciones presentadas y que responden a las preguntas de investigación planteadas en la tesis. Estas analizan 1) el rol de los rankings nacionales de universidades por campos y sus ventajas e inconvenientes como herramienta para la toma de decisiones en política científica y 2) las técnicas de mapeo de la ciencia como herramientas complementarias a los rankings que describen el perfil disiciplinar de las universidades. Para ello estructuramos el capítulo en tres secciones. Las dos primeras responden a las preguntas de investigación planteadas, mientras que la última sección discute las implicaciones que tiene el trabajo presentado.

Los trabajos presentados en la segunda parte de esta tesis profundizan en el uso de herramientas bibliométricas y rankings para la toma de decisiones en politica científica, adoptando una perspectiva crítica que pretende enfatizar tanto sus beneficios como sus limitaciones. Más que pretender plantear la solución definitiva para el desarrollo de rankings basados en indicadores bibliométricos que supere sus limitaciones presentes, analizamos cuestiones que deben ser consideradas y presentamos posibles soluciones que complementen la información que ofrecen los rankings. Adoptamos una perspectiva crítica de los rankings de universidades y detallamos sus limitaciones de cara a su empleo para evaluar el rendimiento científico de las universidades: no tienen en cuenta las diferencias entre disciplinas.

El uso de los rankings de universidades para hacer política científica

El empleo de técnicas y metodologías bibliométricas de cara a la evaluación y el análisis de la actividad científica es común en todo el mundo. Los gestores de investigación confían en ellas cuando evalúan el rendimiento científico de las universidades. Sin embargo, la implementación de evaluaciones poco rigurosas unido a una confianza extrema en sus resultados obliga a los bibliómetras a ser más críticos y alertar sobre las consecuencias negativas de dichas evaluaciones. Las técnicas bibliométricas permiten implementar procesos de evaluación más transparentes que aquellos basados en la revisión por pares. Además, son menos costosos. Esto hacen que sean consideradas como una opción bastante aceptable para sustituir la revisión por pares (i.e., Watkins, 1998). El impacto que tienen los rankings de universidades supone un buen ejemplo de la expansión de estas técnicas en la evaluación científica. Los bibliómetras han publicado numerosos estudios en los que alertan sobre las limitaciones de sus metodologías (i.e., van Raan, van Leeuwen & Visser, 2011) sugiriendo incluso directrices de buenas prácticas con el fin de concienciar a los gestores científicos de la utilidad que tienen para asesorarles pero no para tomar decisiones de manera automática (i.e., van Leeuwen, 2007). En esta sección responderemos a la primera pregunta de investigación que se expuso en el capítulo 3.

¿Son necesarios los rankings nacionales de universidades en un contexto internacional?

Esta pregunta trata una cuestión que se plantea desde la perspectiva de la politica científica al emplear rankings internacionales para evaluar sistemas universitarios nacionales que no están bien posicionados en los mismos. El capítulo 4 trata sobre la necesidad de emplear rankings a nivel nacional. La atracción que los rankings ejercen sobre los gestores científicos hace que su existencia sea inevitable. Ello hace necesario desarrollar rankings sólidos que sean transparentes tanto en el proceso de recogida de datos como en la metodología, permitiendo la reproducibilidad de los mismos y señalando sus fortalezas y debilidades. Ejercicios como el que se presenta en el capítulo 4 o en Robinson-García, Jiménez-Contreras and Delgado López-Cózar (2013) muestran la importancia que tiene combinar y analizar de manera crítica la información que ofrecen tanto los principales rankings internacionales como los nacionales para así obtener una imagen completa y precisa de un sistema nacional universitario.

Para ello es necesario conocer las metodologías, fuentes de datos y variables que emplea cada ranking. Resulta vital identificar las diferencias que hay en este sentido para así entender las posibles discrepancias que haya entre las posiciones de las universidades en los distintos rankings. Además, el uso de rankings globales puede resultar engañoso, ya que presentan sesgos lingüísticos y también hacia ciertos campos científicos (Bornmann, Moya-Anegón & Mutz, 2013; van Raan, van Leeuwen & Visser, 2011). En este sentido, recomendamos el empleo de rankings por campos científicos para así reflejar mejor el rendimiento científico de las universidades y permitir las comparaciones dentro de un mismo campo para así evitar sesgos disciplinares.

Aún así, esto tambiénresulta problemático, tal y como se indica en el capítulo 5. Los usuarios de los rankings y los gestores científicos suelen ignorar cómo se construyen y diseñan dichos campos científicos. Así pues, no son conscientes de que suelen diseñarse a partir de sistemas de clasificacion de revistas y no no representan necesariamente la estructura organizacional de las universidades. Aquí el problema no es metodológico ni de transparencia, sino que se trata de un problema de interpretación y lectura. En este estudio mostramos las dificultades técnicas que se plantean al construir un ranking basado en la estructura organizacional de las universidades y su irrelevancia cuando tratamos con instituciones relativamente nuevas que no siguen una estructura organizacional tradicional. Otro problema destacable es la imposibilidad de emplear una aproximación *bottom-up* en la recogida de datos, lo que permitiría a los autores de los rankings reconstruir la producción global de las universidades mediante la agregación de las unidades básicas organizacionales que conforman una universidad, tal y como se sugieren algunos estudios (Bornmann, Mutz & Daniel, 2013). Aunque lo que esperan ver los gestores universitarios en los rankings por campos es precisamente una estructura similar a la de su institución, esto no es posible.

Las dificultades técnicas que se plantean al intentar asignar cada trabajo científico y cada investigador una una institución específica hace que sea imposible relacionar los datos de entrada (los investigadores o departamentos que producen los artículos) con los datos de salida (las publicaciones científicas). Una solución alternativa sería el empleo de técnicas de visualización de datos multivariantes que nos permitan entrever cómo están relacionados. El estudio presentado en el capítulo 6, muestra la técnica biplot y analiza mediante tres casos de estudios, cómo podría emplearse para asesorar a políticos científicos y gestores. En el primer caso analizamos indicadores de ciencia y tecnología (datos de entrada) con indicadores bibliométricos (datos de salida), comparando y analizando la actividad científica de 21 países europeos. Aunque su uso como complemento de rankings de universidades está más claro en el segundo caso de estudio. Aquí analizamos las primeras 25 universidades en los THE World University Rankings de acuerdo a cuatro de las variables que presentan. Se identifican variables independientes como la perspectiva internacional o el número de citas recibidas, mostrando distintos aspectos que influyen en el posicionamiento de las universidades e ilustran distintos perfiles de universidad (de investigación, docentes, etc.)

Cuestiones relacionadas con el análisis del perfil disciplinar de las universidades

Otra limitación que plantean los rankings de universidades por campos es que 'reafirman la jerarquía del conocimiento tradicional' (Hazelkorn, 2011), extendiendo las prácticas de evaluación, publicación y citación de las Ciencias Básicas y Biomédicas a otras áreas donde su utilidad es limitada (Hammarfelt, 2012). Los capítulos 7 y 8 tratan específicamente este tema y se centran en el uso de técnicas de mapeo de la ciencia para identificar similaridades y diferencias entre universidades que van más allá de los sistemas de clasificación bibliométricos. Estos dos estudios tienen que ver con la segunda pregunta de investigación enunciada en el capítulo 3.

¿Cómo podemos desarrollar herramientas que permitan a los gestores de política científica identificar distintos perfiles disciplinares de universidad?

En el capítulo 6 ya empezamos a explorar técnicas que analizan diferencias disciplinares a nivel institucional. Pero es con la introducción de la metodología del *perfil de publicación en revistas* en el capítulo 7, cuando comenzamos a explorar la posibilidad de clasficar las universidaades españolas de acuerdo a su perfil disciplinar. Dicha clasificación se presenta en el capítulo 8.

Encontramos diferencias disciplinares a distintos niveles y pueden deberse a factores externos como son la proximidad geográfica, una colaboración estrecha o el tipo de universidad. Las técnicas de mapeo de la ciencia nos permiten evitar la Ley de Duhem de Complementariedad Cognitiva, que afirma que son necesarios ciertos niveles de ambigüedad si quieremos hacer afirmaciones con cierto nivel de seguridad sobre un fenómeno específico (Rescher, 2006). Al mismo tiempo, tampoco se limitan a presentar una característica específica, pero que pueden mostrar diversos tipos de relaciones. La metodología de *perfil de publicación en revistas* considera a las revistas comunes en las que dos instituciones publican como un buen enlace para establecer una relación entre ellas. Al emplear la revista como enlace, la relación entre las instituciones queda definida de manera muy amplia, abarcando distintos aspectos como son relaciones colaboración o relaciones temáticas. Asimismo, no depende de un sistema de clasificación, aunque es recomendable emplear esta metodología a nivel meso-, es decir, para el análisis de grandes áreas científicas, por lo que no prescinde del todo del empleo de una clasificación.

A raíz de los análisis presentados empleando esta metodología, hemos identificado cinco tipos de relaciones que caracterizan la clasificación de universidades que presentamos:

- **Relaciones geográficas o históricas.** Las universidades con un origen común o que están geográficamente próximas suelen caracterizarse por tener fuertes lazos de colaboración en determinados campos científicos. Un ejemplo de ello se observa en el capitulo 7, donde nos referimos a las similaridades existentes entre la Universidad de Granada y la Universidad de Jaén en el campo de Ciencias de la Computación (figure 7.3).

- **Relaciones disciplinares.** La metodología propuesta hace que resulte fácil identificar universidades especializadas en un mismo campo. Esto se observa en el capítulo 8 al analizar el área de las Ciencias Sociales (figura 8.1). Aquí, observamos un grupo de tres universidades especializadas en el campo de la Economía (Alicante, Carlos III y Pompeu Fabra), formando un componente separado del resto de la red.

- **Orientación científica.** Una de las principales limitaciones de los rankings de universidades por campos es su incapacidad para distinguir entre las ciencias aplicadas y las ciencias básicas. Nuestra metodología supera esta limitación tal y como se observa en la Figura 7.1 (todas las áreas) y la Figura 8.3 (Ingeniería y Tecnología) donde las universidades politécnicas están separades de las universidades históricas y tradicionales donde se practica una ciencia más básica.

- **Patrones de publicación.** Identificamos de manera rápida a las universidades que presentan un perfil atípico, lo que hace que los gestores de política científica puedan rápidamente analizar estos casos especiales de manera independiente. Este es el caso de la Universidad Pompeu Fabra que parece seguir un comportamiento completamente distinto del resto de universidades en el sistema español. De hecho, este comportamiento atípico nos ha hecho considerarla como una clase de universidad distinta al resto (Tabla 8.5).

- **Dependencia del tamaño.** Parece haber una relación directa entre el tamaño y la especialización. Las universidades más grandes tienden a tener un perfil generalista tal y como revelan nuestros estudios. Tal vez las universidades de menor tamaño obten por concentrar su investigación en determinadas áreas científicas con el fin de destacar del resto de universidades.

Implicaciones y líneas de investigación futuras

Los rankings de universidades son herramientas consideradas en alta estima por parte de los gestores de investigación. Sin embargo, debido a las limitaciones que presentan y las críticas que han recibido, los autores de los rankings están intentando mejorarlos constantemente y convertirlos en algo más que rankings. Tendencias como la publicación de rankings por campos, rankings regionales o nacionales, o la elaboración de productos más transparentes ejemplifican dichos esfuerzos.

Sin embargo, siguen habiendo cuestiones que deben analizarse si queremos ofrecer a los gestores de investigación medidas e indicadores verdaderamente significativos. Por ejemplo, es necesario considerar las limitaciones que plantea la dependencia en un sistema de clasificación que no refleja la investigación interdisciplinar. Aunque es necesario emplear sistemas de clasificación bibliométricos para establecer límites entre los campos científicos y facilitar la interpretación, son constructos que pueden no adaptarse de manera adecuada al realizar análisis a nivel macro tales como los que realizan los rankings, ya que incluyen a variedad muy heterogénea de instituciones. Los rankings son my populares debido a su simplicidad. El hecho de combinar variables que analizan distintos aspectos de la actividad universitaria (tales como la investigación, la educación y la transferencia de conocimiento) para ofrecer una imagen sintética del panorama universitario plantea serios problemas ya que hacen que se muestren resultados poco precisos.

En esta tesis nos centramos en la investigción de las universidades y en un aspecto específico de la misma: el perfil disciplinar de las universidades. No resulta factible reducir a un conjunto de tablas las diferencias disciplinares que hay a nivel institucional. Las técnicas de mapedo de la ciencia plantean una aproximación alternativa para complementar a los rankings. Además, son herramientas fáciles de leer que representan con relativa precisión las relaciones entre universidades. Además, ofrecen una información más rica sobre el rendimiento de las universidades. En otras palabras, se podría decir que mientras que los rankings muestran diferencias verticales, los mapas científicos muestran la estratificación horizontal, permitiendo un análisis más completno no solo a nivel institucional, sino también a nivel de sistema universitario.

Referencias

Bornmann, L., Moya-Anegón, F., Mutz, R. (2013). Do universities or research institutions with a specific subject profile have an advantage or disadvantage in institutional rankings? *Journal of the American Society for Information Science and Technology, 64*(11), 2310-2316.

Bornmann, L., Mutz, R., & Daniel, H.-D. (2013). Multilevel-statistical reformulation of citation-based university rankings: The Leiden Ranking 2011/2012. *Journal of the American Society for Information Science and Technology, 64*(8), 1649-1658.

Hammarfelt, B. (2012). *Following the footnotes: A bibliometric analysis of citation patterns in literary studies* [Doctoral thesis]. Uppsala: Uppsala University.

Hazelkorn, E. (2011). *Rankings and the reshaping of higher education: The battle for World-Class excellence.* Palgrave MacMillan.

Rescher, N. (2006). *Epistemetrics*. Cambridge: Cambridge University Press.

Robinson-García, N., Jiménez-Contreras, E., & Delgado López-Cózar, E. (2013). Las universidades españolas en los principales rankings de universidades. 2013. *EC3 Working Papers, 14*, November 20. Accessible at: http://digibug.ugr.es/bitstream/10481/29083/7/InformeMetaRankingv104-1.pdf

Van Leeuwen, T.N. (2007). Modelling of bibliometric approaches and importance of output verification in research performance assessment. *Research Evaluation, 16*(2), 93-105

Van Raan, T., van Leeuwen, T., & Visser, M. (2011). Non-English papers decrease rankings. *Nature, 469*(7328), 34.

Watkins, P.R.T. (1998). Institutional research rankings via bibliometric analysis and direct peer review: A comparative case stud with policy implications. *Scientometrics, 41*(3), 335-355.