

# TRANSFORMACIÓN DE LA INFORMACIÓN EN CONOCIMIENTO Y DEL CONOCIMIENTO EN DECISIONES ESTRATÉGICAS

Xavier Polanco

Unité de Recherche et Innovation  
Institut de l'Information Scientifique et Technique (INIST)  
Centre National de la Recherche Scientifique (CNRS)  
2, allée du Parc de Brabois - 54514 Vandoeuvre-lès-Nancy Cedex - France  
polanco@inist.fr - <http://www.inist.fr/pri/uri>

## 1 INTRODUCCIÓN

Es evidente que hoy en día nos encontramos en un nuevo contexto desde donde considerar la bibliometría y la cienciometría. Este nuevo contexto se lo conoce bajo la denominación de "sociedad de la información", y se caracteriza por la irrupción de nuevas tecnologías de la información. Estas nuevas tecnologías de la información se orientan, en nuestro campo, a una mejor explotación de los datos. Nos referimos aquí a las tecnologías que permiten la realización de tareas conocidas como "vigilancia tecnológica", "inteligencia económica", "data mining", "text-mining", "knowledge discovery in databases", "knowledge management", y en las cuales la información almacenada no es más un producto final, sino que representa una materia prima que es necesario someter a un proceso de transformación, con el objetivo de extraer los conocimientos que pueden contribuir a mejorar la comprensión de una situación, y la toma de decisiones estratégicas en un determinado campo de actividades.

Es probable que estas nuevas tecnologías de la información signifiquen un "problemshift", es decir, un cambio de problemática en lo que concierne los estudios cuantitativos de la ciencia y de la tecnología. Si este cambio es efectivo, tanto mejor, puesto que nos reconocemos con el ya relativamente antiguo pero hasta ahora no realizado proyecto de una "cognitive scientometrics" (Rip, Courtial, 1984), o "qualitative scientometrics" (Callon, Law, Rip, 1986). Sin embargo, la posición y el punto de vista que adoptamos al respecto son muy diferentes con los representados por los autores citados. Consideramos este proyecto desde la ciencia de la computación o informática, y desde el punto de vista de la inteligencia artificial. Y además, en el contexto al cual hemos hecho referencia en el párrafo anterior.

## 2 UN CAMBIO DE PARADIGMA Y REALIDAD

Un estudio bibliométrico realizado sobre 184 referencias bibliográficas (Duff, 1995), cubriendo el período 1984-1993, y extraídas de 84 revistas científicas, a partir de la interrogación de 4 bases de datos (Information Science Abstracts, SocialSearch, INSPEC, Arts & Humanities Search), muestra que una mayoría de investigadores, en ingeniería, ciencia de la información y ciencias sociales, considera el concepto de "sociedad de la información" como un nuevo paradigma. Observando una baja de frecuencia en la utilización del término en el curso del período estudiado, el autor concluye que, desde un punto de vista bibliométrico, no es posible decir con certitud que la "sociedad de la información" haya alcanzado el estado de paradigma en el sentido de Kuhn (1970); sin embargo, su opinión es que ella representa una manera correcta de comprender el mundo actual.

En realidad, hemos pasado el cabo de la "sociedad de la información". Internet marcaría este pasaje. La historia comienza en 1967 cuando un "packet switching network plan" fue presentado al simposio sobre "Operating System Principle" de la ACM (Association for Computing Machinery), y dos años más tarde el Departamento de la Defensa de los Estados Unidos encarga a la agencia ARPA (Advanced Research Project Agency) el desarrollo de una red teniendo por objetivo conectar los laboratorios de investigación militar. El "arpanet" dió nacimiento a "internet" en los años 1980 siendo utilizado en el campo de la investigación civil en los Estados Unidos. Al comienzo de los años 1990 se produce su difusión internacional y deviene un instrumento accesible al público gracias al World Wide Web (WWW).

Señalemos además el hecho que « *Développer une société de l'information conviviale* » constituye uno de los sub-programas del Quinto Programa Marco de la Unión Europea para las acciones de investigación, de desarrollo tecnológico y demostraciones (1998-2000).

En este contexto, toda organización tiene que aprender a manejar los flujos de información antes de cualquiera toma de decisiones estratégicas. De aquí entonces la aparición de estas nuevas prácticas profesionales conocidas

bajo los nombres de "vigilancia tecnológica", o "científica" o "comercial", y de "inteligencia económica" o "estratégica". Son prácticas ligadas al hecho de que la información "apropiada y fiable" representa en la vida de las organizaciones un valor estratégico primordial.

Las numerosas publicaciones acerca de la "vigilancia tecnológica" y de la "inteligencia económica" se dirigen principalmente al mundo de las empresas. Pero estas nociones pueden ser adaptadas al mundo de la investigación científica para desarrollar otra manera de coleccionar, procesar, distribuir y explotar la información científica y tecnológica. Si la información coleccionada y procesada es únicamente almacenada, no podemos hablar de vigilancia o de inteligencia, puesto que éstas suponen un rol activo en el sentido que la información es coleccionada, procesada, difundida y utilizada en tanto que medio al servicio de la toma de decisiones estratégicas, y para reforzar la acción y la capacidad competitiva de una organización y de sus miembros.

La actitud de "vigilancia/inteligencia" supone entonces la adopción de una posición activa con relación a la información. Se trata de un cambio en la concepción de la información: el pasaje de la "información-adquisición" a aquélla de la "información-acción".

### **3 RECONSIDERACIÓN CRÍTICA DE LA CIENCIOMETRÍA**

La cienciometría puede ser considerada como todo lo que es publicado en la revista de este nombre, es decir, *Scientometrics*. Después de su creación en 1978 esta revista declara que:

"the topics covered are results of research concerned with the quantitative features and characteristics of science. Emphasis is made on investigations in which the development and the mechanism of science are studied by means of (statistical) mathematical methods."

Y como hemos dicho más arriba, la premisa clave desde la cual entendemos la cienciometría es justamente el proyecto de una cienciometría "cognitiva" o "cualitativa". Este proyecto nos parece situarse más allá de los límites de la empresa cienciométrica de acuerdo con la definición de lo que se publica en la revista *Scientometrics*. La ciencia como conocimiento parece estar mucho más cerca del campo conocido como "knowledge discovery in databases" (al respecto véase Fayyad et al., 1996).

Considerar la ciencia como conocimiento representa un cambio con relación a la "research concerned with quantitative features and characteristics of science", en dirección del problema de cómo resolver "la extracción y la gestión de conocimientos". Y es en esta dirección que nosotros nos situamos y estamos trabajando en la Unidad de Investigación e Innovación del INIST-CNRS. Lo cual supone tener presente la distinción entre "conocimiento" y "documento" (ver Brookes, 1980-81) y de acuerdo con esta observación hecha en el campo de los estudios cienciométricos:

« Growth of knowledge must be distinguished from growth of the literature or growth in number of publications. The former is a more abstract concept and hence not so directly assessed. In bibliometrics, growth in number of publications is sometimes taken as a measure or operational definition of growth of knowledge » (Tague et al., 1981, p. 126).

Un aspecto es la gestión de documentos y publicaciones (documentación y biblioteconomía), otra cuestión es la gestión de conocimientos. Una cosa es el análisis estadístico teniendo como unidad de contabilidad las publicaciones y citas, y otra es plantearse el problema de la adquisición, representación y gestión de los conocimientos difundidos bajo una forma escrita a través de las publicaciones científicas y de las patentes.

### **4 UNA NUEVA PROBLEMÁTICA**

La ayuda a la decisión en la organización y la gestión de la investigación en política científica, la vigilancia tecnológica y la inteligencia económica en los sectores industriales, definen el contexto social en el cual se sitúa nuestro trabajo sobre el análisis de la información científica y tecnológica.

El análisis busca identificar la información "útil", aquélla que comporta un interés para el usuario a partir de una cantidad importante de información disponible. El análisis de la información en el contexto de la "vigilancia tecnológica" y de la "inteligencia económica" aparece como el denominador común de todas estas operaciones en donde la información representa, como hemos dicho en la introducción, una "materia prima" que hay que procesar para obtener una información útil. Esto explica la puesta en marcha de una investigación en el campo de la ciencia y de la tecnología de la información, que tiene como principal objetivo la concepción y la producción de instrumentos (es decir, indicadores, métodos y herramientas computacionales) de análisis de la información científica y tecnológica.

**Lingüística computacional.** Desde el momento que se trata de trabajar al nivel del texto integral en grandes corpus, el procesamiento automático del lenguaje natural es una necesidad para operar una extracción terminológica y liberarse de la indización manual. Nos interesamos igualmente por definir indicadores lingüísticos de conocimiento más complejos que las simples palabras claves. La variación y la estabilidad de las expresiones científicas, tal que nosotros podemos observarlas en corpus, pueden servir para definir tales indicadores. Este aspecto ha sido abordado en Polanco et al. (1995a) y (1995b), Royauté et al. (1996).

**Infometría.** Hemos adoptado el término "infometría" para significar de una manera general los análisis métricos de la información (es decir estadísticos, probabilistas y de análisis de datos), aplicados a la producción de una "información de la información" del tipo "quién hace qué, en cooperación con quién, en donde y cuando". Aquí se trata de concebir y desarrollar técnicas de clasificación automática y de representación gráfica (mapas). En este marco, es necesario apoyarse en los recursos que representan las matemáticas aplicadas al análisis de la información.

**Inteligencia artificial.** Cuando se trata de trabajar sobre el conocimiento contenido en los documentos, entonces es necesario apoyarse sobre las técnicas (numéricas y simbólicas) de la inteligencia artificial. Puesto que el objetivo es pasar del nivel de los términos a aquél de los conceptos y de las proposiciones y poder así abordar los problemas de representación de conocimientos y de razonamiento (inferencias). Este aspecto de "ingeniería del conocimiento" lo hemos enunciado en Polanco (1998) desde el punto de vista de una "teoría del conocimiento sin sujeto" (Popper, 1979).

## 5 UN SISTEMA PARA EL ANÁLISIS DE L'INFORMACIÓN

Hemos concebido y desarrollado en la Unidad de Investigación e Innovación del INIST-CNRS un dispositivo permitiendo: [1] el análisis estadístico descriptivo de los datos bibliográficos, [2] la ubicación y extracción terminológica (en francés e inglés) a partir de los títulos y resúmenes, [3] la organización temática de los datos y su visualización bajo la forma de mapas, aplicando procedimientos de clasificación automática y de cartografía, [4] el almacenamiento de los resultados en un SGBD relacional y su acceso a través del Web, mediante una interface especialmente concebida para el análisis de la información organizada en clusters y mapas.

<b>SISTEMA PARA EL ANÁLISIS DE LA INFORMACIÓN</b>	
<b>STANALYST™</b>	
<p><b>[1] ESTADÍSTICAS DESCRIPTIVAS</b></p> <ul style="list-style-type: none"> <li>- MIRIAD (servidor WWW generando páginas HTML) permite hacer estadísticas sobre los datos bibliográficos</li> </ul>	<p><b>[3] CLASIFICACIÓN AUTOMÁTICA Y CARTOGRAFÍA</b></p> <ul style="list-style-type: none"> <li>- NEURODOC (enfoque neuronal) y SDOC (enfoque palabras asociadas: «cword analysis») para la generación de clusters y mapas o cartas</li> </ul>
<p><b>[2] PROCESAMIENTO DEL LENGUAJE NATURAL</b></p> <ul style="list-style-type: none"> <li>- Plataforma ILC, herramienta de ingeniería lingüística para la extracción terminológica (en francés e inglés) a partir del texto integral (lenguaje escrito)</li> </ul>	<p><b>[4] SGBD-WWW</b></p> <ul style="list-style-type: none"> <li>- HENOCH, organiza los resultados de los programas anteriores en un SGBD relacional y asegura el acceso a través de una interface WWW especialmente concebida para analizar la información</li> </ul>

Este sistema ha sido aplicado en la realización de más o menos veinte estudios de vigilancia entre 1995 y 1998. Citemos, por ejemplo, su aplicación en el análisis de 1.857 datos bibliográficos sobre las enfermedades provocadas por los "prions" (es decir, "proteinaceous infectious particles"), extraídos del Science Citation Index (CD-ROM), en vista de un test de softwares de vigilancia científica, a pedido de la Dirección del Departamento Ciencias de la Vida y de la Unidad de Indicadores de Política Científica del CNRS. Esta experiencia nos permite considerar el empleo de un tal dispositivo para la vigilancia y la inteligencia científica al servicio de los laboratorios y direcciones científicas del CNRS.

En el marco de la valorización económica, este mismo sistema ha sido empleado en la realización de un producto comercial de vigilancia científica y tecnológica titulado « *Les Plantes Transgéniques : Rapport de Tendence* » (Octubre 1997)<sup>©</sup>, asociando el INIST y una sociedad privada especializada en el campo de la información, el Bureau Van Dijk.

## 6 ADQUISICIÓN Y ORGANIZACIÓN DE CONOCIMIENTOS

En lugar de coleccionar los conocimientos por medio de entrevistas o encuestas, el enfoque que seguimos representa la alternativa de proceder a la adquisición y organización de los conocimientos en corpus, es decir, a partir de la información presente en un conjunto de textos científico y técnicos del campo considerado. Esta elección está justificada en la medida que los textos (artículos y patentes) son pertinentes para dar cuenta del estado del conocimiento en un campo del saber (científico y tecnológico) a un momento dado de su desarrollo.

Este enfoque ha dado lugar a una metodología de adquisición y estructuración de conocimientos a partir de la información presente en un conjunto de textos. Ella se apoya en la combinación de tres tipos de procesamientos: de lingüística computacional capaz de identificar los términos de una nomenclatura científica dada, de estadística y clasificación algorítmica permitiendo la generación de clusters y mapas, y enseguida la intervención de los expertos para la descripción conceptual de los clusters y el comentario de los mapas. Esta metodología comporta en su realización cuatro fases: [1] adquisición terminológica, [2] control del vocabulario de indexación, [3] clasificación de los términos y documentos (constitución de los clusters y mapas), [4] descripción conceptual de los clusters y comentario de los mapas. Dos fases son enteramente automáticas (1 y 3), mientras que las dos otras (2 y 4) suponen una intervención humana, aquélla de los expertos del campo de aplicación.

El enfoque estadístico (fase 3) proporciona el soporte objetivo para que los expertos precisen cuál es el significado conceptual de los clusters (fase 4). En efecto, la parte más importante del trabajo de los expertos está consagrada en la práctica al análisis de los clusters con el objetivo de determinar su significación conceptual, y caracterizar en función de esta significación el mapa en el cual la posición relativa de los clusters se encuentra representada. Los conocimientos - producidos por este trabajo de "expertise" - deben ahora ser procesados para poder introducirlos en un sistema de gestión de conocimientos como lo hemos expuesto en Polanco et al (1998d). Y de esta manera ir hacia una capitalización de los conocimientos producidos por los expertos en el momento de analizar los clusters y los mapas que han sido generados a partir de los datos mismos.

## 7 PERSPECTIVAS

Sobre la base de lo adquirido y guardando siempre nuestra orientación, a saber la concepción y el desarrollo de tecnologías al servicio del análisis de la información, se trata ahora de hacer que el análisis devenga efectivamente un proceso de adquisición y de apropiación de los conocimientos expresados bajo la forma de datos textuales (artículos, patentes, informes, notas clínicas, etc.). En otras palabras, la adquisición del conocimiento que se expresa por medio del lenguaje escrito. Y por otra parte, como hemos dicho en el parágrafo anterior, proceder igualmente a la capitalización del conocimiento emitido por los expertos cuando analizan los clusters y los mapas.

El fortalecimiento de estas tres direcciones complementarias [1] el procesamiento del lenguaje escrito en tanto que soporte de la información científica y tecnológica, [2] los métodos infométricos de análisis de la información (clasificación y cartografía), [3] la adquisición y gestión de conocimientos a partir de los datos textuales, nos conducen hacia disciplinas como el "descubrimiento de conocimientos en las bases de datos" (*knowledge discovery in databases*), la "explotación de datos" (*data mining*), y la "gestión de conocimientos" (*knowledge management*).

### **a) Explotación de datos textuales (text-data mining)**

Actualmente disponemos de una plataforma de ingeniería lingüística, la plataforma ILC, que permite procesar los datos textuales, y proporciona a los programas infométricos (NEURODOC y SDOC) un medio de extracción terminológica en francés e inglés. Su extensión al español se encuentra en curso de realización. Las investigaciones en lingüística computacional deberán continuarse para hacer aún más eficaz la adquisición de conocimientos y poder dar un sentido a las relaciones entre términos. La asociación de métodos lingüísticos e infométricos es pertinente para extraer la información de un corpus de textos, y para hacer aparecer la estructuración intrínseca de los conocimientos en corpus. Se trata ahora de buscar al nivel de los textos las trazas de las relaciones que existen entre los términos. El resultado de nuestras experiencias es que la información estructurada, comentada por los ingenieros documentalistas especializados en el campo de aplicación, constituye una base parcial de conocimientos. Pretendemos, y tal es el objetivo de nuestra investigación en curso, incorporar un formalismo de representación de conocimientos. Al respecto, Capponi (1999) ha propuesto en su tesis de doctorado un método original de generalización de las estructuras predicativas, presentando igualmente (Capponi y Toussaint, 1998) un método para la interpretación de los clusters basado en una "lógica de descripción" (Napoli, 1998).

## **b) Una plataforma a base de redes neuronales**

Nuestro interés por los algoritmos neuronales se apoya sobre los lazos que existen entre el análisis de datos (enfoque estadístico multidimensional) y el enfoque conexionista en lo que respecta la clasificación automática y las representaciones factoriales (cartografía). El análisis cuantitativo de la información (infometría) puede ser completado por medio de redes neuronales. El objetivo es convertir NEURODOC en una plataforma a base de redes neuronales para la clasificación (creación de clusters) y la cartografía (posicionamiento de los clusters sobre un espacio métrico permitiendo no solamente visualizar sino que también evaluar la posición relativa de ellos). Hemos ya sometido a prueba algunos modelos neuronales (Perceptron, ART-1, Kohonen, Neural-Gas), y comparado sus resultados con los obtenidos por los métodos de análisis de datos que NEURODOC aplica, "k-means axiales" y "análisis en componentes principales". Esta experiencia nos permite considerar, para un futuro próximo, métodos de clasificación y cartografía más robustos y de calidad superior, así que un mejor manejo de la noción de aprendizaje (supervisado y no supervisado). El estado de nuestros trabajos en este campo de investigación está expuesto en Polanco et al. (1998b) y (1998c).

## **c) Un sistema de gestión de conocimientos**

La gestión de conocimientos supone su implementación por medio de un sistema computacional. Creemos que disponemos en HENOCH del principio de este sistema, a saber una arquitectura mixta SGBD-WWW (a propósito de HENOCH, ver Grivel et al. 1997 y 1999). El sistema debe poder acoger los resultados de la estructuración de conocimientos permitiendo así no solamente la exploración y visualización, sino que también la posibilidad de realizar operaciones relativas a estos conocimientos. El sistema debe administrar al menos tres tipos de datos que deseamos poder cruzar: los clusters, los conocimientos y los datos bibliográficos o textuales (que pueden ser por lo demás de diferentes tipos). La idea es de modelar los datos bibliográficos, los clusters obtenidos a partir de estos datos, y las "clases" de conocimientos obtenidos por el análisis lógico de las descripciones de los clusters efectuadas por los expertos. El objetivo es entonces que HENOCH evolue hacia un sistema de gestión de conocimientos (fundado en el modelo relacional u orientado objeto) capaz de asegurar un trabajo de análisis en cooperación (en grupo) y a distancia, de acuerdo con el principio de interfaces "metafóricas", es decir, un modo de relación entre el humano y el computador, que permite al usuario de trabajar sin tener necesidad de un aprendizaje computacional fastidioso de procedimientos y controles.

# **8 CONCLUSIÓN**

1. Hemos considerado las actividades de vigilancia e inteligencia, en el contexto de la "sociedad de la información", como "la transformación de la información en conocimiento, y del conocimiento en acción". En nuestra opinión, esta actitud activa con relación a la información pide más que un sistema de información en su acepción tradicional.
2. Es por esto que hemos subrayado la importancia del "análisis de la información" en el seno de todas las actividades en donde se trata de transformar los datos brutos con el fin de extraer los conocimientos que puedan ser explotados y útiles en un determinado campo de acción. De ahí, entonces, la idea del "análisis asistido por computador" gracias a las tecnologías que hemos expuesto y que llamaremos "tecnologías de la inteligencia", y la puesta en práctica de un cierto enfoque metodológico en cuanto a la "adquisición y organización de conocimiento" en el campo de la información científica y tecnológica (artículos, patentes).
3. Finalmente, hemos presentado las perspectivas de trabajo de la Unité de Investigación e Innovación del INIST-CNRS: [1] la explotación de datos textuales, [2] la aplicación de redes neuronales en clasificación y cartografía, [3] la adquisición, organización, representación y gestión de conocimientos.
4. Tanto el enfoque así mismo que las tecnologías que hemos presentado traducen la voluntad explícita de ir más allá de los indicadores convencionales de la cienciometría tradicional, es decir, del estudio estadístico de las publicaciones científicas y patentes, para interesarnos mucho más por los contenidos de conocimiento que pueden ser útiles y necesarios en la toma de decisiones en ciencia y tecnología.

# **9 BIBLIOGRAFÍA**

- Brookes, B.C. (1980-81) "The Foundations of Information Science", *Journal of Information Science*, vols. 2 y 3.
- Callon M., J. Law, A. Rip (1986) "Qualitative Scientometrics", in M. Callon, J. Law, A. Rip (editors), *Mapping The Dynamics of Science and Technology*. London : The Macmillan Press, pp. 103-123.
- Capponi, C. (1999) *Généralisation de structures prédicatives. Application à l'analyse de l'information*. Tesis de doctorado. Universidad Henri Poincaré - Nancy 1.

- Capponi, C. y Y. Toussaint (1998) "Interprétation de classes de termes par généralisation de structures prédicat-arguments", *Actes IC'98 Ingénierie des Connaissances*, 13-15 mai 1998, Pont-à-Mousson, pp. 41-50.
- Duff, S. (1995) "The information society as paradigm: a bibliometric inquiry", *Journal of Information Science*, vol. 21, n° 5, pp. 390-395.
- Fayyad U.M., y G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy, Editors, (1996) *Advances in Knowledge Discovery and Data Mining*. Menlo Park, Calif. AAAI Press / The MIT Press.
- Grivel, L. y X. Polanco, A. Kaplan (1997) "A Computer System for Big Scientometrics at the Age of the World Wide Web", *Scientometrics*, vol. 40, n° 3, pp. 493-506.
- Grivel, L. (1999) "HENOCH, un outil d'analyse de corpus d'information scientifique et technique", *Le Micro Bulletin Thématique: "L'information scientifique et technique et l'outil Internet. Experiences, recherches et enjeux pour les professionnels de l'IST"*. Paris, CNRS, pp. 27-44.
- Kuhn, T.S. (1970) *The Structure of Scientific Revolutions*. Chicago, The University of Chicago Press, Second Edition, Enlarged.
- Napoli A. (1998) "Une introduction aux logiques de description", in R. Ducournau, J. Euzenat, G. Masini, A. Napoli, *Langages et Modèles Objets*, Le Chesnay: INRIA (Collection Didactique), pp.321-350.
- Polanco, X. y L. Grivel, J. Royauté. (1995a) "How To Do Things with Terms in Informetrics : Terminological Variation and Stabilization as Science Watch Indicators", *Proceedings of the Fifth International Conference of the International Society for Scientometrics and Informetrics*. Edited by M.E.D. Koenig & A. Bookstein. Medford, N.J., Learned Information Inc., pp. 435-444.
- Polanco, X. y J. Royauté, L. Grivel, A. Courgey (1995b) "Infométrie et linguistique informatique : une approche linguistico-infométrique au service de la veille scientifique et technique", *Journées d'études sur les systèmes d'information élaborée de la SFBA*, Ile Rousse, Corse, 19p.
- Polanco, X. (1998a) "Extraction et modélisation des connaissances : une approche et ses technologies (EMCAT)", *Organisation des connaissances en vue de leur intégration dans les systèmes de représentation et de recherche d'information*. [Colloque chapitre français d'ISKO, Lille, France, 16-17 octobre 1997]. Sous la direction de Jacques Maniez et de Widad Mustafa el Hadi. Préface de Gérard Losfeld. Université Charles De Gaulle - LILLE 3, Collection UL3, pp. 101-116.
- Polanco, X. y C. François, J-P. Keim (1998b) "Artificial neural network technology for the classification and cartography of scientific and technical information", *Scientometrics*, vol. 41, n° 1, pp. 69-82.
- Polanco, X. y C. François, A. Ould Louly. (1998c) "For Visualization-Based Analysis Tools in Knowledge Discovery Process : A Multilayer Perceptron versus Principal Components Analysis - A Comparative Study", J.M. Zytkow and M. Quafafou (eds) *Principles of Data Mining and Knowledge Discovery*. Second European Symposium, PKDD'98, Nantes, France, 23-26 September 1998. Lecture Note in Artificial Intelligence 1510. Subseries of Lecture Notes in Computer Science. Berlin, Springer, pp. 28-37.
- Polanco, X. y C. François, J. Royauté, L. Grivel, D. Besagni, M.Dejean, C. Otto (1998d) "Organisation et gestion des connaissances en veille scientifique et technologique", VSST'98 (Veille Stratégique Scientifique et Technologique), Toulouse, 19-23 octobre 1998, pp. 328-337.
- Popper, K.R. (1979) *Objective Knowledge*. Oxford, The Clarendon Press, Revised Edition.
- Rip A., J-P. Courtial, "Co-Word Maps of Biotechnology : An Example of Cognitive Scientometrics", *Scientometrics*, vol. 6, n° 6 (1984), pp. 381-400.
- Royauté, J. y C. Muller, X. Polanco. (1996) "Une approche linguistique infométrique de la variation terminologique pour l'analyse de l'information", *Colloque ILN'96, Informatique et Langage Naturel*, Institut de Recherche en Informatique de Nantes, Université de Nantes, 9-10 octobre 1996, pp. 563-581.