Jointly published by Akadémiai Kiadó, Budapest
and Springer, Dordrecht

Scientometrics, Vol. 79, No. 2 (2009) 365–375
DOI: 10.1007/s11192-009-0424-0

Citation structure of an emerging research area on the verge of application

HENRY SMALL,[a] PHINEAS UPHAM[b]

[a] *Thomson Reuters, 3501 Market St., Philadelphia, Pa. 19104, USA*
[b] *The Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, Pa. 19104, USA*

A case study of an emerging research area is presented dealing with the creation of organic thin film transistors, a subtopic within the general area called "plastic electronics." The purpose of this case study is to determine the structural properties of the citation network that may be characteristic of the emergence, development, and application or demise of a research area. Research on organic thin film transistors is highly interdisciplinary, involving journals and research groups from physics, chemistry, materials science, and engineering. There is a clear path to industrial applications if certain technical problems can be overcome. Despite the applied nature and potential for patentable inventions, scholarly publications from both academia and industry have continued at a rapid pace through 2007. The question is whether the bibliometric indicators point to a decline in this area due to imminent commercialization or to insurmountable technical problems with these materials.

Introduction

Organic semiconductors have been known since the 1940s, and the first transistor based on an organic semiconductor was reported in 1986 [COLLINS, 1986]. The 2000 Nobel Prize in Chemistry recognized the contributions to this area by Heeger, MacDiarmid, and Shirakawa for work done in the late 1970s. They were able to create semiconductors by doping the polymer polyacetylene. Conducting plastics have already been used in a number of applications, such as light emitting diodes. However, research in this field has recently picked up pace with new discoveries and the introduction of new materials.

Organic thin film transistors hold out the promise of the inexpensive manufacture of electronic and computing devices by printing on plastic. They have attracted much interest due to their potential applications in low-cost, light-weight, flexible, large-area applications such as smart cards, radio-frequency identification tags, flat panel displays, and computing in clothing. Hence, this is an example of a basic science field with great practical potential. However, despite this progress and potential, several technical issues remain to be resolved before organic thin film transistors can be in wide scale use.

Received December 5, 2007

Address for correspondence:
HENRY SMALL
E-mail: henry.small@Thomson.com

0138–9130/US $ 20.00

The field of organic thin film transistors came to our attention because it was the largest emerging cluster in an analysis of Web of Science data from the 1999–2004 period. This cluster first appeared in a research front analysis in August of 2004, when it consisted of only three highly cited papers. By year-end 2004 it had expanded to 26 highly cited papers.

Method

This area was identified using the method of co-citation clustering [SMALL, 2006] during a routine analysis. An outline of this methodology is given in the following link: http://www.esi-topics.com/RFmethodology.html.. To briefly review this process, the top 1% of highly cited papers is selected from each of 22 broad fields of science defined by journal sets for each year within a six year rolling file. The highly cited papers are subject to a single-link cluster analysis, adjusting a cosine normalized co-citation threshold to create a cluster within a prescribed size limit. This process creates approximately 5,000 clusters, and is repeated every two months with updated data for the six-year rolling file. Each set of clusters for one time period is matched against the clusters for the prior time period to look for continuing areas as well as "new" clusters, that is, those containing highly cited papers that did not appear in any cluster in the prior period. These constitute the so-called "emerging research areas". Clusters are linked across time periods if they share continuing highly cited papers. Following these links across time allows us to create what are called cluster strings, the idea for which derives from Garfield's historiographs [GARFIELD & AL., 2003]. The methodology of co-citation combined with single linkage is particularly well suited to obtaining a rapid and broad overview of the research areas represented in a large set of highly cited documents. This provides a rough delineation and indicator of currently active specialties that can be refined using other more rigorous techniques, such as those described by CHEN [2006] and MORRIS [2003]. It is not claimed that the methodology used here provides the most accurate or refined picture possible of a research area, but rather an important initial view, and a broad-based science monitoring strategy.

Additionally, each cluster can be displayed as a two-dimensional configuration generated using a force-directed placement algorithm [SMALL, 2006]. This procedure works by setting up attractive forces between co-cited papers and repulsive forces between all papers. The attractive force is directly proportional to the inter-paper distance multiplied by the cosine of the co-citation strength, and the repulsive force is inversely proportional to the square of the distance. The configuration is formed by successive addition of co-citation links, starting with only the strongest links that form a minimal spanning tree, and adding links successively in a series of iterations, each iteration adding the next strongest link for each paper. In each iteration the positions of papers are adjusted to minimize the force on each paper. The amount of residual force per node at the end of the process is used as a goodness of fit measure. The limitations of this methodology are that weak co-citation links may not be taken into account in the positioning of nodes, that as in other ordination methods, such as multidimensional scaling, local minima may be reached, and as in any attempt to represent complex networks in two dimensions only approximate fits may be achieved.

Emerging cluster on organic thin film transistors

The co-citation structure of the organic thin film transistor cluster is shown in Figure 1. The cluster consists of 26 highly cited papers. Each circle is a paper whose area is proportional to its citation count. Papers with the most recent publication year are darker in color. Papers are connected only by the strongest normalized co-citation links per paper (solid lines), plus additional weaker links sufficient to connect the remaining components (dashed lines). This forms a minimal spanning tree, with 25 links connecting the 26 papers. Actually there are 275 co-citation links among these 26 papers, that is, 85% of the theoretically possible connections. The residual force per node for this map is 0.064.

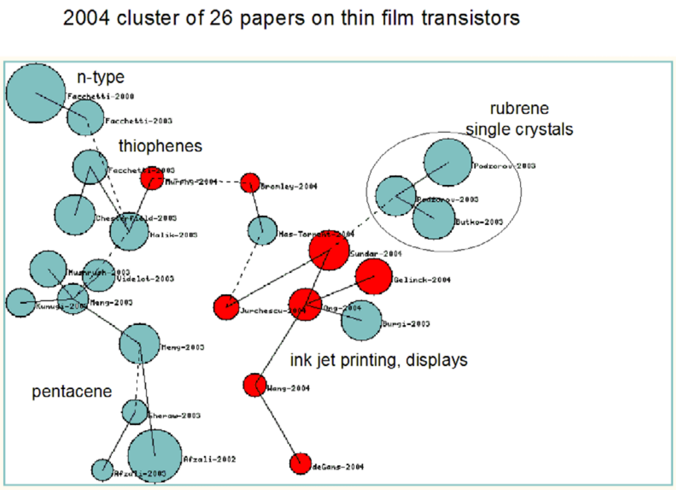

Figure 1. Map of the initial 2004 cluster on organic thin film transistors

The cluster can be labelled by region according to the major themes of research. The oldest paper in the cluster from 2000 at the upper left is a demonstration of an n-type, or electron conducting transistor in an organic thin film. By contrast most organic transistors are p-type or hole conducting. Research tends to group around specific organic materials with transistor potential such as thiophenes, pentacene and rubrene. In addition, there is a grouping on the lower right dealing with fabrication methods such as printing on plastic, and applications such as flat panel displays. Many of these themes will continue in the next few years.

Co-citing contexts were examined for each pair of papers along the path of the spanning tree [SMALL, 1986]. The majority of links were what might be termed rhetorical modifications (paper A does X but paper B does Y). In other words an author would cite one paper as having somehow modified, qualified, or improved upon the results of the other paper. This finding illustrates the high degree of competition in the field regarding the best technical approach to creating organic transistors. A smaller number of links might be characterized as applications (A has applied B), or redundancies (A is equivalent to B).

It is also of interest to view the position of this specialty within nanoscience as a whole. This can be accomplished by clustering clusters in a second and third iteration [SMALL, 2006]. The following map (Figure 2) was derived for the six year period ending in 2006. The organic thin film transistor specialty is situated at the lower right linked to solar cells and adjacent to light emitting diodes. LEDs were the first practical devices based on plastic electronics, while solar cells are considered to be a very promising area of application for organic transistors.

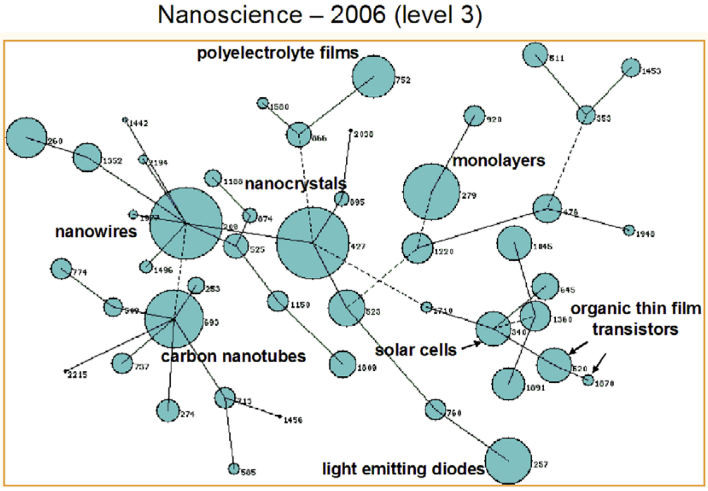

Figure 2. Level 3 map showing the context of organic thin film transistors within nanoscience in 2006

Citation and social characteristics

To examine some citation and social characteristics of the cluster we analyze the citing papers for the set of 26 highly cited papers. Earlier studies have shown that the authors within these clusters tend to know one another and form social groups (Crane). There are 228 citing papers, and 15 of these (6.6%) are in the set of 26 highly cited papers in the six-year period. This percent is a measure of the inbred nature of the cluster. The 228 citing items cite the highly cited papers 460 times. Seventeen percent of these citations were self-citations. The authors on the highly cited papers were also authors on 40% of the citing papers, accounting for about 50% of the citations. This means that highly cited authors remain involved in writing current papers and cite each other frequently. Thus, the leading authors are still active in the field, underlining the small-world nature of this group, where citing and cited populations are highly overlapped. This measure also gives a sense of how a small set of leading authors are driving this specialty forward. Finally, the currency of the cluster is indicated by the percentage of cited items in the most recent year (2004), as well as the percentage of citing items in that year, as shown in Table 1.

Table 1. Citation characteristics of emerging cluster on thin film transistors

	Thin-film transistors
% self-citations (all author)	16.9%
% citing papers by cited authors	39.9%
% cited items in most recent year	30.7%
% citing items in most recent year	78.9%
% citing items that are cited items	6.6%

The drive toward applications is evident if we examine the cited and citing papers. All but five of the 26 highly cited papers specifically mention possible applications in their first paragraphs. Ten of the 26 papers also point to “significant recent progress” in their first paragraphs. This is echoed by the citing authors. The progress alluded to is often the increased charge carrier mobility, an important characteristic for the organic materials. The papers reveal a great deal of tinkering with systems in order to get better charge mobility. All investigators appear to be racing toward the same goal – finding a practical and economically viable device. As noted above, the specialty not only mixes disciplines but also performing sectors. Table 2 shows the institutional mix of author addresses across sectors, both for the cited and citing populations.

Table 2. Sector of cited and citing authors for the initial cluster

Sector	Percent of cited addresses	Percent of citing addresses
Academic	60.7%	60.8%
Industry	33.6%	27.5%
Government	5.6%	11.6%

Interviews

Eight authors of highly cited papers were interviewed via email questionnaire. All interviews were posted at http://www.esi-topics.com/otft/. The authors were asked to comment on their papers and on the state of the field of organic thin film transistors. Almost all interviewees noted that the field was very close to commercial application, listing numerous potential products, but pointing out that some technical hurdles remained. Most mentioned the need to continue to improve the charge carrier mobility of the organic materials, the importance of developing low cost fabrication methods, and improving device stability. They also noted recent improvements in the understanding of the basic physics which allows them to better predict performance and determine the limits of their devices.

Most mentioned how interdisciplinary collaboration had been critical to the development of the field, involving either cross-training in different disciplines or recruiting new team members from different fields. Some commented on the competition from silicon based devices which often had higher performance characteristics, but felt that organic materials were potentially more versatile, cheaper to produce and offered new and expanding possibilities. Some felt that the field needed a niche product that silicon materials could not duplicate.

A telephone interview was also conducted with an industrial scientist not among the highly cited group. This individual confirmed many of the views expressed by the high cited authors and reported that the field was currently poised between basic research and applications, what the interviewee called the “development phase,” with some significant obstacles to overcome before commercialization can begin. The focus on applications and economic payoff is reflected in the involvement of authors from companies such as Lucent, IBM, Infineon, Philips, and Xerox, often in collaboration with university partners.

The respondent also pointed to the fact that large companies are entering the field and setting up expensive multidisciplinary research teams similar to those in academia, including materials scientists, physicists, engineers, and chemists. It was suggested that these new interdisciplinary teams may, in part, account for the upsurge in number of publications. However, the race for commercialization has also made companies and universities more secretive about their work, for example, making collegial interactions at conferences more difficult. In addition, often patents are applied for prior to the publication of results, which adds to the time lag for dissemination [MURRAY].

Thus, the emergence of this front in 2004 has to do with both the increase in activity sparked by new advances in basic and applied research, as well as the increased investment of resources by academia and industry to overcome the remaining technical obstacles to commercialization. Ironically, if publication activity declines in the future, it might signal either the privatization of the field as patenting replaces publication, or alternatively the failure of the field to overcome the remaining technical barriers. Therefore, it is of interest to trace the subsequent history of this area up to the present to look for signs of commercialization, or intellectual decline.

Cluster evolution 2004–2007

A cluster string (Figures 3 and 4) is used to represent the evolution of the organic thin film transistors area [SMALL, 2006]. The string represents the complex splitting and merging pattern of cluster development from December 2004 through February 2007. Each cluster is represented by a circle with the number of highly cited (core) papers shown in the circle. For example, the initial cluster of 26 highly cited papers is shown on the left. The lines connect clusters in adjacent years that share continuing papers. The number of common papers is indicated on the line. Links are normalized using the cosine formula (number of common papers divided by the square root of the product of numbers of papers in the linked clusters). Only links with a cosine of 0.1 or higher are drawn. This breaks only one weak link that leads into organic chemistry. Line thickness indicates the strength of the inter-year connection. Note that summing the cluster sizes gives a growth in the specialty size of from 30 highly cited papers in December 2004 to 53 by August 2006 to 61 in February 2007.

The initial cluster of 26 splits and most of its papers lead to a large cluster of 32 papers in December 2005. The cluster of 32 papers from 2005 splits in June 2006 into three groups leading predominantly into a cluster consisting of 24 papers. This group is mainly devoted to n-type devices. After only two months, in August of 2006, this main cluster of 24 papers fragments (see Figure 3), one branch leading to what appears to be a new dominant area of 19 papers which has formed by the merging of four independent lines of research previously separate from the main path of development. Thus, in August 2006 there appears to be a reorganization and coalescence of research around a new approach, and a fragmentation and diminution of the previously dominant group. The largest of the resulting fragments from the break-up of the prior dominant group is a cluster of eight papers dealing almost exclusively with n-type semiconductors, with papers originating primarily from two research labs. Such short-term disruptions have been observed in prior co-citation studies [SULLIVAN & AL., 1977; SMALL, 1977].

The new dominant area of 19 papers receives its main influx of papers from a cluster of nine papers from June 2006 focused on what is called regioregular organic structures which also deals with the role of molecular weight of organic compounds in improving charge mobility. All nine papers from this prior group are incorporated into the cluster of 19 papers. From the previous dominant group of 24 papers, the new group of 19 papers receives three papers concerned with printing methods.

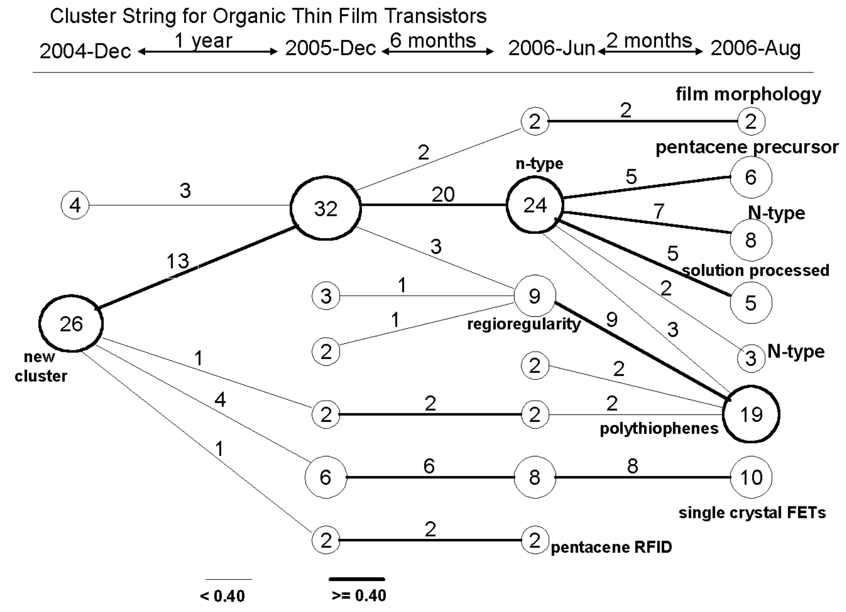

Figure 3. Cluster string representation of organic thin film transistors (Dec 2004 through Aug 2006)

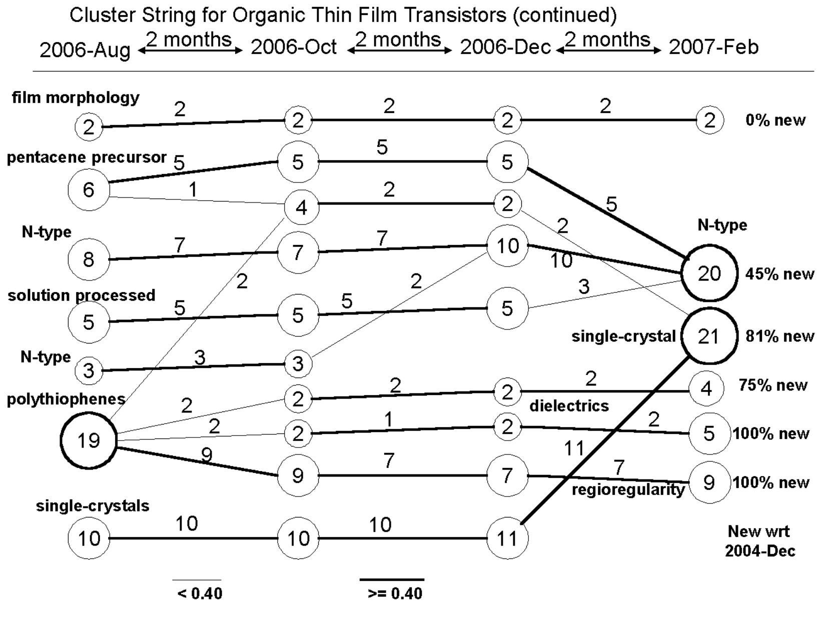

Figure 4. Cluster string representation of organic thin film transistors (Aug 2006 through Feb 2007)

Two smaller groups that merge into the cluster of 19 may hint at the growing sense that practical applications are not far off. One of these small groups deals with "device physics" and "new applications". The other group deals with "high performance thiophenes", suggesting that there is a new emphasis on devices with high performance.

Figure 4 tracks the research area starting in August 2006 (also shown in Figure 3) and ending in February 2007, updating every two months. The group of 19 from August 2006 labelled polythiophenes is shown to splinter in October but the main lines of research persist until there is a new consolidation in February 2007. The largest consolidated areas are n-type conductors and single-crystal organic transistors. The third largest area deals with regioregularity. Tracking these clusters through June 2007 (data not shown) does not reveal any major shifts from the February configuration.

One way of thinking about the rate of knowledge change in these lines of research is to focus on the proportion of new and continuing core papers with respect to some base year. Using December 2004 as the base, it is seen that the regioregular cluster is the most dynamic of the larger groups with 100% of its core papers new with respect to 2004. The large group on single-crystals also shows rapid change with 81% new papers, whereas the n-type group moves more slowly with 45% new. Two small clusters dealing with dielectrics are also fairly new with respect to 2004. New micro-areas show that the field is still innovating. However, the absence of large scale change might also be interpreted as a sign of lack of progress.

Conclusions

It is too soon to say whether any of these developments signals a significant breakthrough, a clear path toward applications, or the reaching of a dead end. The overall pattern of development in the field of organic thin film transistors has been one of twigging or the spawning of new lines of research. This is to be expected in an area where knowledge is expanding rapidly. Only one area seems to have ceased, namely pentacene RFIDs, a clear area of application.

The opposite of twigging, namely coalescence, is also evident at specific points in time and may be more critical for the development of the specialty. For example, there is the merging in June 2006 leading to the formation of the cluster of nine papers, and in August 2006 with the formation of the cluster of 19 papers. In these cases the coalescence or merging, which takes place within an overall pattern of twigging, signals the formation of a new dominant theme of research. A number of specific concepts seem to be connected with this coalescence, including, "regioregularity" in the polymer structure, the use of molecular weight to improve performance, and the applications of insulators and gate dielectrics. Another notable consolidation event occurred in February 2007 with the coalescence of n-type organic transistors and of single-crystal transistors.

Throughout both the splitting and merging processes of specialty evolution it is clear that certain structures of strongly co-cited documents are conserved, rearranged, and exchanged from one cluster to the next. These units of exchange may consist of only a few papers, are highly topic specific, and behave as collective concept markers for the specialty. These conserved groups are reminiscent of the memes of DAWKINS [1976], units of cultural transmission. Examples are paper sets on insulator/gate dielectrics, regioregular structure, n-type conducting, and printing. How this phenomenon relates to Chen's "pivot points" remains to be investigated [CHEN, 2006]. We also see that new small clusters can merge with larger groupings later on and be incorporated into their structures, representing perhaps a delayed recognition of their relevance. Other patterns are more stable and represent relatively isolated groups that continue without much change from year to year, for example, the single crystal field effect transistor group which splits off from the original cluster and continues as a separate area. Whether in forecasting developments it is best to track the new micro-clusters or the larger established groupings remains a topic for future research.

Since we know that research in this area has been going on for many years, what accounts for the sudden emergence of this cluster at the end of 2004? The immediate cause was of course a sudden surge in publication and citation rates, bringing paper citation and co-citation counts above our detection thresholds. From interviews we can speculate that this surge was due to a number of factors, including an explosion of new knowledge, the infusion of funding from industry, the productivity of newly formed multidisciplinary teams, and the expectation that the first practical devices incorporating organic thin film transistors were just in the offing – a kind of sprint to the finish line phenomenon. While industry seems to be playing an increasing role in this research, significant contributions continue to come from academia, and we might speculate that those involved in developing applications are dependent upon the basic researchers to solve some of the remaining technical hurdles to commercialization such as charge mobility, stability and fabrication methods. In addition, it is likely that academia will remain the main source of new novel organic materials with semiconducting properties that will be relied upon by industry well into the future. This study again demonstrates that the path to applications from basic science is not a straight line but rather a complex interplay of ideas and technology where both research and development are moving in parallel.

If we apply the rule that a cessation of publication would signal that the field has moved into an applications phase, where patenting and industrial secrecy would hold sway, we do not see as yet any evidence of this happening. On the contrary, the response from active investigators indicates that the pace of basic research is still strong with the expectation that this research will ultimately pay off with new applications and devices having major economic impact.

References

CHEN, C. (2006), CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57 (3) : 359–377.

COLLINS, G. P. (2004), Next stretch for plastic electronics. *Scientific American*, 291 : 74v81.

CRANE, D. (1972), *Invisible Colleges Diffusion of Knowledge in Scientific Communities*. Chicago, IL, University of Chicago Press.

DAWKINS, R. (1976), *The Selfish Gene*. Oxford, UK: Oxford University Press.

GARFIELD, E., PUDOVKIN, A. I., ISTOMIN, V. S. (2003), Why do we need algorithmic historiography? *Journal of the American Society for Information Science and Technology*, 54 (5) : 400–412.

MORRIS, S. A., YEN, G., WU, Z., ASNAKE, B. (2003), Timeline visualization of research fronts. *Journal of the American Society for Information Science and Technology*, 55 (5) : 413–422.

MURRAY, F. S. (2005), Do intellectual property rights hinder the free flow of scientific knowledge? *NBER Academic Science and Entrepreneurship Conference.*

SMALL, H. (1977), A co-citation model of a scientific specialty: a longitudinal study of collagen research. *Social Studies of Science*, 7 : 139–166.

SMALL, H. (1986), The synthesis of specialty narratives from co-citation clusters. *Journal of the American Society for Information Science*, 37 (3) : 97–110.

SMALL, H. (2006), Tracking and predicting growth areas in science. *Scientometrics*, 68 (3) : 595–610.

SULLIVAN, D., WHITE, D. H. BARBONI, E. J. (1977), Co-citation analyses of science: an evaluation. *Social Studies of Science*, 7 (2) : 223–240.

UPHAM, P., SMALL, H. (2007), Emerging research fronts in science and technology: Patterns of new knowledge development. To be published.