A new approach to explore the knowledge transition path in the evolution of science & technology: From the biology of restriction enzymes to their application in biotechnology

Xiaojun Hu [a], Ronald Rousseau [b,c,]*

[a] *School of Medicine, Zhejiang University, Hangzhou, 310058, China*
[b] *Facultair Onderzoekscentrum ECOOM, KU Leuven, Naamsestraat 61, Leuven, B-3000, Belgium*
[c] *Faculty of Social Sciences, University of Antwerp (UA), Antwerp, B-2000, Belgium*

A R T I C L E   I N F O

A B S T R A C T

In this contribution, we develop a new approach to explore the process of knowledge transition from discovery-oriented science to technological fields, via applications-oriented research, including a mediator set. This trajectory is referred to as the D-A-T trajectory. It is shown how it can be constructed and measures are proposed to characterize the relational strength among different environments (discovery oriented research, applications-oriented research and patents) and the speed of evolution. Our approach is illustrated by a case study of three fundamental restriction enzymes articles. Among other results we found that 387 patents cited 124 of the 988 articles (a share of 12.55%) in the mediator set. Defining the non-patent references (NPR) transition rate as the number of citing patents divided by the number of articles in the mediator set yields a value 0.392. Our results suggest that the D-A-T path acts as a backbone and reveals important "invisible contributions" of an original scientific work during its evolution from discovery oriented research to outside academia. Our contribution provides a useful tool for bridging the existing gap in detecting the transition of knowledge between science and technology.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

The evolution of science and technology can be seen as a process of knowledge transition between fields, from basic science to applied research or vice versa (Glänzel & Meyer, 2003). Price even stated that ".. . historically the arrow of causality is largely from the technology to the science" (de Solla Price, 1983). New basic knowledge may lead to a knowledge flow to applied fields and technology, and, in this way, contributes to progress in industry and society (Li, Azoulay, & Sampat, 2017).Of course, sometimes knowledge may also flow from technology to new data, leading to new questions and when these are solved, to new basic knowledge, astronomical instruments, such as the Hubble telescope, being a case in point (http://hubblesite.org/). Knowledge flows manifest themselves sometimes as successions of incremental improvements. Yet, it also happens that progress is temporarily stopped and then revived during periods of revolutionary transitions (Pan, Sinha, Kaski, & Saramäki, 2012). Such knowledge flows can be tracked by roadmaps, showing interactive movements and

---

* Corresponding author at: Facultair Onderzoekscentrum ECOOM, KU Leuven, Naamsestraat 61, Leuven, B-3000, Belgium.
E-mail addresses: xjhu@zju.edu.cn (X. Hu), ronald.rousseau@kuleuven.be, ronald.rousseau@uantwerpen.be (R. Rousseau).

illustrated by evolutional networks between science and technology or vice versa (Boyack, Klavans, & Börner, 2005; Boyack, Klavans, Small, & Ungar, 2014; Li et al., 2017; Pan et al., 2012).

The study of scientific evolution is performed in many ways. Among these we mention:

(1) Determining knowledge diffusion among scientific fields by citations over time (Frandsen, 2004; Frandsen, Rousseau, & Rowlands, 2006; Lewison, Rippon, & Wooding, 2005; Liu & Rousseau, 2010; Rowlands, 2002). Such an approach stays largely within academia and can be performed using databases such as the Web of Science (WoS), Scopus or PubMed.
(2) Measuring knowledge transfer between science and technology through the interaction of scientific articles and patents (Gao, Ding, Teng, & Pang, 2012; Li et al., 2017; Meyer, 2006). This approach goes one step further and moves to the realm of technological applications. These studies are often performed using a patent database, but, to provide a full picture they also need data related to academic research.
(3) When focusing on medicine and the health sciences one often uses the term translational research. The objective of translational research is to harness knowledge from basic research (biomedical) to clinical practice and the development of new drugs and treatments, leading to better patient care (Luwel & van Wijk, 2015).

The former two in this list are studied at least since the 1960s (Sherwin & Isenson, 1967), while the third one dates from somewhat later (the 1990s) (Seiken, 1993). The notions "translational applications" and "translational effect" are hot topics nowadays (leading to specialized journals such as *Science Translational Medicine*, since 2009). Yet, sensible metrics for them are still on their way (Luwel & van Wijk, 2015).

Knowledge transfer is often seen as a largely linear evolution of scientific or technical knowledge through time. In both cases the original piece of knowledge may need to be transformed or adapted to be useful under new circumstances. Knowledge transfer from the basic sciences to technical applications may, e.g., result from contract research between academia and industry, via patenting or via spin off activity. Yet, knowledge transfer may also occur between different fields of science such as from the natural sciences to the social sciences, or from subfields of physics to subfields of chemistry. The term transition is used in the case of a link between different types of fields such as basic and applied fields, as defined by Tijssen (2010). This is explained in more detail further on.

One may observe that among the two major types of research on knowledge diffusion and transfer, the former one maps knowledge transfer within the realm of science according to the records in a science citation database, while the latter type explores the links in particular fields between scientific knowledge and technological inventions via a patent database. Most of them applied a direct citation analysis (Hall, Jaffe, & Trajtenberg, 2005) but this may provide an incomplete picture (Hu, Rousseau & Chen, 2011, 2012). For this reason Kuan, Huang, and Chen, (2018) included bibliographic coupling in their patent analyses. Still there may be many "invisible elements" that stay largely unnoticed as a contribution to science and technology. Tracking the influence of pioneering work during scientific evolution has only recently been investigated seriously, e.g. by the use of Reference Publication Year Spectroscopy (RPYS) (Marx, Bornmann, Barth, & Leydesdorff, 2014; Marx, Bornmann, & Barth, 2013) or by trying to detect so-called under-cited influential articles (Hu & Rousseau, 2016, 2017, 2018).

The article that is the most related to our investigation is Ahmadpoor and Jones' (2017). These authors study the whole articles-patents citation network, starting from patents that directly cite journal articles. They refer to such citations as the "paper-patent boundary". Then they determine for all other papers and patents the minimum citation distance to this boundary. One of their main findings is that a majority of patents (60.5%) is connected (directly or indirectly) to science and engineering papers, while a substantial majority of articles (79.7%) can be linked to patents(also directly or indirectly). They further find that some patent classes such as molecular biology, superconductivity technology and combinatorial chemistry are close to the paper-patent boundary while others, such as locks, buttons, fasteners, chairs are situated at a much larger distance. On the other side of the boundary (papers) they found nanoscience and nanotechnology, materials science and computer science hardware & architecture, while mathematics is situated very far away.

Yet, we think that much more can be done in studying knowledge transfer. Traditionally the evolution of knowledge has mostly been studied either within a scientific context or from a technological point of view. As it is clear that applied sciences provide the (missing?) link between basic science and technology the problem lies in the difficulty in operationalizing the notion of applied science as distinct from basic science.

Detailed knowledge of the backbone of knowledge transfer via a successive process from basic research to technological fields is still lacking. In this article we try to contribute to this aspect of the evolution of science and technology.

We propose a new approach to detect the otherwise "invisible" or at least partially hidden, path of knowledge from an original scientific work during its evolution from science to technology. The term 'invisible' is used here because only studying non-patent-references (NPRs) does not lead to the origin of the used knowledge. This statement implies that we focus on knowledge transfer via documents and this from science to technology, considering other forms of knowledge transfer or of knowledge diffusion out of the scope of this contribution, i.e. we do not consider knowledge transfer via more meandering paths in Fig. 1 or via social media or word of mouth. Particularly, we focus on different environments of knowledge utilization in the process of knowledge transfer via citation networks, hence including indirect citations, focusing on an applications-oriented evolution. Concretely, we will try to find answers to the following questions:

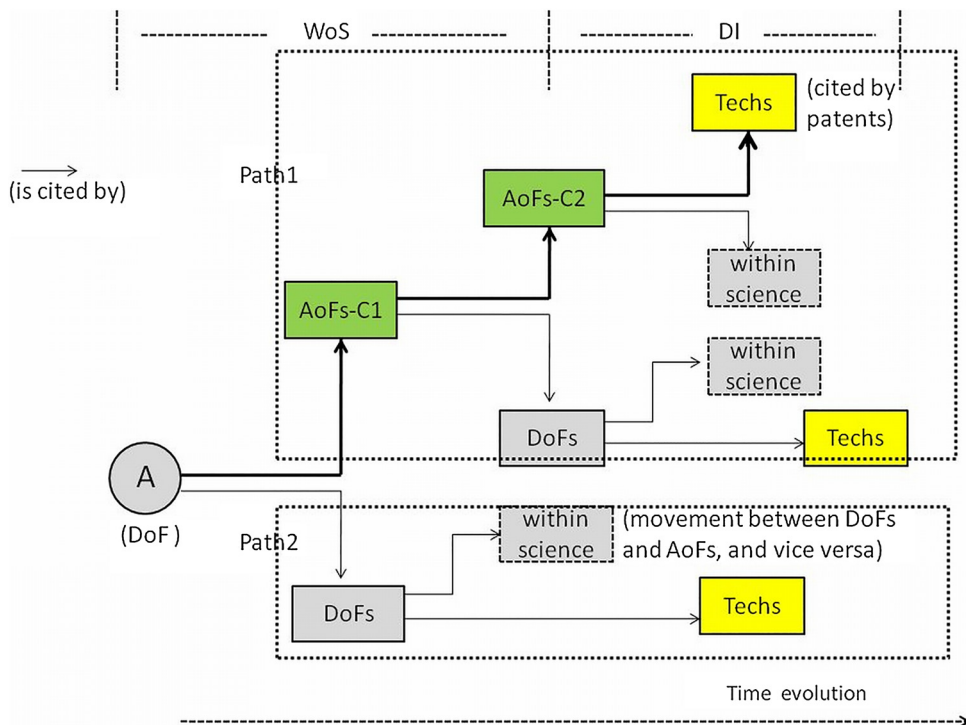(1) How to construct a new framework to detect transfer of knowledge from basic science to technology?

Fig. 1. Roadmap showing the different states of knowledge utilization during a D-A-T evolution.

(2) Is it feasible to show the main features of the movement of the knowledge included in an original scientific work during its evolution towards possible industrial application? Can this be done in another way rather than by tracking direct citations of basic science in patents?

(3) How to construct and apply indicators to characterize the relational strength among different environments (discovery oriented research, application oriented research and patents)?

(4) How to measure the speed of evolution?

2. A short literature review

Citation-based diffusion is one of the basic approaches to study scientific evolution (Frandsen, 2004; Frandsen et al., 2006; Lewison et al., 2005; Liu & Rousseau, 2010; Rowlands, 2002). In this context, specific indicators for measuring diffusion have been constructed or adapted from other uses, such as measuring interdisciplinarity (Liu, Rafols, & Rousseau, 2012). In Liu and Rousseau (2010), the authors proposed the indicators "field diffusion breadth" (*FDB*) to define the number of Essential Science Indicators (ESI)fields in which a set of articles (which may be a singleton set) is cited, and "field diffusion intensity" (*DINT*), defined as the number of citing articles in one particular *ESI* field. These indicators provide useful instruments to determine knowledge transfer across fields.

To study the connection between science and technology, Narin and his colleagues published a series of fundamental articles (Narin & Hamilton, 1996; Narin & Noma, 1985; Narin & Olivastro, 1998). They used the age of a scientific paper as an indicator to define the time lag between the year of publication and the year an article is first cited in a patent.

Following the lead of Narin and his colleagues, the interaction of knowledge transfer between science and technology became a hot topic in the latest decades (Gao et al., 2012; Meyer, 2006; Van Looy et al., 2011; Li et al., 2017). Research often focused on the links between science and particular technological fields (Gao et al., 2012; Meyer, 2006), and tried to uncover the underlying mechanism of knowledge transfer (Sudhindra, Ganesh, & Arshinder, 2017; van Raan, 2017). These colleagues used e.g., co-citation analysis to map scientific knowledge sources and technology fronts (Gao et al., 2012). In order to observe the characteristics of Sleeping Beauties (SBs) van Raan (2017) developed an approach in different steps to explore the cognitive environment of SBs cited in patents, using the EPO Worldwide Patent Statistical Database (PATSTAT). He used "the first year of citation in a patent" to calculate the time lag between publication and patent citation. Based on the references of articles cited in patents (second generation of backward citations), Gherardini and Nucciotti (2017) mapped the invisible colleges on which "Knowledge Transfer" as a discipline is based. They pointed out that although the entire scientific domain has a strongly connected - international - dimension, it still manifests a persistent fragmentation too. With the rise in the use of the term "knowledge transfer", researchers became aware that a lack of consensus about the conceptualization of knowledge transfer has resulted in a diversity of terms being used to describe this phenomenon. Some authors conceive

knowledge transfer as a process, others describe it as an activity, an approach, an interaction, or a transformation (Gervais, Marion, Dagenais, Chiocchio, & Houlfort, 2016), leading to possible misunderstandings related to the term and meaning of knowledge transfer in the literature (Gervais et al., 2016; Tangaraja, Rasdi, Samah, & Ismail, 2016). We see knowledge transfer mainly as a process.

3. Methodology

In this section, we describe a new approach to explore the "invisible" influence of an original scientific work in applications-oriented knowledge flows over time. As the term 'invisible' refers to the fact that the total influence of a discovery in basic science cannot be found when studying patents and their references alone, we do more than just taking the crossover from sciences to technology into account. Concretely, we take the following steps.

3.1. Theoretical principle

The evolution of science and technology is an ongoing process of knowledge transition. Hence, the notions *environment of knowledge utilization* and *applications-oriented* research as proposed by Tijssen (2010) and further elaborated in (Tijssen & Winnink, 2016) provide the main guideline for this investigation.

3.2. Applications-oriented evolution

Based on the notions of *environment of knowledge utilization* and *applications-oriented* proposed by Tijssen (2010), we develop a roadmap reflecting an applications-oriented evolution to detect different states of knowledge utilization over time. As a first step we identify discovery-oriented science and applications-oriented science, see sections 3.3 and 3.4.

3.3. Journal Application Domain (JAD) categories

Based on a "Knowledge Utilization Triangle" Tijssen (2010) proposed three types of coexisting knowledge application domains: clinical, industrial and civic. Then he continued by developing a system for classifying scholarly journals according to their degree of application orientation. Employing this system, all journals indexed by the Web of Science (WoS) were classified into six JAD categories: academic journals; industry relevant journals; industry practice journals: clinical relevant journals; clinical practice journals and industry-clinical relevant journals. For details about this classification we refer the reader to (Tijssen, 2010).

3.4. The notions of discovery-oriented and applications-oriented science

(a) Discovery-oriented science: When using the "Knowledge Utilization Triangle" Tijssen (2010), p. 1844) pointed out that application domains are without clear-cut boundaries. Journals residing in the center of this triangle are seen as 'discovery oriented', with academic contributors employed by universities or other public sector research-performing organizations. He also pointed out that the majority of WoS-indexed journals have an academic authorship (Tijssen, 2010, p. 1845), strongly suggesting to us that the journals fed from discovery-oriented research environments should be classified as discovery-oriented science. All fields that belong to discovery-oriented science are indicated as discovery-oriented fields, DoFs in short.

(b) Application-oriented science: Each journal in the other five JAD categories can be classified and ranked according to their degree of relevance in terms of application orientation (Tijssen, 2010). From this, Tijssen determined the Top 25 Application-Oriented Fields (AoFs) in science, consisting of the Top 10 Industrial-oriented, the Top 10 Clinical practice-oriented, and the Top 5 Industrial /clinical-oriented ones. These fields, i.e. WoS Subject Categories, are shown in Appendix A.

Tijssen (2010) also found that there are only 31 fields with 'non-patent references'(NPR) scores above 0.25% in WoS indexed journals. Here a field's NPR score is defined as the number of NPR to publications (in journals) in this WoS Subject Category divided the total number of publications in this Category (using a 7 years' publication window - all types of publications - in the WoS and the same citation window in PATSTAT patents).Following Tijssen (2010) we think that an applications-oriented research environment is a sufficiently robust proxy for identifying applications-oriented science. In order to apply the JAD categorization in a cautious way, we operationalize the notion of Application-Oriented Fields (AoFs) as those WoS Subject Categories belonging to the 25 AoFs pinpointed in (Tijssen, 2010). Hence, from now on the term AoF will always refer to one of those 25 Web of Science Subject Categories (Tijssen, 2010) and shown in Appendix A.

3.5. Path construction

In this construction we follow the direction: discovery-oriented fields (DoFs) → applications-oriented fields (AoFs) → technological areas (Techs). We will use the general term *environment* to denote any of these three fields or areas. This line

of thought will be referred to as the D-A-T evolution. In order to come to the path we want to construct, we first have a closer look at possible paths in citation networks in science and between science and technology.

(a) Path identification: As shown in Fig. 1, two paths (Path 1 and Path 2) of knowledge transfer from an article in a DoF to technological areas (Techs) may exist during an applications-oriented evolution. Path 2 represents the conventional roadmap to study the evolutional process from basic research to a technological field. Yet, according to Tijssen (2010, p. 1848), "the vast majority of the WoS-indexed journals will never be cited within patents, especially the 'basic' journals within fields of science that are disconnected from technological development". Hence, low transition rates are expected from DoFs to Techs when one tracks the relationships between science and technology via Path 2. One may also note that there is a reverse sub-path included in Path 1 when going up the second generation (AoFs to DoFs and further to Techs). Such a sub-path can occur, but as described above, it is expected to play a minor role in an applications-oriented evolution. A direct AoFs-to-DoFs path is actually a reverse movement in the D-A-T evolution. In view of our study of the applications-oriented evolution of knowledge, Path 1 can be said to be the backbone to track the forward movement of knowledge diffusion and transition: it evolves as a successive process from DoFs to AoFs, and from AoFs to Techs.

(b) Finding the main path of a D-A-T evolution: In order to come to the path we want to construct, we need a database that covers academic and technological/industrial documents in a sufficient manner. For this we combine data from the Web of Science (WoS) with Derwent Innovation (DI), which is based on the Derwent World Patents Index (DWPI). Hence citations will always be collected from one of these two databases. A reviewer observed that we could have used Dimensions (www.dimensions.ai) which combines publications and patent data. Finding out pros and cons of these two (and possibly other) options would, however, lead us too far. Following the D-A-T evolution roadmap (Fig. 1), we construct a citation path starting from the original scientific publication A to technological areas via indirect citation relations mined from the WoS and Derwent Innovation (DI). To map "invisible elements" in the process of evolution, we look up the citations of article A in the first forward generation and identify the citing articles in AoFs (denoted as AoFs-C1). If a citing article appears in a journal that is classified in one of these AoFs and also in another field (any, not necessarily application oriented) then the citing article is included. Next we track the citations of AoFs-C1 up to A's second forward generation and identify AoFs articles among them (denoted as AoFs-C2). Finally, we form the union of AoFs-C1 and AoFs-C2, denoted as AoFs-C. Note that in the terminology of (Hu, Rousseau, & Chen, 2011) we work with sets (not multisets), and by taking the union it does not matter if we work with overlapping or non-overlapping generations. The set AoFs-C will be called A's mediator set. Further on we will apply this approach to a set of articles, instead of just one (such as A). If this set is set S we will call the corresponding AoFs-C set S's mediator set (or mediator in short, when it is clear about which set S we are talking). Next we inspect all articles in the mediator to see if they are cited in patents (as covered by DI). We choose DI here because DI provides access to global patent data. Moreover, non-patent references cited by inventors as well as examiners are included in DI (Clarivate, 2017 https://clarivate.com/products/derwent-innovation/). Again we refer the reader to Fig.1.

3.6. Indicators to determine the D-A-T evolution

We develop several indicators to analyze the characteristics of evolution: some are related to the outcome dimension and some to the time dimension. We recall that the Top 25 AoFs are based on WoS categories (Tijssen, 2010), which is an overlapping system. This implies that a citing article in the mediator may belong to more than one AoF, leading to a rather fuzzy picture. For this reason we also include ESI fields, which are non-overlapping. Concretely we determine the ESI field of each article in the mediator. We recall that ESI fields have already been used for mapping knowledge diffusion, see e.g., (Liu & Rousseau, 2010).

In this way, each article in the mediator is classified according to two schemes: by AoFs and by ESI fields, and hence belongs to two sets: an AoF and an ESI-field. Note that articles in one ESI field may belong to different AoFs. Although one article may only belong to one ESI field, it is possible that one AoF contains articles belonging to different ESI fields. When counting the number of articles in a set (an AoF, an ESI-field or a mediator) we use whole counting.

Next we come to the definition and description of the actual indicators. The first two are structure related.

(a) The relational strength (RS): Each article in a mediator belongs to one ESI field and to one or more of the AoFs from Tijssen's list. To characterize the relation between these ESI fields and the AoFs, we define the following relational strength indicators.

The relational strength between an ESI field E(i) and an AoF A(j) is defined as

$$RS_{ij} = \frac{\#(E(i) \cap A(j) \cap Mediator\,Set)}{\#(Mediator\,Set)} \tag{1}$$

where # means "number of elements in". The numerator denotes the number of articles in the mediator belonging to E(i) and to A(j): it reflects a co-occurrence relation. The denominator is equal to the total number of articles in the mediator. Note that, because we use whole counting and some articles belong to more than one AoF,

$$\sum_{i,j} RS_{ij} \geq 1$$

Similarly, we define the relational strength of an ESI-field E(i) with respect to the mediator or the relational strength of an AoF A(j) with respect to a mediator as:

$$RS(i) = \frac{\#(E(i) \cap Mediator\,Set)}{\#(Mediator\,Set)} \tag{2}$$

$$RS(j) = \frac{\#(A(j) \cap Mediator\,Set)}{\#(Mediator\,Set)} \tag{3}$$

Analogously to the observation made above we note that $\sum_i RS(i) = 1$ and $\sum_j RS(j) \geq 1$. In Appendix B we provide an illustration of these two relations. Other relations between indicators follow similarly.

(b) Relative strength of knowledge transition (*SKT*): Let P be the union of all patents that cite at least one article in the corresponding mediator under study. Here we use the following weighting of citations: if a patent citing an article in the mediator set, belongs to two (or more) patent classes this citation is counted twice or more; similarly, if the cited article belongs to two (or more) AoFs it is counted for each AoFs (whole counting). The sum of all weighted citations between P and the mediator set is denoted as M. If P(k) is one patent class and A(j) is one AoF field then the relative strength of knowledge transition between P(k) and A(j) is defined as:

$$SKT_{kj} = \frac{C_{kj}}{M} \tag{4}$$

where C_{kj} denotes the number of (weighted) citations from patents in class P(k) to articles in field A(j), and, of course, belonging to the mediator. Because of the counting method we have:

$$\sum_{k,j} SKT_{kj} = 1$$

Similarly, the relative strength of knowledge transition of a patent class P(k) is:

$$SKT_k = \frac{C_k}{M} \tag{5}$$

where C_k denotes the number of citations from patents in P(k) to any article in the mediator; analogously we obtain:

$$SKT_j = \frac{C_j}{M} \tag{6}$$

SKT_j denotes the relative strength of knowledge transition of the AoF field A(j). The symbol Cj denotes the total number of citations received by articles in the mediator that belong to field A(j), coming from any patent class.

(c) The first transition time (*FT*): The number of years it takes an article in a DoF to receive its first citation from an AoF. *FT* is set to zero if the first citation is received in the article's publication year. This indicator is a variant on the first-citation idea as introduced in (Rousseau, 1994).

(d) The age of transition (*AGT*): The average number of years for a set of articles in Environment E to receive citations from Environment F in the applications-oriented evolution (averaged over all citations and all articles in the set), reflecting the time needed to reach a phase transition, either from a DoF to an AoF (denoted as *AGT-DA*) or from an AoF to a patent (Techs) (denoted as *AGT-AT*).

(e) The article time-span of D-A-T transition (*TS-DAT*): The time-span of a discovery-oriented (DoF) article in the D-A-T transition is defined as the difference in years between the publication of a DoF article and a citation in a patent, via a cited (by the patent) and a citing AoF article (of the DoF article).

(f) In order to track the process of knowledge transition during the D-A-T evolution, we apply the indicators *FDB* and *DINT* proposed by Liu and Rousseau (2010) as defined earlier. Recall that these indicators were introduced in the context of ESI-fields.

(g) The cumulative transition rate between DoF and AoF: this indicator is defined as the ratio of the cumulative number of citations received by a set of DoF articles from articles in the AoF set, over the cumulative total number of received citations by this same set. As this ratio can be obtained each year it is a yearly cumulative transition rate.

(h) Finally, the NPR transition rate is defined as the number of citing patents divided by the number of articles in the mediator set.

The age of transition (*AGT*), *FDB*, *DINT*, the cumulative transition rate and the NPR transition rate, are indicators for sets of articles (which can in principle also be applied to a singleton set), while the first transition time (*FT*) and the time-span of the D-A-T transition (*TS-DAT*) are article indicators.

4. A case study on restriction enzymes in biotechnology

It is well known that restriction enzymes laid the foundation for the physical mapping of DNA: one can rightly say that the biotechnology industry would not have flourished without them (Roberts, 2005). For this reason we use three pioneering

Table 1

The diffusion breadth (*FDB*) and intensity (*DINT*) of the A-S-N set (first generation citations) over time, based on ESI categories (cumulative data).

	DINT																FDB
	AS	BB	CH	CM	CS	ENG	EN/EC	IM	MICRO	MBG	MUL	NB	PT	PAS	PP	SS	
Y+9	2	171	4	15	0	1	0	5	197	103	70	0	4	3	0	0	11
Y+19	3	213	9	25	0	1	1	8	229	125	82	3	11	6	1	0	14
Y+29	3	230	11	42	1	2	1	9	241	135	91	3	14	8	1	2	16
Y+39	4	250	12	47	5	3	2	13	249	142	96	3	16	9	1	4	16
Y+48	4	270	22	53	6	3	3	18	254	149	103	4	17	11	1	6	16

AS = Agricultural sciences; BB = Biology &Biochemistry; CH = Chemistry; CM = Clinical Medicine; CS = Computer Science; ENG = Engineering; EN/EC = Environment/Ecology; IM = Immunology; MICRO = Microbiology; MBG = Molecular Biology & Genetics; MUL = Multidisciplinary; NB = Neuroscience & Behavior; PT = Pharmacology & Toxicology; PAS = Plant & Animal Science; PP = Psychiatry/Psychology; SS = Social Sciences, General.

basic articles on restriction enzymes as the topic of our investigation of the transition between discovery-oriented sciences, via application oriented fields to patents to be used in technological applications.

A famous trio of articles: Arber and Linn (1969); Smith and Wilcox (1970) and Danna and Nathans (1971) provided the foundation for the understanding and application of restriction enzymes. These three articles all belong to DoFs: Arber and Linn (1969) as well as Smith and Wilcox (1970) are published in the category Biochemistry & Molecular Biology (the first one is nowadays included in the ESI field: Biology & Biochemistry; the second one in the ESI-field Molecular Biology & Genetics), while Danna and Nathans (1971) is classified as Multidisciplinary Sciences (ESI field: Multidisciplinary). In these articles the authors provided a theoretical framework describing the biology of restriction enzymes and its application to problems in molecular genetics. More specifically, Arber and Linn (1969) were the first to successfully isolate a so-called type I restriction enzyme. In the next year, Smith discovered the first type II restriction enzyme (Smith & Wilcox, 1970). Yet it was Nathans who set the stage for the use of restriction enzymes for mapping and manipulating DNA in different scientific applications (Danna & Nathans, 1971). For their work on the discovery and characterization of restriction enzymes Arber, Nathans and Smith received the 1978 Nobel Prize for Physiology or Medicine (www.nobelprize.org).

Next, we describe the details of our case study: the research objects, the data extraction method and the measures used to describe the D-A-T evolution. Data collection was performed in December 2017.

(1) We start from the three original DoF articles (Arber & Linn, 1969; Danna & Nathans, 1971; Smith & Wilcox, 1970), denoted as the Arber-Smith-Nathans set (A-S-N set, in short), and determine the articles (in the WoS) that cite at least one of the articles in the A-S-N set. The set of these citing articles is denoted as C1.
(2) We developed a program to determine all articles in C1 that belong to journals classified in the Top 25 AoFs (based on WoS Subject Categories),as determined by Tijssen (2010). If a journal is classified in one of the categories in the list of Top 25 AoFs and also in another field (any, not necessarily application oriented) then its articles are included. This set of articles is denoted as AoFs-C1.
(3) Next, we determine all articles (again in the WoS) that cite at least one of the articles in AoFs-C1. This set is denoted as C2.
(4) We repeat step (2) for the articles in C2, leading to the set AoFs-C2.
(5) We form the union of AoFs-C1 and AoFs-C2, hence removing duplicates. This set of articles is called the A-S-N mediator.
(6) For each article in the A-S-N mediator, we also determine the ESI field to which it belongs. This step is done automatically, using another software program developed by ourselves. We recall that ESI fields are non-overlapping.
     In this way we collected all articles published in an AoF that cited at least one of the articles in the A-S-N set directly or indirectly (two forward citation generations). Next, we construct a matrix of relationships between AoFs and ESI fields (based on articles) to determine relational strengths (*RS*) as defined in the method section.
(7) Finally, we determine all patents (from DI) that cite at least one of the articles in the A-S-N mediator set. In this way we obtain for each article in the mediator the number of patent citations, the corresponding International Patent Classification codes (IPC codes), and the application year for each patent. Then, we construct a matrix of relationships between AoFs and IPC-3 codes (based on their citation links). Here, only different IPC-3 codes are counted.
(8) At last, we calculate the values of the *FDB, DINT, RS*, *SKT, FT, AGT*, *TS-DAT* indicators and transition rates according to the definitions and formulas described above.

5. Results

5.1. Direct knowledge transfer in the first citation generation

By the end of December 2017 the three articles of interest were directly cited by 1152 different articles. Among these 924 (80.2%) originate from outside fields, i.e. ESI fields different from those these three articles belong to. Table 1 shows the field diffusion breadth (*FDB*) and the field diffusion intensity (*DINT*) based on ESI categories over time. Here, year Y refers to the publication year of an article. As shown in Table 1, the (ESI) field diffusion breadth (*FDB*) is 16 covering more than 2/3 of

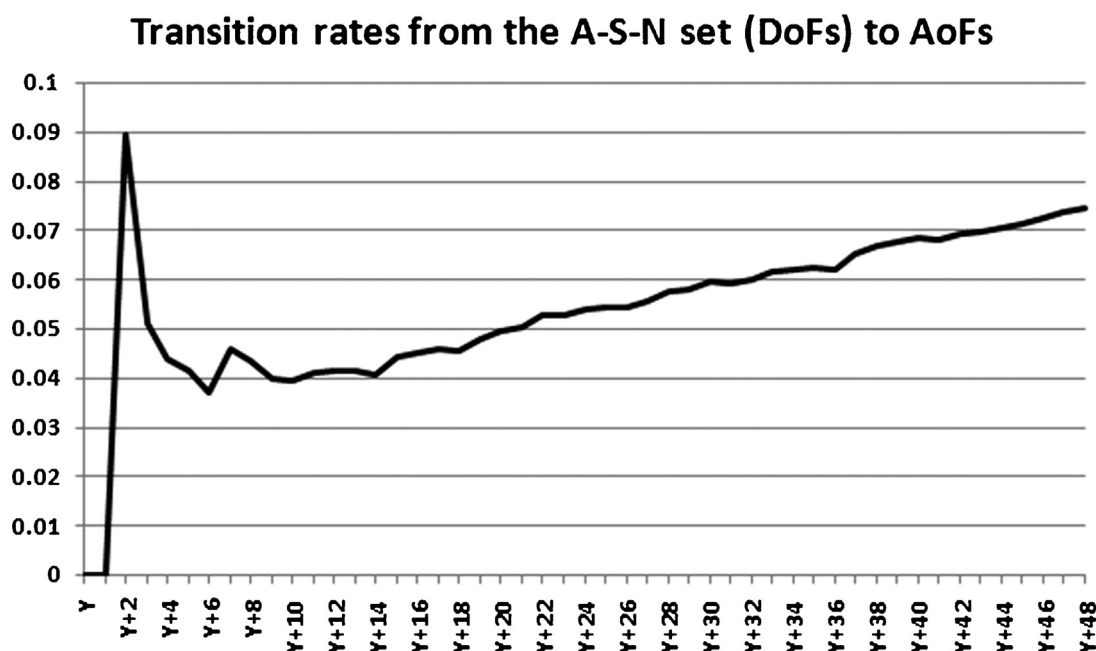## Transition rates from the A-S-N set (DoFs) to AoFs

Fig. 2. Yearly cumulative transition rates from the A-S-N set (DoFs) to AoFs in the first citation generation (cumulative data); Year Y refers to the publication year of an article.

the ESI fields. Biology & Biochemistry, Microbiology, Molecular Biology & Genetics, and Multidisciplinary are the fields with the highest *DINT*-values, reflecting that direct diffusion occurred mainly between fields close to those of these three articles.

Among those 1152 citing articles only 86 belong to an AoF determined by Tijssen. More precisely these first citation generation (AoFs-C1) articles are distributed over 11 of the 25 AoFs. These are Biochemical Research Methods, Medicinal Chemistry, Energy & Fuels, Medical Laboratory Technology, General & Internal Medicine, Legal Medicine, Nanoscience & Nanotechnology, Pediatrics, Pharmacology & Pharmacy, Surgery, and Toxicology. They belong to 8 ESI fields. Fig.2 illustrates the evolution of the yearly cumulative knowledge transition rate from the A-S-N set (DoFs) to AoFs. Clearly, for the A-S-N set knowledge transition through direct citation is rather low.

5.2. The backbone of knowledge-transition of the A-S-N set and the relationships between ESI fields and AoFs in the mediator set

Among 3870 articles citing articles in AoFs-C1 we tracked 913 articles in AoFs (or 23.6%). They form AoFs-C2. Combined with the items in AoFs-C1, we find 988 articles forming A-S-N's mediator set. Articles in this mediator set belong to 17 different AoFs; 6 new AoFs Critical Care Medicine, Engineering, Electrical & Electronic, Gastroenterology & Hepatology, Materials Science, Biomaterials, Rheumatology, and Urology & Nephrology have been added in the second generation data.

5.2.1. The roadmap of AoF transition over time
Fig.3 displays transitions from the A-S-N set (the dot in the middle) into AoFs per five-year period. This map has two important characteristics. First, we see that the number of AoFs and the strength of lines are growing over time: even in the latest five years, the A-S-N set is linked to eight AoFs. This reveals that the indirect influence of the three DoF articles under study – although almost 50-years old – is still very important in the applications-oriented evolution. Second, we note that during the early years the main broad lines connecting to AoFs are related to clinical fields, a situation that has changed in the period 2004-2008. From then biochemical research methods became the most active AoFs during the movement of knowledge utilization. In this same period also Nanoscience & Nanotechnology emerged, revealing that the A-S-N set led to new applications in biotechnological fields. Although there are no new AoFs in the latest period, there was one in the period 2009–2013, namely Materials Science, Biomaterials.

5.2.2. Relationships between ESI fields and AoFs
To reveal relative dependency relationships between ESI fields and AoFs, the relational strength (*RS*) between an ESI field and an AoF, the *RS* of an ESI field, and the *RS* of an AoF are calculated based on their co-occurrence matrix (within the mediator set) and formula (1). Results are shown in Table C1 (Appendix C). In order to show the relationships in a clearer way, we use abbreviations in the tables and figures below. Full names of AoFs are in the legend of Table C1, while these of the ESI fields were already given in Table 1.

Fig. 3. The roadmap of AoFs transitions in the mediator set per 5-year period.

The *RS* values for a whole field using formulae (2) and (3) are shown by the row and column with heading *Total*. As shown by these values ESI fields Biology & Biochemistry, Clinical Medicine, Pharmacology & Toxicology, play major roles with respective *RS* values of 0.362, 0.336 and 0.134. The last row illustrates that AoFs Biochemical Research Methods, General & Internal Medicine, Toxicology, Pharmacology & Pharmacy are main actors in the phase transition with respective *RS* values of 0.423, 0.179, 0.145 and 0.123.

The value of *RS* in each cell reflects the relational strength between an ESI-field and an AoF. In this way we can easily see that the strongest links are: Biology & Biochemistry to Biochemical Research Methods (with *RS* value equal to 0.319); Pharmacology & Toxicology to Toxicology (0.163); Clinical Medicine to General & Internal Medicine (0.142); Pharmacology
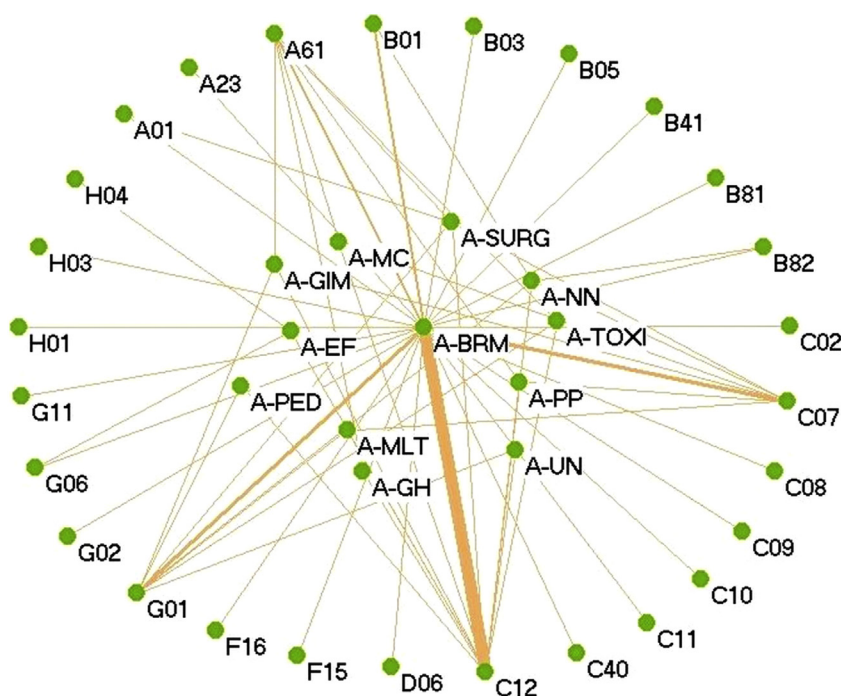
Fig. 4. Transition relationships between AoFs and technological classes.

& Toxicology to Pharmacology & Pharmacy (0.106); Clinical Medicine to Pediatrics (0.068) and Clinical Medicine to Surgery (0.046). They reflect the relational strength in the evolution between ESI fields and AoFs.

5.3. Transition relationships between AoFs and technological classes

We obtained the following key indicator values of the A-S-N set: 387 patents were tracked citing 124 of the 988 articles (a share of 12.55%) in the mediator set. We recall that this set is a union of AoFs-C1 and AoFs-C2, hence duplicates have been removed. Defining the NPR transition rate as the number of citing patents divided by the number of articles in the mediator set yields a value of 387/988 = 0.392.This percentage reflects a huge "invisible contribution" of the three fundamental restriction enzyme articles to technological fields. We use the term 'invisible' because only studying NPRs does not lead to the origin of the used knowledge, in this case the A-S-N set.

Table 2 and Fig.4 display the strength of relationships between AoFs and the IPC-3 of patents, and the detailed values (based on formula (4)) in the phase transition of knowledge utilization from applications-oriented fields to technological classes. These codes cover seven industrial sections from A to H, section E "Fixed Constructions" (WIPO, 2017) being the exception. These results reflect the fact that knowledge of the A-S-N set has transferred into most technological sections. More precisely, we should state that it has the potential of being transferred into technology based industries, as we are aware that being granted a patent is not at all the same as being actually applied in an industrial process.

The *SKT* value shown in the last row of Table 2 indicates the relative strength of knowledge transition of an AoF field to all patent classes included in the set of patents (WIPO, 2017). Biochemical Research Methods plays the leading role in this movement, with a value as high as 0.845.The AoF with the smallest value shown in the last low of Table 3 is Gastroenterology & Hepatology, with a value of 0.003. To put this value in perspective we recall that there are only 31 WoS fields with NPR scores above 0.25% (Tijssen, 2010). From the values of *SKT* displayed in the last column, we understand that the technological class C12 (Biochemistry; Beer; Spirits; Wine; Vinegar; Microbiology; Enzymology; Mutation or Genetic engineering) (WIPO, 2017), has the largest strength (0.426) in intake from related AoFs.

The *SKT* values in the matrix reflect the relative strength of knowledge transition between an AoF and a technological class. Clearly, many technological classes are highly dependent on the field of Biochemical Research Methods, such as C12 (Biochemistry etc.), C07 (Organic chemistry), G01 (Measuring; Testing), A61 (Medical or veterinary science; Hygiene), B01 (Physical or chemical processes or apparatus in general) (WIPO, 2017). Values are respectively 0.373, 0.113, 0.111, 0.084, and 0.074. One may further observe that A61, C07, and C12 too are closely related to the field General & Internal Medicine, with respective values of 0.023, 0.016, and 0.016.

Table 2

The strength of knowledge-transition between AoFs and Technological classes (*SKT*-values). Row and column totals and highest cell values are indicated in grey.

AoFs* IPCs	A-BRM	A-MC	A-EF	A-GH	A-MLT	A-GIM	A-NN	A-PED	A-PP	A-SURG	A-TOXI	A-UN	Total
A01	0.0002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.0003
A23	0.0002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0002
A61	0.0084	0.002	0.000	0.002	0.000	0.0023	0.000	0.000	0.0013	0.003	0.002	0.000	0.127
B01	0.0074	0.000	0.000	0.000	0.000	0.000	0.0002	0.000	0.000	0.000	0.000	0.000	0.076
B03	0.0002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0002
B05	0.0002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0002
B41	0.0002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0002
B81	0.0005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0005
B82	0.0003	0.000	0.000	0.000	0.000	0.000	0.0002	0.000	0.000	0.000	0.000	0.000	0.0005
C02	0.0002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0002
C07	0.113	0.002	0.000	0.000	0.003	0.0016	0.002	0.000	0.003	0.003	0.000	0.000	0.142
C08	0.0005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0005
C09	0.0002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0002
C10	0.0003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0003
C11	0.0008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0008
C12	0.373	0.002	0.005	0.002	0.011	0.0016	0.002	0.002	0.006	0.002	0.002	0.005	0.426
C40	0.0021	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.021
D06	0.0002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0002
F15	0.0003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0003
F16	0.0002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0002
G01	0.111	0.000	0.000	0.000	0.005	0.002	0.003	0.003	0.000	0.002	0.002	0.005	0.132
G02	0.0002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0002
G06	0.0013	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.015
G11	0.0003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0003
H01	0.0008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0008
H03	0.0002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0002
H04	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0002
Total	0.845	0.005	0.008	0.003	0.019	0.056	0.010	0.005	0.023	0.011	0.005	0.010	1.000

*A-BRM = Biochemical Research Methods; A-MC = Chemistry, Medicinal; A-EF = Energy & Fuels; A-GH = Gastroenterology & Hepatology; A-MLT = Medical Laboratory Technology; A-GIM = Medicine, General & Internal; A-NN = Nanoscience & Nanotechnology; A-PED = Pediatrics; A-PP = Pharmacology & Pharmacy; A-SURG = Surgery; A-TOXI = Toxicology; A-UN = Urology & Nephrology.

5.4. The speed of knowledge transfer in the D-A-T evolution from the A-S-N set

In this section, we discuss the results of time-related indicators.

5.4.1. The transition speed and the age of transition of articles from the A-S-N set to its mediator set

Considering the first transition time (*FT*), we see that the fastest movement from the A-S-N set to an AoF occurred for Biochemical Research Methods, General & Internal Medicine, and Pharmacology & Pharmacy, with an *FT*-value of two years. However, the average age of articles in most AoFs is more than 30 years, reflecting that the indirect influence of the trio

Table 3
From the A-S-N set to its mediator set: first transition time and the average age of citing articles.

AoFs[a]	FT	AGT-DA	Received citations
A-BRM	2	37.32	418
A-MC	4	38.67	39
A-CCM	29	32	2
A-EF	34	44.25.	32
A-EEE	37	40	3
A-GH	14	31.37	19
A-MSB	43	44.25	4
A-MLT	3	27.64	42
A-GIM	2	19.48	143
A-LG	17	21.33	3
A-NN	35	43.05	17
A-PED	7	20.24	75
A-PP	2	24.56	125
A-RHE	9	26.8	6
A-SURG	6	30	44
A-TOXI	4	25.7	44
A-UN	15	29.8	10

[a] The full names of the AoFs are shown in the legends of Table 2 and Table C1.

restriction enzyme articles in the applications-oriented transition of knowledge utilization is still continuing, particularly in the field of Biochemical Research Methods, see Table 3. Note that we use 1969+k, 1970+k and 1971+k (depending on the article) to calculate the values for the year Y + k for *FT,* and use the average of 1969+k, 1970+k and 1971+k to calculate the values for the year Y + k for *AGT-DA*, k = 0, . . .,48. The column *Received citations* refers to the number of citations received by the A-S-N set from articles belonging to the AoF mentioned in the first column (recall that data were collected in December 2017).

One might assume that when a first citation occurs soon, then the total number of items may tend to be larger. Stated otherwise: the smaller *FT* the larger the number of received citations. We checked this by determining the Pearson correlation between these two columns. A value of -0.461 partially confirms our assumption; removing the outlier Biochemical Research Methods, even leads to a correlation of -0.604.

5.4.2. The speed of knowledge transition from the mediator set to technological areas and the time-span of the D-A-T path

In most technological classes the age of articles cited from AoFs (*AGT-AT*) is less than 10-years (6.14 on average), reflecting that the phase transition of restriction enzyme in knowledge utilization from AoFs to technological fields is very fast, see Table 4.

The value of *TS-DAT* in Table 4 is the difference in years between a citation in a patent and the publication of one of the A-S-N articles, following a D-A-T path, through a citation in AoF. By the shortest time we mean the smallest number of years needed to complete a D-A-T path. The mean in the fourth column is the average value of the time-span of an article in the D-A-T transition, based on the number of patents in a technological class. The shortest D-A-T path appears in C12; moreover the mean *TS-DAT* is close to 30 years.

6. Discussion and conclusions

In this contribution, we developed a new approach to explore the process of knowledge transition from discovery-oriented science to applications-oriented science and from applications-oriented science to being a part of knowledge flows in technological fields. Using two databases, the D-A-T trajectory was constructed to identify the different states of knowledge utilization in the evolution of discovery-oriented science to technology. We are convinced that our method will provide a useful tool for bridging the existing knowledge gap in detecting evolutions between science and technology (Narin & Noma, 1985).

As a preliminary study to track knowledge utilization in an applications-oriented evolution, we did not show all cases of knowledge transfer in citation networks. Yet, as illustrated by Fig.1, a conventional route (Path 2) plays a minor role in the process of knowledge transition. This fact was supported by the evidence from Tijssen (2010) that the 'basic' journals within fields of science are largely disconnected from technological development. Moreover, our results shown in Fig. 2 also reflect that the transition rates from the DoFs to AoFs via direct citations are very low (86 articles in AoFs among 1152 items).

In contrast, the D-A-T evolution illustrated by Path 1, taking AoFs transition directly or indirectly (two forward citation generations) into account, determined 988 articles in AoFs starting from the A-S-N set (DoFs), which is more than ten times the number of items detected via Path2. Moreover 387 patents were successively tracked from the articles in AoFs, forming a huge "invisible contribution" of the three fundamental restriction enzyme articles to scientific and technological fields. We found that 12.55% of the articles in the mediator set were cited by patents and that the NPR transition rate is 0.392. Here we want to recall that Tijssen found only 31 AoFs with a NPR-score of at least 0.25%. We admit that our approach and corresponding calculation methods are quite different from Tijssen's. Nevertheless, the percentages we obtained show

Table 4
The age of transition of articles from the mediator set and the time-span of the D-A-T path.

IPCs	AGT-AT	TS-DAT		Number of patents
		The shortest time	Mean	
A01	14	15	33.5	2
A23	4	43	43	1
A61	8.17	19	37.2	79
B01	4.44	27	33.46	48
B03	13	30	30	1
B05	10	43	43	1
B41	6	28	28	1
B81	9	29	31	3
B82	6.33	28	32.33	3
C02	4	32	32	1
C07	6.92	19	36.85	88
C08	3	32	36.67	3
C09	5	41	41	1
C10	0.5	38	38.5	2
C11	15.4	29	34.4	5
C12	5.19	13	39.4	264
C40	13	28	38.7	13
D06	19	30	30	1
F15	7	29	29	2
F16	7	29	29	1
G01	6.99	18	35	82
G02	6	28	31.67	3
G06	3.67	28	40.56	9
G11	7.5	28	29.5	2
H01	8	27	32.4	5
H03	2	46	46	1
H04	3	45	45	1

that our approach is more focused on the articles that really matter, i.e., those in the mediator set. Yet, both approaches are valuable, reflecting different points of view on the study of knowledge transition.

Our observations suggest that Path 1 in the D-A-T evolution is the backbone to track knowledge transfer originating from of a scientific work during its evolution outside the realm of academia; the D-A-T path is particularly useful for real-world tracking of the contribution of a pioneering work or of transformative research to progress made in S&T. Originality has, indeed, been a major point of attention and discussion in studies of research and its evaluation over the latest years (Grayson, 2015; Hu & Rousseau, 2016; Van Noorden, 2010).Only a minority of original, basic research can make it into patents and subsequent industrial applications on their own. In most cases a step in between via applications-oriented research is needed. This does not diminish the innovativeness of the original source. On the contrary: our approach serves to bring "sources of originality" to the limelight.

One limitation of this work is that we use only the Top25 AoFs based on JAD categories (Tijssen, 2010). However, we consider the results of Tijssen (2010) about applications-oriented research environments to be the most rational one available, so that the Top 25 AoFs are sufficiently robust proxies for identifying applications-oriented science. A possible way to lift this restriction is applying the distance metric proposed by Ahmadpoor and Jones (2017) who state that this is more precise than using traditional descriptors such as "basic" and "applied" science. Another point of discussion is that we started from a set of basic articles and followed their trace leading to patents. A complete picture would emerge if this information were augmented by a backward study of the resulting patents leading to all original basic knowledge (as suggested by a reviewer).

The results from our study suggest that it is feasible to develop a new roadmap tracking knowledge transfer starting from an original scientific work to its final transition in patents. Not only did we describe the methodology, outside the traditional route, but we also provided detailed results for an important case study, involving Nobel Prize winning publications.

By tracking the path of D-A-T evolution for three fundamental restriction enzymes articles, we determined the relationships between knowledge transition in ESI fields and AoFs, and understand that ESI fields Biology & Biochemistry, Clinical Medicine, Pharmacology & Toxicology, play major roles in this movement. Moreover, the AoFs Biochemical Research Methods, General & Internal Medicine, Toxicology, Pharmacology & Pharmacy are main actors in the phase transition. Biochemical Research Methods played the most important role in the knowledge transfer from AoFs into technological areas.

To determine the key characteristics of knowledge transition through the D-A-T path, some new (and old) measures were proposed and applied in the case study. Through the case study of the famous trio of articles related to restriction enzymes and the results shown, we come to the final conclusions of our study:

(1) The D-A-T tracking path makes it feasible to describe the evolution of science and technology from an applications-oriented point of view.

(2) Compared to direct citation analysis influenced by the idea of a separation of science and technology, the D-A-T path, taking the intermediate step via applied-oriented research into account, is the real backbone to reveal important "invisible contributions" of an original scientific work during its evolution outside academia.

(3) By the calculated values of the used indicators we not only understand the process of knowledge transition in the successive movements of knowledge utilization environments, but also know quantitative dependency relationships among ESI fields, applications-oriented fields, and technological classes. Moreover, the speed in phase transition, and the time-span of the D-A-T path were determined.

Finally, we mention that our study could be expanded using e.g. mapping methods such as Garfield's HistCite™ (Garfield, Pudovkin, & Istomin, 2003).

Author's contribution

Xiaojun Hu: Conceived and designed the analysis; Collected the data; Contributed data or analysis tools; Performed the analysis; Wrote the paper.

Ronald Rousseau: Performed the analysis; Wrote the paper.

Acknowledgements

Appendix A.

Table A1

Table A1
Top 25 applications-oriented fields based on WoS Subject Categories (Tijssen, 2010).

Top 10 Industrial-oriented	Top 10 Clinical practice-oriented
Petroleum Engineering	Emergency Medicine
Telecommunications	Orthopedics
Engineering, Electrical & Electronic	Critical Care Medicine
Chemistry, Medicinal	Surgery
Materials Science - Coatings & Films	Pediatrics
Nanoscience & Nanotechnology	Rheumatology
Imaging Science & Photographic Technology	Medicine, General & Internal
Computer Science - Hardware & Architecture	Medical Laboratory Technology
Manufacturing Engineering	Urology & Nephrology
Energy & Fuels	Gastroenterology & Hepatology
Top 5 Industrial/clinical-oriented	
Pharmacology & Pharmacy	
Materials Science - Biomaterials	
Toxicology	
Biochemical Research Methods	
Medicine - Legal	

Appendix B.

We provide a simple example illustrating the relations $\sum_i RS(i) = 1$ and $\sum_j RS(j) \geq 1$.

Assume that the Mediator Set contains three articles: x, y and z. Assume that x belongs to E(1), while y and z belong to E(2). Recall that the ESI fields are non-overlapping so that none of these three articles can belong to two ESI fields at the same time. Now $RS(1) = \frac{1}{3}$ while $RS(2) = \frac{2}{3}$. Their sum is clearly equal to 1.

The AoF sets A(.) on the other hand are overlapping. So we may assume that x belongs to A(j) and to A(k); y belongs to A(m) and z belongs to A(j) and to A(m). Now $RS(j) = \frac{\#\{A(j) \cap Mediator\ Set\}}{3} = \frac{\#\{x,z\}}{3} = \frac{2}{3}$; $RS(k) = \frac{\#\{A(k) \cap Mediator\ Set\}}{3} = \frac{\#\{x\}}{3} = \frac{1}{3}$; $RS(k) = \frac{\#\{A(k) \cap Mediator\ Set\}}{3} = \frac{\#\{x\}}{3} = \frac{1}{3}$; and finally $RS(m) = \frac{\#\{A(m) \cap Mediator\ Set\}}{3} = \frac{\#\{y,z\}}{3} = \frac{2}{3}$. Their sum is $\frac{5}{3} \geq 1$.

Appendix C.

Table C1
The relational strength between ESI fields and AoFs in the mediator set. Row and column totals and largest cell values are indicated in grey.

AoF \ ESI-F	A-BRM	A-MC	A-CCM	A-EF	A-EEE	A-GH	A-MSB	A-MLT	A-GIM
BB	0.319	0.012	0.000	0.026	0.001	0.000	0.002	0.001	0.000
CH	0.051	0.011	0.000	0.001	0.000	0.000	0.000	0.000	0.000
CM	0.002	0.000	0.002	0.000	0.001	0.019	0.000	0.040	0.142
CS	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ENG	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.000
EN/EC	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.000	0.000
IM	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000
MS	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000
MICRO	0.020	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MBG	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
NB	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.000
PT	0.004	0.016	0.000	0.000	0.000	0.000	0.000	0.000	0.002
PHYS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
PAS	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Total	0.423	0.040	0.002	0.032	0.003	0.019	0.004	0.044	0.145

AoF \ ESI-F	A-LG	A-NN	A-PED	A-PP	A-RHE	A-SURG	A-TOXI	A-UN	*Total*
BB	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.362
CH	0.000	0.009	0.000	0.001	0.000	0.000	0.000	0.000	0.069
CM	0.003	0.001	0.068	0.001	0.006	0.046	0.000	0.003	0.336
CS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.008
ENG	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
EN/EC	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.005
IM	0.000	0.000	0.007	0.000	0.000	0.000	0.000	0.000	0.011
MS	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.005
MICRO	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.020
MBG	0.000	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.019
NB	0.000	0.000	0.001	0.012	0.000	0.000	0.008	0.000	0.020
PT	0.000	0.001	0.000	0.106	0.000	0.000	0.163	0.000	0.134
PHYS	0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.004
PAS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004
Total	0.003	0.018	0.076	0.123	0.006	0.046	0.179	0.003	

A-BRM = Biochemical Research Methods; A-MC = Chemistry, Medicinal; A–CCM = Critical Care Medicine; A-EF = Energy & Fuels; A-EEE = Engineering, Electrical & Electronic; A-GH = Gastroenterology & Hepatology; A-MSB = Materials Science, Biomaterials; A-MLT = Medical Laboratory Technology; A-GIM = Medicine, General & Internal; A-LG = Medicine, Legal; A-NN = Nanoscience & Nanotechnology; A-PED = Pediatrics; A-PP = Pharmacology & Pharmacy; A-RHE = Rheumatology; A-SURG = Surgery; A-TOXI = Toxicology; A-UN = Urology & Nephrology.

References

Ahmadpoor, M., & Jones, B. F. (2017). The dual frontier: Patented inventions and prior scientific advance. *Science*, *357*(6351), 583–587.

Arber, W., & Linn, S. (1969). DNA modification and restriction. *Annual Review of Biochemistry*, *38*(July), 467–500.

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

Boyack, K. W., Klavans, R., Small, H., & Ungar, L. (2014). Characterizing the emergence of two nanotechnology topics using a contemporaneous global micro-model of science. *Journal of Engineering and Technology Management*, *32*(2), 147–159.

Clarivate. (2017). *Clarivate*. https://clarivate.com/products/derwent-innovation/

Danna, K., & Nathans, D. (1971). Specific cleavage of simian virus 40 DNA by restriction endonuclease of Hemophilus influenzae. *Proceedings of the National Academy of Sciences USA*, *68*(12), 2913–2917.

de Solla Price, D. J. (1983). Of sealing wax and string: A philosophy of the experimenters craft and its role in the genesis of high technology. In *Little science, big science and beyond*. pp. 237–253. New York: Columbia University Press.

Frandsen, T. F. (2004). Journal diffusion factors: A measure of diffusion? *Aslib Proceedings*, *56*(1), 5–11.

Frandsen, T. F., Rousseau, R., & Rowlands, I. (2006). Diffusion factors. *Journal of Documentation*, *62*(1), 58–72.

Gao, J. P., Ding, K., Teng, L., & Pang, J. (2012). Hybrid documents co-citation analysis: Making sense of the interaction between science and technology in technology diffusion. *Scientometrics*, *93*(2), 459–471.

Garfield, E., Pudovkin, A. I., & Istomin, V. S. (2003). Why do we need algorithmic historiography? *Journal of the Association for Information Science and Technology*, *54*(5), 400–412.

Gervais, M.-J., Marion, C., Dagenais, C., Chiocchio, F., & Houlfort, N. (2016). Dealing with the complexity of evaluating knowledge transfer strategies: Guiding principles for developing valid instruments. *Research Evaluation*, *25*(1), 62–69.

Gherardini, A., & Nucciotti, A. (2017). Yesterday's giants and invisible colleges of today. A study on the 'knowledge transfer' scientific domain. *Scientometrics*, *112*(1), 255–271.

Glänzel, W., & Meyer, M. (2003). Patents cited in the scientific literature: An exploratory study of reverse citation relations. *Scientometrics*, *58*(2), 415–428.

Grayson, M. (2015). Raising standards. *Nature*, *520*(7549), S10–S12.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *The Rand Journal of Economics*, *36*(1), 16–38.

Hu, X. J., & Rousseau, R. (2016). Scientific influence is not always visible: The phenomenon of under-cited influential publications. *Journal of Informetrics*, *10*(4), 1079–1091.

Hu, X. J., & Rousseau, R. (2017). Nobel prize winners 2016: Igniting or sparking foundational publications? *Scientometrics*, *110*(2), 1053–1063.

Hu, X. J., & Rousseau, R. (2018). A warning for Chinese academic evaluation systems: Short-term bibliometric measures misjudge the value of pioneering contributions. *Journal of Zhejiang University-Science B*, *19*(1), 1–5.

Hu, X. J., Rousseau, R., & Chen, J. (2011). On the definition of forward and backward citation generations. *Journal of Informetrics*, *5*(1), 27–36.

Hu, X. J., Rousseau, R., & Chen, J. (2012). A new approach for measuring the value of patents based on structural indicators for ego patent citation networks. *Journal of the American Society for Information Science and Technology*, *63*(9), 1834–1842.

Kuan, C.-H., Huang, M.-H., & Chen, D.-Z. (2018). Missing links: Timing characteristics and their implications for capturing contemporaneous technological developments. *Journal of Informetrics*, *12*(2), 259–270.

Lewison, G., Rippon, I., & Wooding, S. (2005). Tracking knowledge diffusion through citations. *Research Evaluation*, *14*(1), 5–14.

Li, D., Azoulay, P., & Sampat, B. N. (2017). The applied value of public investments in biomedical research. *Science*, *356*(6333), 78–81.

Liu, Y. X., & Rousseau, R. (2010). Knowledge diffusion through publications and citations: A case study using *ESI*-fields as unit of diffusion. *Journal of the American Society for Information Science and Technology*, *61*(2), 340–351.

Liu, Y. X., Rafols, I., & Rousseau, R. (2012). A framework for knowledge integration and diffusion. *Journal of Documentation*, *68*(1), 31–44.

Luwel, M., & van Wijk, E. (2015). The' translational' effect in medical journals: Bridging the gap? *Research Evaluation*, *24*(1), 19–29.

Marx, W., Bornmann, L., Barth, A., et al. (2013). Detecting the historical roots of research fields by Reference Publication Year Spectroscopy (RPYS). In J. Gorraiz (Ed.), *Proceedings of the 14th International Conference on scientometrics and informetrics* (pp. 493–506). Vienna: AIT Austrian Institute of Technology.

Marx, W., Bornmann, L., Barth, A., & Leydesdorff, L. (2014). Detecting the historical roots of research fields by Reference Publication Year Spectroscopy (RPYS). *Journal of the Association for Information Science and Technology*, *65*(4), 751–764.

Meyer, M. (2006). Measuring science-technology interaction in the knowledge- driven economy: The case of a small economy. *Scientometrics*, *66*(2), 425–439.

Narin, F., & Hamilton, K. S. (1996). Bibliometric performance measures. *Scientometrics*, *36*(3), 293–310.

Narin, F., & Noma, E. (1985). Is technology becoming science? *Scientometrics*, *7*(3-6), 369–381.

Narin, F., & Olivastro, D. (1998). Linkage between patents and papers: An interim EPO/US comparison. *Scientometrics*, *41*(1-2), 51–59.

Pan, R. K., Sinha, S., Kaski, K., & Saramäki, J. (2012). The evolution of interdisciplinarity in physics research. *Scientific Reports*, (2), art. 551

Roberts, R. J. (2005). How restriction enzymes became the workhorses of molecular biology. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(17), 5905–5908.

Rousseau, R. (1994). Double exponential models for first-citation processes. *Scientometrics*, *30*(1), 213–227.

Rowlands, I. (2002). Journal diffusion factor: A new approach to measuring research influence. *Aslib Proceedings*, *54*(2), 77–84.

Seiken, J. (1993). Grant program encourages collaboration. *Scientist (Philadelphia, Pa)*, *7*(4), 21-21.

Sherwin, C. W., & Isenson, R. S. (1967). Project hindsight - A defense department study of the utility of research. *Science*, *156*(3782), 1571–1577.

Smith, H. O., & Wilcox, K. W. (1970). A restriction enzyme from haemophilus-influenzae. 1. Purification and general properties. *Journal of Molecular Biology*, *51*(2), 379–391.

Sudhindra, S., Ganesh, L. S., & Arshinder, K. (2017). Knowledge transfer: An information theory perspective. *Knowledge Management Research & Practice*, *15*(3), 400–412.

Tangaraja, G., Rasdi, R. M., Samah, B. A., & Ismail, M. (2016). Knowledge sharing is knowledge transfer: A misconception in the literature. *Journal of Knowledge Management*, *20*(4), 653–670.

Tijssen, R. J. W. (2010). Discarding the 'basic science/applied science' dichotomy: A knowledge utilization triangle classification system of research journals. *Journal of the American Society for Information Science and Technology*, *61*(9), 1842–1852.

Tijssen, R. J. W., & Winnink, J. (2016). Twenty-first century macro-trends in the institutional fabric of science: Bibliometric monitoring and analysis. *Scientometrics*, *109*(3), 2181–2194.

Van Looy, B., Landoni, P., Callaert, J., van Pottelsberghe, B., Sapsalis, E., & Debackere, K. (2011). Entrepreneurial effectiveness of European universities: An empirical assessment of antecedents and trade-offs. *Research Policy*, *40*(4), 553–564.

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

van Raan, A. F. J. (2017). Sleeping beauties cited in patents: Is there also a dormitory of inventions? *Scientometrics*, *110*(3), 1123–1156.

WIPO. (2017). *International patent classification (version 2017)* Available at. http://www.wipo.int/classifications/ipc