European Commission

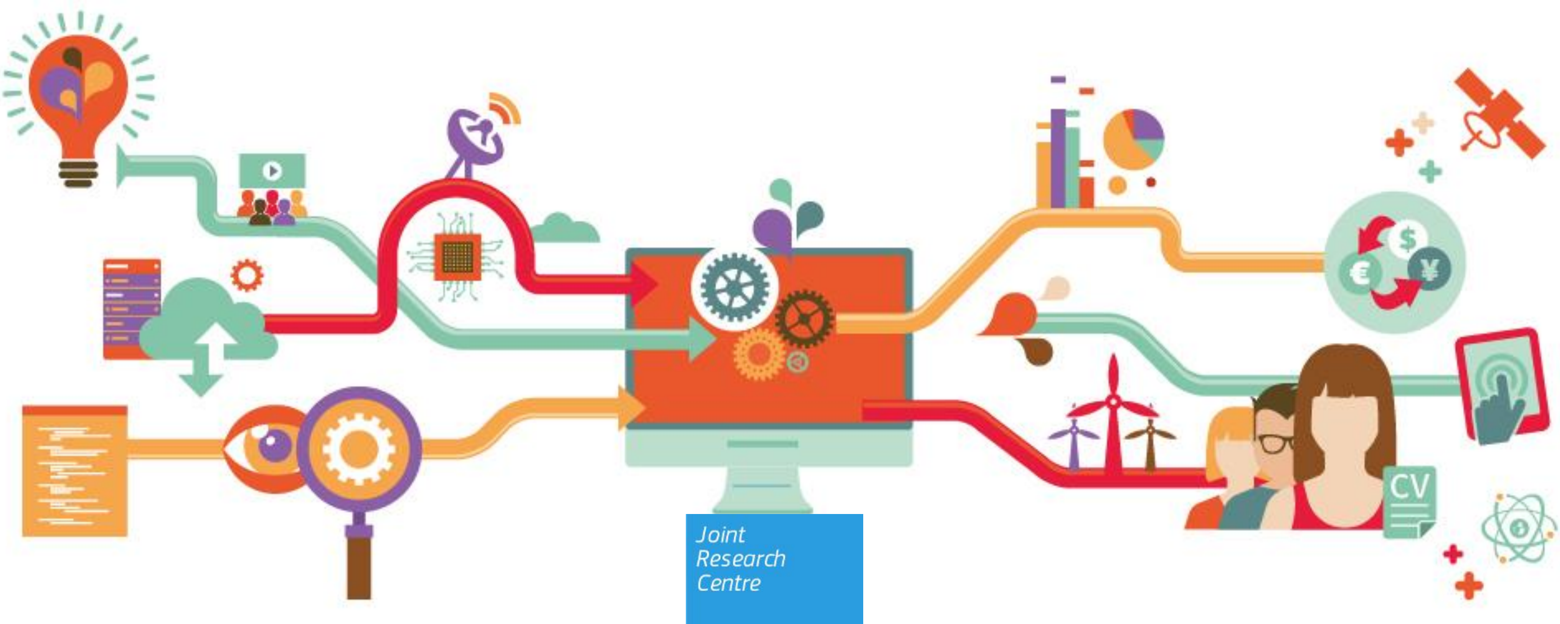
JRC TECHNICAL REPORTS

Technological diffusion as a recombinant process

JRC Working Papers on Corporate R&D and Innovation No 07/2016

Petros Gkotsis, Antonio Vezzani

2016



Joint Research Centre

Contact information
Fernando Hervás Soriano
Address: Edificio Expo. c/ Inca Garcilaso, 3. E-41092 Seville (Spain)
E-mail: jrc-ipts-kfg-secretariat@ec.europa.eu
Tel.: +34 954488463
Fax: +34 954488316

JRC Science Hub
https://ec.europa.eu/jrc

JRC102638

ISSN 1831-9408 (online)

Seville, Spain: European Commission, 2016

How to cite: Gkotsis, P. and Vezzani, A. (2016). Technological diffusion as a recombinant process. JRC Working Papers on Corporate R&D and Innovation No 07/2016, Joint Research Centre.

The JRC Working Papers on Corporate R&D and Innovation addresses economic and policy issues related to industrial research and innovation and to the competitiveness of the European industry. Mainly addressed to policy analysts and the academic community, these are policy relevant early-stage scientific articles highlighting policy implications. These working papers are meant to communicate to a broad audience preliminary research findings, generate discussion and attract critical comments for further improvements. All papers have undergone a peer review process.

Technological diffusion as a recombinant process[1]

Petros Gkotsis and Antonio Vezzani

European Commission, Joint Research Centre, Seville, Spain

Abstract

In this work we analyse patterns of technological development using patent applications at the United States Patent and Trademark Office (USPTO) over the 1973-2012 period. Our study focuses on the combinations of technological fields within patent documents and their evolution in time, which can be modelled as a diffusion process. By focusing on the combinatorial dimension of the process we obtain insights that complement those from counting patents. Our results show that the density of the technological knowledge network increased and that the majority of technological fields became more interconnected over time. We find that most technologies follow a similar diffusion path that can be modelled as a Logistic or Gompertz function, which can then be used to estimate the time to maturity defined as the year at which the diffusion process for a specific technology slows down. This allows us to identify a set of promising technologies which are expected to reach maturity in the next decade. Our contribution represents a first step in assessing the importance of diffusion and cross-fertilization in the development of new technologies, which could support the design of targeted and effective Research & Innovation and Industrial policies.

Keywords: technological diffusion, patents, knowledge.

JEL Classification: O33; O31; C10.

[1] We are grateful to Alex Coad and the participants of the Eurkind #GCW2016 Conference "Innovation, Employment, the Environment" (Valencia, 2016) and the 16th Schumpeter Conference on Evolutionary Economics (Montreal, 2016) for their useful insights and comments. We are responsible for any omission or remaining mistake. This Working Paper is issued in the context of the **Key Enabling and Emerging Technologies for Competitiveness (KeyTEC)** activities, carried out by the European Commission's Joint Research Centre (JRC).

1. Introduction

In the Schumpeterian system, business cycles are related to the occurrence of innovations in time (Schumpeter, 1939). Innovations tend to appear in clusters that shape technological development and the way the economy works. However, an explanation of the way new technologies emerge was beyond the scope of Schumpeter's work and as a result the way these clusters arise is still undetermined. Recently, in an attempt to foresee the occurrence of new innovation waves, a growing body of innovation literature focused on the emergence of new technologies. A common understanding on what an emerging technology is, and how it can be detected, has not been reached yet. However, the relative fast rate of growth of a technology (or scientific field) is one of the most frequent attributes considered as a condition for emergence (Rotolo et al., 2015). Nevertheless, recent contributions recognize the importance of looking at the co-development of technologies as well (e.g. Dernis et al., 2015). Verhoeven et al. (2016) highlight the superior performance of patents combining IPC codes which had never been combined before; recombining knowledge in original ways may lead to superior innovative performances.

The idea of knowledge recombination can be traced back to Weitzman, who described innovation as an endogenous combinatoric process where new knowledge is created by re-combining previous one. Thus, the number of possible *"untried combinations of existing ideas eventually grows much faster than anything else in the economy"* (Weitzman, 1996 p.211), but the capacity of exploring and realizing new combinations grows at a much lower pace, constraining the knowledge generation process (and growth). Despite the interest on this idea, an analysis of its actual features is still missing. Indeed, the focus of the new micro-level studies has been largely on the cognitive dimension of knowledge recombination. For example, Gruber et al. (2013) analyse the relation between the inventors backgrounds (scientist or engineer) and their capacity to produce inventions spanning over a broader set of technologies, while Jones (2008) discussed the possible educational burdens posed by the increasing technological complexity.

In a knowledge recombination framework, the more a technology is combined with others the greater its importance within the overall production of new knowledge. As a result, the probability for a technology to emerge could be related to its diffusion in the technological development process. Therefore, insights on the characteristics of technological diffusion in the knowledge space may provide a basis to better understand how technologies emerge. Here diffusion is seen as a time dependent stochastic process causing a spread of a specific technology in the knowledge space. In this respect, diffusion should be understood differently from the concept of "innovation diffusion" (Rogers, 1983), which represent the common understanding in innovation studies. For us diffusion is an attribute defining the spread of technologies from a "production" rather than an adoption point of view. Understanding the combinatorial dimension of technological diffusion will shed new light on technological development. In what follows we will analyse the degree of integration of different technologies in the knowledge base relying on the standard technological classification used to classify patent documents.

2. Recombinant knowledge and technological diffusion

Technological competition has become global and the race for the top has moved from the control of natural resources to the development of high technology products allowing pushing further the technological frontier. Technological competition has increased and the rate of patenting invention increased drastically since the nineties. This attracted the attention of academics and policy makers because understanding technological development and forecasting new promising (emerging) technologies is key to designing targeted interventions - especially in case of limited resources - aiming at fostering countries' competitiveness.

Despite the interest they have attracted from the research community, a consensus of what emerging technologies are, has not been reached yet. For example Porter et al. (2002) stressed the potential impact of emerging technologies on the economy, while others focused on the uncertainty characterizing the process of emergence (Boon and Moors, 2008) or on their novelty and growth potential (e.g. Small et al., 2014). Recently, Rotolo et al. (2015) contributed to the discussion by proposing a framework to conceptualize emerging technologies. In their view, these share five basic characteristics: radical novelty; fast growth; coherence; potential socio-economic impact; uncertainty.

Clearly, in order to grow fast (or faster than others), technologies should be developed and adopted by an increasingly large number of inventors and users. In this respect emergence is intrinsically related to the concept of diffusion, which in economics dates back to 1957 when Griliches analysed the diffusion of hybrid corns based on epidemic models. Mansfield (1963) discussed the rates at which a firm adopting new techniques proceeds to substitute old ones with these new ones and concluded that in order for economies to benefit from innovation, the diffusion process should proceed at a sufficiently fast pace. As pointed out by Dosi (2013) the basic forces driving technological diffusion are the spread of information/knowledge and the expectation of profits, while development/adoption costs and the uncertainty surrounding new technologies represent barriers to diffusion. Rogers defines diffusion as *"the process by which an innovation is communicated through certain channels over time among the members of a social system. It is a special type of communication, in that the messages are concerned with new ideas."* (Rogers, 1983 p.5). In his view, time is a crucial element when analysing the diffusion of knowledge or new ideas. The rate of adoption here is understood as the speed at which an innovation is adopted by members of a social system.

Given that innovations are linked, to a certain extent, to patents, this conceptualization of the innovation diffusion process could lead to an assessment of technological performances based on different metrics based on patent counts. However, this approach can be affected by differences in patent propensity across industries, products and time. For example the iPhone - introduced in 2007 - was protected with a bundle of about 200 patens, while the Airbus filed around 380 patents during the development of the A380 (commercialized in the same year). In 2013 Apple held more than 1300 patents for the iPhone5 and its related software.[2] An alternative approach is to rely on information contained in patent documents in order to build metrics measuring the number of combinations of distinct technologies in the knowledge space.

In fact, innovation can be understood as an endogenous combinatoric process where new knowledge is created by re-combining previous one (Weitzman, 1996). According to Weitzman (1998) new knowledge builds itself upon combining existing knowledge in useful ways non-previously conceived. An analogue to the production of new ideas from the biological field is the development of new plant varieties by cross-pollinating existing ones. In this view, the technological discontinuities characterizing the rising of new technological paradigms (Dosi, 1982 1988) may be seen as part of the cumulative knowledge process, representing novel ways of combining existing knowledge.

In this view technological change can be seen as a macro process driven by the diffusion of specific technologies in the technical knowledge space via the formation of new (successful) technological combinations at the micro scale (e.g. patent document level). In other words, the focus shifts on the extent to which a given technology is combined with others to give rise to new applications. Diffusion therefore could be observed in the (increasing) number of ways a given technological field is combined with others.

This conceptualization of the knowledge creation process could be also linked to the idea of cross-fertilization of technologies in the development of new products. Patent analysis is particularly suitable in this framework, because patents are linked to the creation of new (technical) knowledge and can be seen as precursors of new products/processes. Moreover, Dernis et al. (2015), comparing publication and patent data, found that in some cases the acceleration in the development of science may follow the acceleration in the development of technologies. The authors also highlighted the importance of cross-fertilisation of scientific domains in identifying the emergence of new technologies. However, taking into account the uncertainty involved in the technological development (Martino, 2003) especially in the case of technologies at their infancy, an ex-ante identification of (successful) new technologies could represent a prohibitive task. Given the difficulties in forecasting emerging technologies, developing a methodology to identify those technologies closer to the maturity phase is of great importance for the support of the policy process. These technologies are expected to drive economic growth in the medium term.

In this work, we attempt to contribute to the discussion on technological development by analysing the spread of technologies in "knowledge production". Our point of departure is the combination of technological fields in patent documents and their evolution over time. In other words, we start from the co-occurrences of technologies (IPC codes) within patent documents to analyse the degree of interconnection of each specific technology in the technological knowledge base. Co-occurrences of technologies in patent documents have been already used in the literature (Breschi et al., 2003; Bar and Leiponen, 2012) to compute relatedness or distance between technologies. However, this approach does not take into account the dynamic nature

[2] For more information: http://www.ipeg.com/intellectual-property-in-our-daily-lives.

of the knowledge creation process, where a technology may change its relatedness (or position) with the rest of the system because of diffusion. Probably, the papers closer to ours are those of Krafft et al. (2011) - where network analysis measures were applied to study the knowledge base evolution of biotechnologies – and Youn et al. (2015) which attempt to provide a quantitative characterization of the combinatorial process underpinning inventive activity, with a focus on the combinations that have been already realized among the theoretically possible ones. Our rationale is that the increase of combinations mirrors the increasing importance of a specific technological field in the development of new technological applications and possibly of new technological knowledge as well.

3. Data and methodology

We analyse the technological development using patent applications at USPTO over the last 40 years (1973-2012). USPTO was chosen due to the availability of data covering a long period of time, not available in other large Intellectual Property Offices. The choice of applications over granted patents is due to the lag in granting that limits the availability of data in the recent years. Until 2000, the proportion of granted applications was around 90%. However, the recent surge in patent filings resulted in a longer lag between application and granting year, this in turn drastically lowered the share of granted patents in most recent years (to 60% in 2010).

In order to study how the combinations of technological fields evolve along time, we use some simple metric from the network analysis theory. In particular, for each year considered we build network graphs using the International Patent Classification (IPC) technological classes at the four digit level as nodes and their co-occurrences within patent documents as edges.[3]

From these networks we then compute the degree of node i, $d(n_i)$, which counts the number of ties (connections) incident on a node (IPC class).[4] The degree of a node measures the activity of the entity it represents (Wasserman and Faust, 1994) and can be interpreted as a measure of the immediate risk/probability that a distinct technology becomes relevant for the rest of the network. In this way we can investigate how different technological solutions emerge, diffuse, grow and decline over time. Indeed, although we focus on technological diffusion by monitoring the number of connections between the nodes of the network (the degree), we also consider the growth dimension in terms of patent filings.

[3] In network analysis, the structure of a network is characterized in term of nodes (the entities within the network) and ties or edges (relationships or interactions) that connect them.

[4] Although more sophisticated network measures are available, we opted for the degree because of its simplicity and fitness with the concept discussed in the paper. We are aware that the degree may suffer from 'spurious' co-occurrences, but we expect that even if this is the case, these to co-occurrences would disappear in the following year. Moreover, 'spurious' co-occurrences may represent unsuccessful combinations and we cannot exclude them *a priori*.

Our focus on the degree of a technology in the technical knowledge network and its evolution over time closely resembles the diffusion process as generally understood in the innovation literature. The adoption of innovations by consumers/firms has been analysed as a diffusion process by a number of researchers (e.g. Rogers, 1983; Rosenberg, 1976) who have highlighted a number of stylized facts recently reviewed and discussed by Dosi and Nelson (2013); these are also relevant in this case.

Diffusion is a time dependent process that can generally be represented by s-shape curves. The shape of the curve is defined by the rate of adoption, which may vary greatly among technologies; some new ideas diffuse relatively rapidly, showing a quite steep s-curve. However, only some innovations succeed in diffusing and among them there are some with very asymmetric profiles. In other words, innovations show very different lag profiles between their introduction and the start of the diffusion. The analysis of our processed data reveals similar behaviours for the technological fields represented by the IPC classes.

In order to analyse the diffusion process of technologies, we fit the evolution of their degree by using two different functional forms. In particular we use and compare the logistic (L) and the Gompertz (G) functions. These distributions have been normally used in the diffusion literature because of their suitability to fit s-shaped processes. However, differently from the Logistic, the s-shaped curve from the Gompertz function is not symmetric around its inflection point, where the concavity changes (Berger, 1980). In other words, the Gompertz function is more appropriate to fit asymmetric diffusion processes.

The logistic and Gompertz curves could be written as:

$L(t) = \frac{D}{1+a*exp^{-s_1(t-t_0)}}$ and $G(t) = De^{-b*exp^{-s_2 t}}$, with $lim_{t\to\infty}L(t), lim_{t\to\infty}G(t) = D$

where s_1 and s_2 represent the steepness (or growth paramenter) of the curve, t is the time (with t_o representing the sigmoid's midpoint), a and b are two constants. Finally, D represents the maximum possible degree for a given technology. Estimating maximum values in diffusion processes is a notoriously difficult problem; in this case the two most straightforward values for D are represented by the theoretical and empirical maximum degree of the network. The former is represented by the number of existing IPC4 codes (642), while the latter is represented by the highest degree observed in the data (465). An alternative approach would be to let D as an additional parameter to be estimated; however our tests have shown that for a number of IPC4 this would result in D values much higher than the possible theoretical maximum. Given this and the non-linear nature of the problem which requires initial guesses for the parameters, we decided to set D equal to 465 in order to reduce the uncertainty related to the initial choice of the parameters.[5] The choice resulted in a more reliable distribution of the estimated parameters.

[5] The maximum empirical degree (465) was observed in 2004 for code G06F. We have tested both maximum values (empirical and theoretical) and the results of the estimations do not differ significantly. The time to maturity is slightly longer when assuming that all technologies will finally become fully interconnected.

By comparing the goodness of fit obtained from the different functional forms, we are able to choose which one is the most appropriate for each specific IPC4. Once the functional form has been selected, we can derive our estimates for the current rate of the diffusion process and the time to maturity. The former is computed as the first derivative of the function in the last year of observation, 2012. The first derivative provides the slope of the function, measuring the rate of change of the degree: the number of new IPC codes combined with a given IPC code in 2012. The latter is instead obtained by projecting the functional form from 2013 on and computing the second derivative of the function in each year. The second derivative measures how the rate of change is itself changing, therefore a value lower than zero indicates that the process decelerates. We classify a technology as mature in the year when the second derivative is negative for the first time. By doing so we define a technology as mature when its diffusion process decelerates, meaning that it can still continue to diffuse (albeit at a slower pace) and, most important, to give economic returns since we do not link technological diffusion/maturity with market performances.

4. Results and discussion

4.1 Describing the diffusion process

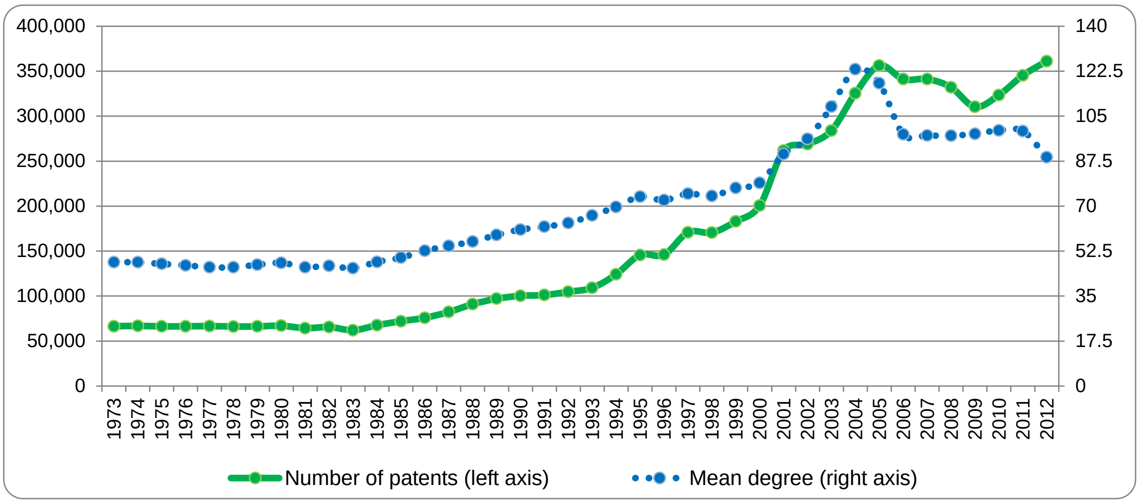
The number of active IPC codes in USPTO has slightly increased during the period of analysis. In 1973, 619 different IPC4 codes have been used within patent documents; this number has reached its maximum value (632) in 2004 to slightly decrease since then to 626 in 2012. This change in the number of network nodes would suggest using the normalized degree[6] to compare data between different years. However, the change in IPC codes is rather small and the correlation between the degree and the normalized degree is extremely high (0.999). Therefore, we decided to present degree statistics that allow to better understand the actual dimension of the phenomenon under study. Figure 1 shows the number of patents filed at the UPSTO each year between 1973 and 2012 (left axis), together with the mean degree of the corresponding IPC4 network (right axis). Over the years both the mean degree and the total number of patents have increased. However, after 2004 the two seem to diverge: while patent filings decreased between 2004 and 2008 to then recover, mean degree experienced a big drop between 2004 and 2006 and then stagnated.

The drop in the mean degree is unexpected and followed a long period during which the density of the technology network steadily increased. This result would suggest that the complexity of the technological knowledge generation process was limited in recent years. However, such a conclusion would require additional and more specific evidence. Moreover, we should point out that this fact coincided with the end of the

[6] The normalized degree is obtained by dividing the degree by the number of network nodes minus one; this allows comparisons among networks of different sizes.

reform period for the last update of the IPC classification, although after each revision of the classification patent documents are reclassified accordingly.[7]

Figure 1: Patent applications and mean degree of IPC codes



Note: *authors' calculations on USPTO data, 1973-2012.*

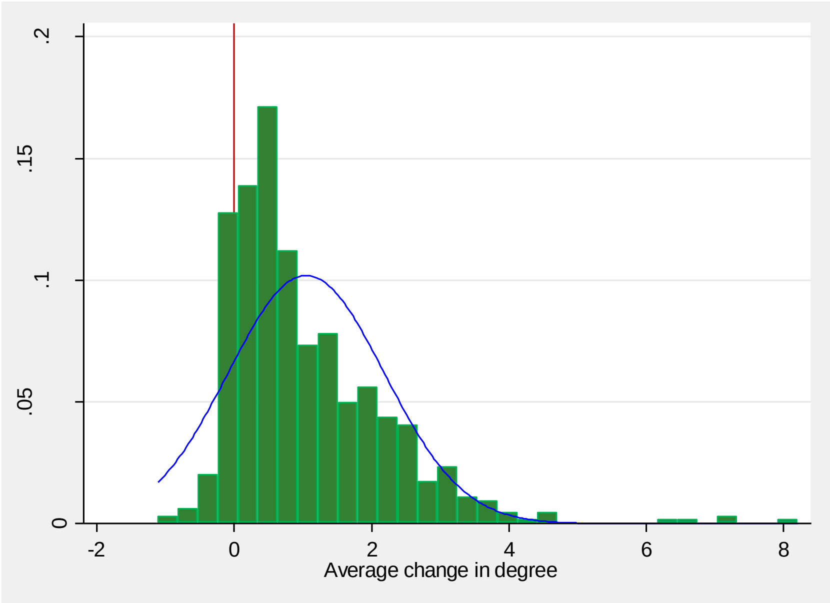
At the IPC4 level the correlation between patent filings and degree is not particularly strong (0.487). In order to test for predictive causality between the two data series (Diebold, 2001), we ran a Granger causality test (Granger, 1969) on the annual changes of the degree (Δd) and number of patents (Δp) for each IPC4 code. In 58.4% of cases the null hypothesis that Δp (Δd) does not Granger-cause Δd (Δp) cannot not be rejected at the usual 5% significant level; for these technologies lagged changes in diffusion are not statistically related with present changes in patent applications, and vice versa. For 21.8% of technologies Δd "Granger-causes" Δp but not the other way round. In 11.4% of cases Δp "Granger-causes" Δd and, finally, in the remaining 8.4% of cases we find evidence of Granger causality in both directions. These results suggest that our approach of introducing a social network perspective in the analysis of patent data complements patent counting statistics providing additional insights for the understanding of the technological diffusion process.

Our results suggest that the majority of IPC codes have increased in degree over the period considered. The average value of change in degree is 1.048, the median is 0.718, implying an increase in the complexity of technologies developed during the 40 years considered (figure 2). Only about 10% of IPC classes experienced a decrease in their degree. Moreover, the distribution is right skewed, with a few technologies diffusing much more than the rest.

---

[7] The transitional revision period started in 1999 and in 2005 the basic period of reform was completed.

Figure 2: Average degree changes of IPC codes distribution



Note: *authors' calculations on USPTO data, 1973-2012.*

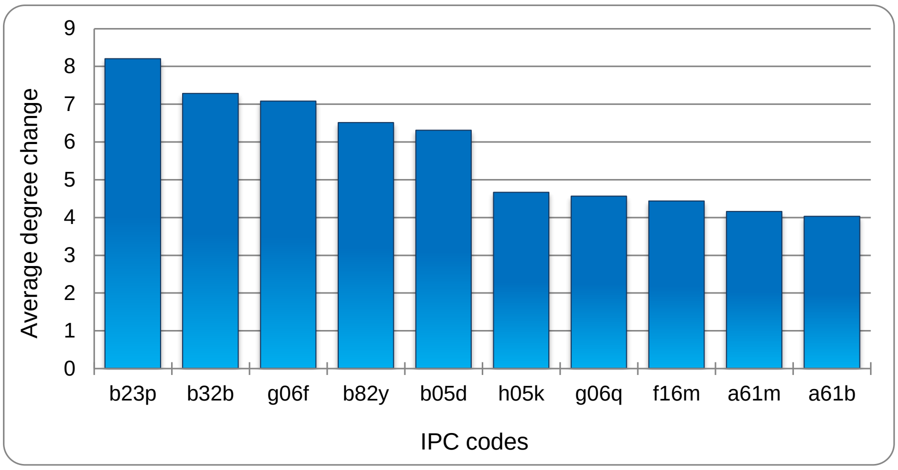
The increased density of the network of IPC4 codes, which is linked to an increase in complexity of technologies/applications over time, is evident in figures 1 and 2. This is the result of a general increase in complexity for most technological fields, with a small number of technologies experiencing a particularly high increase. For illustration purposes we report the 10 technological fields with the highest average degree increase in figure 3. It is interesting to note that 5 among these 10 technologies are related to machinery and material development and testing (b23p, b32b, b82y, b05d, f16m)[8], 2 directly linked to data processing (g06f, g06q), 2 to medical applications (a61m, a61b) and one to electronics/electric components manufacturing (h05k).

Our top technological fields partly overlap with those reported as bursting by Dernis et al. (2015). However, the focus there was on the acceleration in the co-development of patented technologies given by the number of patents related to specific IPC4 (or

[8] G06f – Electric digital data processing; B32b – Layered products [i.e. products built-up of strata of flat or non-flat (e.g. cellular or honeycomb) form]; B23p – Other working of metal; combined operations; universal machine tools; B82y - Specific uses or applications of nano-structures; measurement or analysis of nano-structures manufacture or treatment of nano-structures; B05d - Processes for applying liquids or other fluent materials to surfaces, in general; H05k - Printed circuits; casings or constructional details of electric apparatus; manufacture of assemblages of electrical components; G06q - Data processing systems or methods, specially adapted for administrative, commercial, financial, managerial, supervisory or forecasting purposes; systems or methods specially adapted (same); F16m - Frames, casings, or beds, of engines or other machines or apparatus, not specific to an engine, machine, or apparatus provided for elsewhere; stands or supports; A61m - Devices for introducing media into, or onto, the body (...); devices for transducing body media or for taking media from the body (…); devices for producing or ending sleep or stupor; A61b - Diagnosis; surgery; identification.

IPC7) pairs. Moreover, the unit of analysis in their study was the IP5 Patent families[9] rather than USPTO patent.

Figure 3: The 10 IPC codes with the highest average change in degree



Note: *authors' calculations on USPTO data, 1973-2012.*

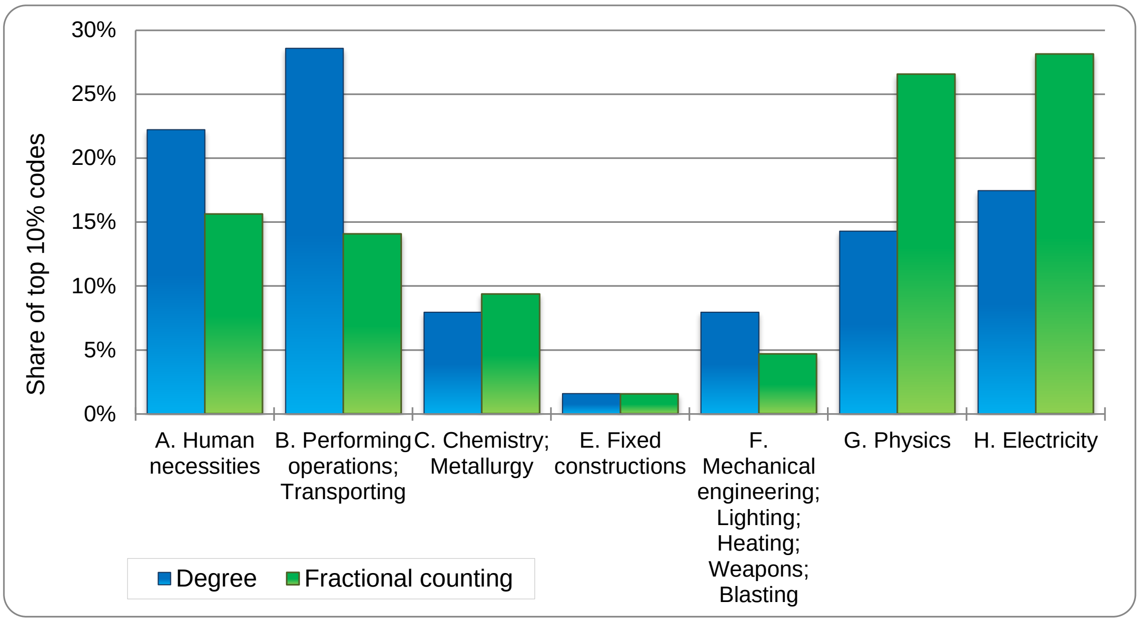
What is common in patent literature and in studies to support policy making is to count patents per IPC code as a metric to assess technological performances. However, this approach can be affected by differences in patent propensity across industries. In particular, the patent propensity, measured by the number of patents over R&D investments, largely depends on the costs associated to the development of new applications. We argue that our method, by focusing on the connections among technologies should be less prone to such type of bias. Here we perform a simple test to demonstrate the differences between the two approaches. For this test, we rank IPC4 codes according to the average increase in degree and number of patents over the last forty years. The top 10% of codes from the two rankings are then selected to count the number of occurrences for each IPC1 class; the relative number of occurrences (share) is presented in figure 4.

Fractional counting of patents results in the high growth of "Physics" and "Electricity" related codes (generally understood as information and communication technologies, ICT), with more than 50% of high growth codes belonging to these 2 categories. Considering high diffusing codes provides very different insights: "Human necessities" and "Performing operations; Transporting" (which includes materials) are the categories that stand out from the others.

[9] IP5 Patent families are defined as families of patents filed in at least two IPOs, one of which should be amongst the top patenting offices worldwide: European Patent Office (EPO), Japan Patent Office (JPO), Korean Intellectual Property Office (KIPO), United States Patent and Trademark Office (USPTO), State Intellectual Property Office of the People's Republic of China (SIPO).

Figure 4: Distribution of top 10% growing codes aggregated at IPC1 level

Degree average growth versus patent application average growth



Note: *authors' calculations on USPTO data, 1973-2012.*

A large part of technological development (which also led the ICT revolution) is related to the development of new materials and apparatus. This is also reflected by the stronger focus put on developing new medical instruments in the recent decades.

4.2 Fitting the diffusion process

In this section we first compare the results obtained by fitting the diffusion paths with the logistic and Gompertz distributions and then discuss the level of maturity expressed by the time to maturity of each IPC code. In particular, we fit the empirical data at the IPC4 level with both functional forms and then select the appropriate one on the basis of the goodness of fit; because the number of estimated parameters in the same for both (two), we select the functional form which provide the lowest residual sum of squares (RSM). Some tests to select between the two distributions have been proposed (e.g. Frances, 1994), however they assume that data has a monotonic behaviour, which is not the case for many of the IPC4 codes.

The results of our test show that the Gompertz is more appropriate in 53% of cases, while the Logistic should be preferred in the remaining 47%. However, it is also interesting to note that the correlation between the year of maturity from the two functions is very high ($\rho = 0.926$) and both yield the same year of maturity in 15% of cases.

Table 1 shows descriptive statistics in order to compare between the two functional forms and to assess the overall results obtained by systematically selecting the best fitting function (BestFit). Times to maturity obtained from the Logistic tend to be longer than those estimated with the Gompertz. This is particularly evident when considering the tenth percentile (p10) of the time to maturity distribution: in case of the logistic the corresponding value is positive (not yet mature), while in the

Gompertz case it is negative suggesting that at least 10% of technologies have already reached their maturity phase. In both cases there is a number of technologies for which the time to maturity is estimated at 157, this is due to the fact that we compute diffusion patterns over a theoretical period of 200 years. For technologies with a flat/decreasing diffusion profile it is not possible to estimate a maturity year and therefore this is set equal to 157.[10] However, the times to maturity obtained from the Gompertz show higher variability than those obtained from the Logistic. By systematically selecting the best fitting curve we balance the results from the two curves.

The goodness of fit is particularly good, with an average R-squared above 0.9 for both the Logistic and the Gompertz curves. The median value is even higher than the average one and the interquartile range (difference between the 25th and 75th percentile) is really small; this suggests that it is only for a very low number of IPC4 codes that the two curves do not perform well.[11]

Table 1: Times to Maturity, Goodness of fit and Diffusion Rates for IPC4 codes

Results for Logistic, Gompertz and their combination

Descriptive Statistics	Maturity			R-squared			Diffusion rate		
	Logistic	Gompertz	BestFit	Logistic	Gompertz	BestFit	Logistic	Gompertz	BestFit
Average	60.9	54.0	56.2	0.924	0.932	0.933	2.5	2.1	2.4
p10	1	-16	-5	0.864	0.875	0.877	0.2	0.3	0.3
Median	45	39	41	0.959	0.959	0.960	2.1	1.8	2.0
p90	157	157	157	0.984	0.984	0.984	5.1	4.3	4.8
Coeff. Var.	0.91	1.15	1.04	0.13	0.10	0.10	0.91	0.84	0.91
IQR	83	100	88	0.051	0.047	0.048	2.8	2.3	2.5

Note: *authors' calculations on USPTO data, 1973-2012.*

Finally the average diffusion rate, calculated in the last year available, is slightly higher than 2 (Degrees per year). Consistent to the time to maturity almost 10% of the technologies are not diffusing or are diffusing at a really low rate.

In Table 2 we present the distribution of technologies based on their stage of maturity. Based on the results obtained previously with the best fit we attempt to single out technologies which may reach maturity in the next three decades. About 15% of IPC4 codes have been classified as already mature. The most interesting cases are represented by the IPC4 codes which are expected to mature within the next 10 years. This period represents a reasonable interval to get reliable estimates.

These correspond to about 11% of IPC4 classes and are reported in the Appendix (table A.1), ordered by the estimated diffusion rate. On the top of the table we find: *i)* wind motors - F03D; *ii)* biocidal, pest- repellant, pest attractant or plant growth regulatory activity of chemical compounds or preparations - A01P; *iii)* analogue

---

[10] In these cases the resulting fits resemble a straight line.

[11] This is further supported by plotting and visually checking the data and the corresponding fits for each IPC4 code. In the few cases where the fit is poor data appears to be randomly scattered.

computers - G06G; *iv)* apparatus for enzymology or microbiology - C12M; and, *v)* propulsion of electrically- propelled vehicles - B60L. Among these codes F03D and C12M were also classified as long run fast growing technologies by Evangelista et al. (2015).[12]

Table 2: Classifying technologies based on their time to maturity

Stage	Share
Mature	15.3%
t <= 10	11.1%
10< t <= 20	9.5%
20 < t <= 30	7.4%
Other	56.7%

Note: *authors' calculations on USPTO data, 1973-2012.*

In figure A.1 we also provide an example of fitting for nanotechnologies and data processing systems or methods, where the Logistic and Gompertz are projected in time. These technologies are good examples of diffusing technologies at different rates. Nanotechnologies are of particular interest, because they are relatively new with the first filing in 1987 and the pace at which they are combined with other technologies is rapidly growing.

About 17% of the IPC4 codes are expected to reach maturity between the next ten and thirty years. In table 2, codes associated to times to maturity longer than thirty years or for which is not possible to compute a change in sign of the second derivative within the time span considered are classified as others.

5. Conclusions

In this paper, we considered diffusion as an attribute defining the spread of technologies from a technology production point of view. More specifically, diffusion of a technological field was proxied by the number of ways it is combined with others within patent documents. The main aim was to analyse diffusion patterns in order to provide new evidence on the technological development process, which we think is crucial for detecting the rise of new technological paradigms.

During the 1973-2012 period, we observe an increase in the average degree (number of connections), which suggests that the density of the technological knowledge network increased. This implies an increase in the number of combinations for each technology, which suggest an increased complexity in the development of new technological applications and possibly of new technological knowledge. This

[12] These technologies were identified among those with the highest growth in term of patent applications for 2 consecutive periods (over the 1992/95-2008/11 interval) at the EPO.

increased complexity poses new challenges on creating (the right) high qualified jobs profiles and calls for the design of educational policies that will enable people to adapt to the upcoming technological paradigm(s). This can be also linked to the idea of an increasing complexity of the product space discussed by Hidalgo et al. (2007) where it is argued that more-sophisticated products are developed in countries/regions which form a densely connected core. In this framework, countries/regions *"move through the product space by developing goods close to those they currently produce"* (p. 482).

Our approach of technological diffusion as a combinatorial process provides results which complement those obtained by patent counting. However, by focusing on connections among technologies our results should be less affected by differences in patent propensity across industries/technologies. A simple comparison between the combinatorial and patent counting perspective suggest that the former gives more weights to "Human necessities" and "Performing operations; Transporting" (which includes materials) related technologies, rather than to "Physics" and "Electricity" related codes as the latter. However, the two approaches are not completely unrelated and some of the results are overlapping. We believe that our framework of technological diffusion can support policy making by focusing on a selected group of technologies that are expected to become central in the technological development in the next years.

To this end we model the technological diffusion process with two functional forms normally used in the literature, the Logistic and the Gompertz. Consistently with previous literature we found that many technologies follow a similar pattern of progress but at different diffusion rates. Our empirical application shows that the two distributions generally provide very good fits and that there is not a one-fits-all (better) distribution to apply to all technological fields. In fact, based on the goodness of fit we selected the Gomperzt in 53% of cases and the Logistic in the remaining 47%. However, in most cases the results we obtain do not change significantly when selecting one over the other, the main differences being that the Logistic tends to estimate a longer time to maturity. The time of maturity was calculated for all technologies assuming a maximum degree equal to the observed historical one. This is an oversimplification of the diffusion phenomenon because the maximum degree may depend on the specific technology. Defining the maximum for the diffusion process is a notoriously difficult problem and we are currently assessing to what extent it is possible to set maximum degrees specific to each technological macro class. Based on the calculated time of maturity, technologies were classified to identify those that show some potential for maturity in the next decade.

The identification of a narrow set of promising technologies is of particular interest for policy making and can allow the design of targeted and effective Research and Innovation (as well as Industrial) policies. This contribution represents a first step in directly evaluating the diffusion of technologies and its importance in the creation of new technological knowledge. Highly diffusing technologies may be linked to enabling and emerging technologies, however more studies are required to understand them and the way they diffuse in the technology production. Further analysis of the combinatorial structure of each technology from a network perspective can give insights on the actual products related to them and the way new technological combinations arise.

"Knowledge is not simply another commodity. On the contrary. Knowledge is never used up. It increases by diffusion and grows by dispersion."
(D.J. Boorstin)

References

Bar, T., Leiponen, A. (2012). A measure of technological distance. *Economics Letters*, 116(3), 457-459.

Berger, R.D. (1981). Comparison of the Gompertz and Logistic Equations to Describe Plant Disease Progress. *Phytopathology*, 71(7), 716-719.

Boon, W., Moors, E. (2008). Exploring emerging technologies using metaphors–a study of orphan drugs and pharmacogenomics. *Social Science & Medicine*, 66(9), 1915-1927.

Breschi, S., Lissoni, F., Malerba, F. (2003). Knowledge-relatedness in firm technological diversification. *Research Policy*, 32(1), 69-87.

Dernis, H., Squicciarini, M., de Pinho R. (2015a). Detecting the emergence of technologies and the evolution and co-development trajectories in science (DETECTS): a 'burst' analysis-based approach. *The Journal of Technology Transfer*, 1-31.

Dernis, H., Dosso, M., Hervás, F., Millot, V., Squicciarini, M., Vezzani, A. (2015b). World Corporate Top R&D Investors: Innovation and IP bundles. A JRC and OECD common report. Luxembourg: Publications Office of the European Union.

Diebold, F.X. (2001). Elements of Forecasting (2nd ed.). Cincinnati: South Western. p. 254.

Dosi, G. (1982). Technological paradigms and technological trajectories. A suggested interpretation of the determinants and directions of technical change. *Research Policy*, 11(3), 147-162.

Dosi, G. (1988). Sources, procedures and microeconomic effects of innovation. *Journal of Economic Literature*, 26(3), 1120-1171.

Dosi, G., Nelson, R.R. (2013). The evolution of technologies: an assessment of the state-of-the-art. *Eurasian Business Review*, 3(1), 3-46.

Evangelista, R., Meliciani, V., Vezzani, A. (2015). The Specialisation of EU Regions in Fast Growing and Key Enabling Technologies (No. JRC98111). Institute for Prospective and Technological Studies, Joint Research Centre.

Franses, P.H. (1994). A method to select between Gompertz and Logistic trend curves. *Technological Forecasting Social Change*, 46, 45–50.

Granger, C.W.J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37, 424-438.

Griliches, Z. (1957). Hybrid corn: an exploration in the economics of technical change. *Econometrica*, 25(4), 510-522.

Hidalgo, C.A., Klinger, B., Barabási, A.L., Hausmann, R. (2007). The product space conditions the development of nations. *Science*, 317(5837), 482-487.

Jones, B.F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies*, 76(1), 283-317.

Krafft, J., Quatraro, F., Saviotti, P.P. (2011). The knowledge-base evolution in biotechnology: a social network analysis. *Economics of Innovation and New Technology*, 20(5), 445-475.

Mansfield, E. (1963). Intrafirm Rates of Diffusion of an Innovation, *Review of Economics and Statistics*, 45, 348-354.

Martino, J.P. (2003). A review of selected recent advances in technological forecasting. *Technological Forecasting and Social Change*, 70(8), 719-733.

Porter, A.L., Roessner, J.D., Jin, X.-Y., Newman, N.C. (2002). Measuring national emerging technology capabilities. *Science and Public Policy*, 29(3), 189-200.

Rogers, E.M. (1983). *Diffusion of innovations*. The Free Press, New York.

Rosenberg, N. (1976). Perspectives on technology. Cambridge: Cambridge University Press.

Rotolo, D., Hicks, D., Martin, B.R. (2015). What is an emerging technology?. *Research Policy*, 44(10), 1827-1843.

Schumpeter, J.A. (1939). *Business cycles. A Theoretical, Historical and Statistical Analysis of the Capitalist Process*. McGraw-Hill Book Company.

Small, H., Boyack, K.W., Klavans, R. (2014). Identifying emerging topics in science and technology. *Research Policy,* 48 (8), 1450–1467

Verhoeven, D., Bakker, J., Veugelers, R. (2016). Measuring technological novelty with patent-based indicators. *Research Policy,* 45(3), 707-723.

Wasserman, S., Faust, K. (1994). Social network analysis: Methods and applications. Cambridge university press.

Weitzman, M.L. (1996). Hybridizing growth theory. *The American Economic Review,* 86(2), 207-212.

Weitzman, M.L. (1998). Recombinant growth. *Quarterly journal of Economics*, CXIII(2), 331-360.

Youn, H., Strumsky, D., Bettencourt, L.M.A., Lobo, J. (2015). Invention as a combinatorial process: evidence from US patents. *Journal of The Royal Society Interface*, 12(106), 20150272.

APPENDIX

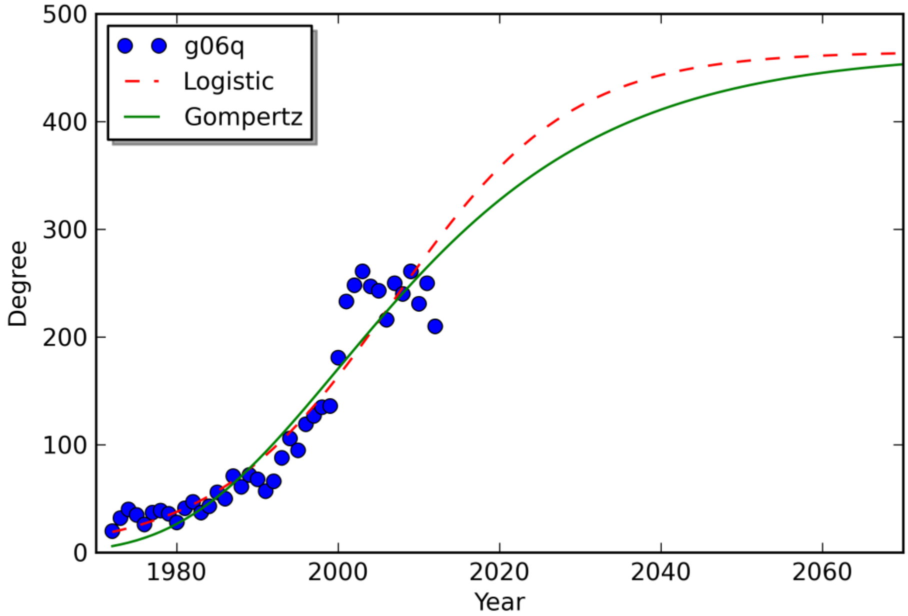
Table A.1: List of IPC4 with estimated time to maturity within 10 years

IPC Code	IPC label
f03d	Wind motors
a01p	Biocidal, pest repellant, pest attractant or plant growth regulatory activity of chemical compounds or preparations
g06g	Analogue computers
c12m	Apparatus for enzymology or microbiology
b60l	Propulsion of electrically-propelled vehicles
b67d	Dispensing, delivering, or transferring liquids, not otherwise provided for
g09g	Arrangements or circuits for control of indicating devices using static means to present variable information
a47b	Tables; desks; office furniture; cabinets; drawers; general details of furniture
b28b	Shaping clay or other ceramic compositions, slag or mixtures containing cementitious material (e.g. plaster)
a62b	Devices, apparatus or methods for life-saving
h04w	Wireless communication networks
f16m	Frames, casings, or beds, of engines or other machines or apparatus, not specific to an engine, machine, or apparatus provided for elsewhere; stands or supports
g01c	Measuring distances, levels or bearings; surveying; navigation; gyroscopic instruments; photogrammetry or videogrammetry
h02j	Circuit arrangements or systems for supplying or distributing electric power; systems for storing electric energy
g01j	Measurement of intensity, velocity, spectral content, polarisation, phase or pulse characteristics of infra-red, visible or ultra-violet light; colorimetry; radiation pyrometry
b82b	Nano-structures formed by manipulation of individual atoms, molecules, or limited collections of atoms or molecules as discrete units; manufacture or treatment thereof
b64c	Aeroplanes; helicopters
b81c	Processes or apparatus specially adapted for the manufacture or treatment of micro-structural devices or systems
c40b	Combinatorial chemistry; libraries, e.g. chemical libraries, in silico libraries
b60q	Arrangement of signalling or lighting devices, the mounting or supporting thereof or circuits therefor, for vehicles in general
h04r	Loudspeakers, microphones, gramophone pick-ups or like acoustic electromechanical transducers; deaf-aid sets; public address systems
e04b	General building constructions; walls, e.g. partitions; roofs; floors; ceilings; insulation or other protection of buildings
g01m	Testing static or dynamic balance of machines or structures; testing of structures or apparatus, not otherwise provided for
b64d	Equipment for fitting in or to aircraft; flying suits; parachutes; arrangements or mounting of power plants or propulsion transmissions in aircraft
b60k	Arrangement or mounting of propulsion units or of transmissions in vehicles; arrangement or mounting of plural diverse prime-movers in vehicles; auxiliary drives for vehicles; instrumentation or dashboards for vehicles; arrangements in connection with cooling, air intake, gas exhaust or fuel supply of propulsion units in vehicles
a01g	Horticulture; cultivation of vegetables, flowers, rice, fruit, vines, hops, or seaweed; forestry; watering
h02p	Control or regulation of electric motors, electric generators or dynamo-electric converters; controlling transformers, reactors or choke coils
g03b	Apparatus or arrangements for taking photographs or for projecting or viewing them; apparatus or arrangements employing analogous techniques using waves other than optical waves; accessories therefor
g01l	Measuring force, stress, torque, work, mechanical power, mechanical efficiency, or fluid pressure
a01n	Preservation of bodies of humans or animals or plants or parts thereof (preservation of food or foodstuff a23); biocides, e.g. as disinfectants, as pesticides or as herbicides (preparations for medical, dental or toilet purposes which kill or prevent the growth or proliferation of unwanted organisms a61k); pest repellants or attractants; plant growth regulators (mixtures of pesticides with fertilisers c05g)
f28f	Details of heat-exchange or heat-transfer apparatus, of general application
g06n	Computer systems based on specific computational models
f25d	Refrigerators; cold rooms; ice-boxes; cooling or freezing apparatus not covered by any other subclass
b62d	Motor vehicles; trailers

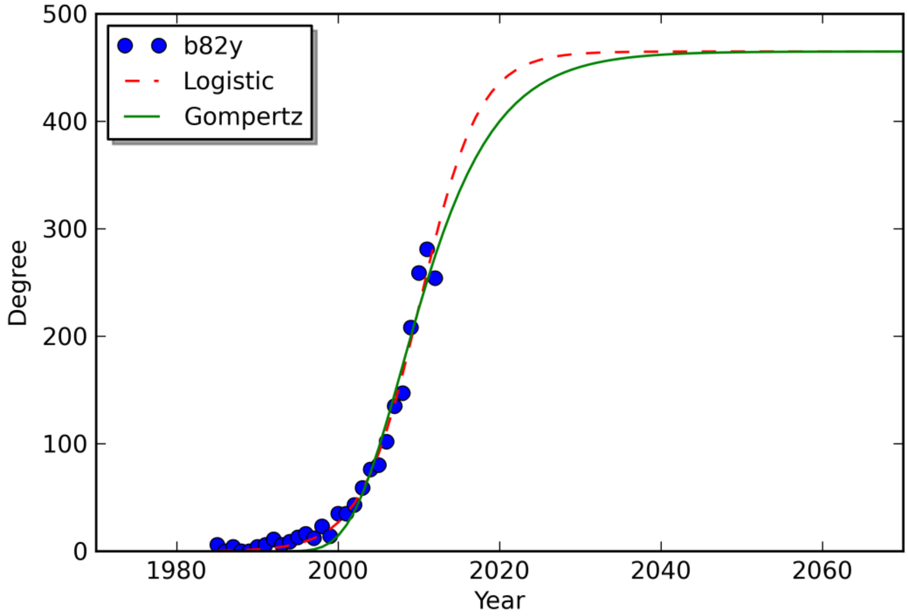
c07h	Sugars; derivatives thereof; nucleosides; nucleotides; nucleic acids
e04h	Buildings or like structures for particular purposes; swimming or splash baths or pools; masts; fencing; tents or canopies, in general
h05h	Plasma technique (ion-beam tubes h01j 27/00; magnetohydrodynamic generators h02k 44/08; producing x-rays involving plasma generation h05g 2/00); production of accelerated electrically- charged particles or of neutrons (obtaining neutrons from radioactive sources g21, e.g. g21b, g21c, g21g); production or acceleration of neutral molecular or atomic beams
f04b	Positive-displacement machines for liquids; pumps
g01f	Measuring volume, volume flow, mass flow, or liquid level; metering by volume
a41d	Outerwear; protective garments; accessories
c08f	Macromolecular compounds obtained by reactions only involving carbon-to-carbon unsaturated bonds
h01b	Cables; conductors; insulators; selection of materials for their conductive, insulating or dielectric properties
e05b	Locks; accessories therefor; handcuffs
b60n	Vehicle passenger accommodation not otherwise provided for
e06b	Fixed or movable closures for openings in buildings, vehicles, fences, or like enclosures, in general, e.g. doors, windows, blinds, gates
f21s	Non-portable lighting devices or systems thereof
a45d	Hairdressing or shaving equipment; manicuring or other cosmetic treatment
e21b	Earth or rock drilling (mining, quarrying e21c; making shafts, driving galleries or tunnels e21d); obtaining oil, gas, water, soluble or meltable materials or a slurry of minerals from wells
h01s	Devices using stimulated emission
a61j	Containers specially adapted for medical or pharmaceutical purposes; devices or methods specially adapted for bringing pharmaceutical products into particular physical or administering forms; devices for administering food or medicines orally; baby comforters; devices for receiving spittle
a61c	Dentistry; apparatus or methods for oral or dental hygiene
b29l	Indexing scheme associated with subclass b29c, relating to particular articles
a63h	Toys, e.g. tops, dolls, hoops, building blocks
a61g	Transport, personal conveyances, or accommodation specially adapted for patients or disabled persons (appliances for aiding patients or disabled persons to walk a61h 3/00); operating tables or chairs; chairs for dentistry; funeral devices
e04f	Finishing work on buildings, e.g. stairs, floors
g09b	Educational or demonstration appliances; appliances for teaching, or communicating with, the blind, deaf or mute; models; planetaria; globes; maps; diagrams
b29d	Producing particular articles from plastics or from substances in a plastic state
b60j	Windows, windscreens, non-fixed roofs, doors, or similar devices for vehicles; removable external protective coverings specially adapted for vehicles
f16d	Couplings for transmitting rotation
c11d	Detergent compositions; use of single substances as detergents; soap or soap-making; resin soaps; recovery of glycerol
g07f	Coin-freed or like apparatus
c08g	Macromolecular compounds obtained otherwise than by reactions only involving carbon-to-carbon unsaturated bonds
f16h	Gearing
g01v	Geophysics; gravitational measurements; detecting masses or objects; tags
b29k	Indexing scheme associated with subclasses b29b, b29c or b29d, relating to moulding materials or to materials for reinforcements, fillers or preformed parts, e.g. inserts
g01p	Measuring linear or angular speed, acceleration, deceleration or shock; indicating presence or absence of movement; indicating direction of movement
a61q	Specific use of cosmetics or similar toilet preparations
b41m	Printing, duplicating, marking, or copying processes; colour printing
b22f	Working metallic powder; manufacture of articles from metallic powder; making metallic powder (making alloys by powder metallurgy c22c); apparatus or devices specially adapted for metallic powder
g10k	Sound-producing devices (sound-producing toys a63h 5/00); methods or devices for protecting against, or for damping, noise or other acoustic waves in general; acoustics not otherwise provided for

Figure A.1: Examples of diffusion process fitting

Data processing systems or methods[13]



Nanotechnologies



[13] The full label is: *data processing systems or methods, specially adapted for administrative, commercial, financial, managerial, supervisory or forecasting purposes; systems or methods specially adapted for administrative, commercial, financial, managerial, supervisory or forecasting purposes, not otherwise provided for.*

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.