The release of this working paper has been authorised by Andrew Wyckoff, OECD Director for Science, Technology and Innovation and by Anthony Cox, OECD Director for Environment.

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the authors. Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcomed, and may be sent to Directorate for Science, Technology and Innovation, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

Note to Delegations:
This document is also available on O.N.E under the reference code:
DSTI/STP/BNCT(2017)2/FINAL

This document, as well as any data and any map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

© OECD (2018)
You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for commercial use and translation rights should be submitted to rights@oecd.org.
Note by the BNCT and NESTI Secretariats

This Report is a collation of the work conducted by both the BNCT and the NESTI secretariat in collecting and analysing statistics and indicators in biotechnology and nanotechnology.

The work presented in this “Report on Statistics and Indicators of Biotechnology and Nanotechnology” is in part based on indicators of biotechnology and nanotechnology that have been previously collected and analysed by the OECD (e.g. patent statistics), and in part on new exploratory indicators (e.g. text-mining analysis) that have been chosen to delineate the two technology fields and to potentially help with the identification of areas of convergence.
The "Report on Statistics and Indicators of Biotechnology and Nanotechnology" brings together the latest available patenting and bibliometric activity data on biotechnology, nanotechnology and related emerging and converging technologies. The data has been collected by the secretariat to the OECD Working Party on Biotechnology, Nanotechnology and Converging Technologies (BNCT).

The OECD Working Parties on Biotechnology (OECD WPB) and Nanotechnology (OECD WPN), in collaboration with the NESTI, have for several years set, measured and published indicators and statistics for biotechnology and nanotechnology (van Beuzekom, 2001[1]; van Beuzekom and Arundel, 2006[2]; van Beuzekom and Arundel, 2009[3]; Palmberg, Dernis and Miguet, 2009[4]). With the creation of the new BNCT in January 2015, the OECD re-confirmed indicators and measurement methodology for these multidisciplinary and partially overlapping technologies with a view to establishing uniquely accurate and relevant datasets.

In order to guarantee comparability both to traditionally measured statistics (backward comparability) and within the context of current national and supranational developments focused on a specific subject, indicators must be identified on the basis of their policy relevance, analytical soundness, statistical quality and measurability. These criteria are of specific importance when setting new ‘experimental’ indicators, aiming to indicate new trends and dynamics, and to stimulate and focus new policy debates.

This report combines highlights of the indicators and statistics for both biotechnology and nanotechnology, collated and analysed by the BNCT. The bulk of the data will be made available for self-searching and download via established OECD portals, such as the Innovation Policy Platform (IPP) (www.innovationpolicyplatform.org/content/statistics-ipp), and the BNCTs website (www.innovationpolicyplatform.org/cstp/bnct), as well as the main OECD website (www.oecd.org).

The report provides a summary of trends and developments in biotechnology and nanotechnology over the last 20 to 25 years, based on the longitudinal analyses of patents. In addition, an experimental indicator, which aims to identify trends and developments of research areas of biotechnology and nanotechnology, was introduced: publication titles of the two technology fields were analysed using a text-mining approach, and an overlay map of the most often use keyword co-occurrences was created, in order to visualise the developments over time in each research area.

Structure of the document

This report is divided into two parts: the first part of the report provides patent statistics of biotechnology and nanotechnology, and the second part of the report introduces a new exploratory indicator on keywords used in articles published in pre-defined biotechnology and nanotechnology journals.
Acknowledgements

This report was prepared by Steffi Friedrichs of the BNCT secretariat with the OECD’s Directorate for Science, Technology and Innovation. Brigitte van Beuzekom of the NESTI secretariat provided valuable input and helped with the access to the bibliometric data. Hélène Dernis of the STI Micro-data Lab provided the patent data, as well as valuable input to the report. Fernando Galindo-Rueda of the NESTI secretariat supported this report with guidance and valuable input.
Table of Contents

Note by the BNCT and NESTI Secretariats ................................................................. 3
Foreword ...................................................................................................................... 4
Executive Summary ................................................................................................... 8
1. Introduction ............................................................................................................ 10
2. Patent statistics of biotechnology and nanotechnology ........................................ 11
3. Developments in the contents of the fields of biotechnology and nanotechnology .... 25
4. Conclusions .......................................................................................................... 48

Tables
Table 3.1. Selected Subject Areas and Subject Categories of the SCImago Journal & Country Ranking 26

Figures
Figure 2.1. Development in biotechnology patent applications, 1990-2012 ....................... 12
Figure 2.2. Development in biotechnology patent applications in the BRIICS, 1990-2012 ..... 13
Figure 2.3. Share of biotechnology patents by main economies patenting in biotechnology, 1990-2012 14
Figure 2.4. Development in nanotechnology patent applications, 1990-2012 .................... 14
Figure 2.5. Share of nanotechnology patents by main economies patenting in nanotechnology, 1990-2012 ................................................................. 15
Figure 2.6. Revealed Technological Advantage (RTA) index in biotechnology and nanotechnology, 1990-2013 ................................................................. 17
Figure 2.7. Share of biotechnology patents by sub-areas, 1990-2012 .............................. 18
Figure 2.8. Share of nanotechnology patents by sub-areas, 1990-2012 .......................... 19
Figure 2.9. Selected specialisation indices in biotechnology sub-areas for EU28, Japan, USA and Germany, 1990-1993 and 2010-2012 ........................................... 20
Figure 2.10. Selected specialisation indices in biotechnology sub-areas for Korea, China, Australia and the Netherlands, 1998-2001 and 2010-2012 ................................. 22
Figure 2.11. Selected specialisation indices in nanotechnology sub-areas for EU28, Japan, USA and Korea, 1990-1993, 1998-2001 and 2010-2012 ................................. 24
Figure 3.1. Changing occurrence of the most important keywords in biotechnology literature, 1996–2014 ........................................................................................................ 28
Figure 3.2. Changing occurrence of the most important keywords in nanotechnology literature, 1996–2014 .......................................................................................... 29
Figure 3.3. Keywords in the field of biotechnology, 1996-2014 ...................................... 30
Figure 3.4. Time-series of the developments of keywords in the field of biotechnology, 1996, 2003, 2010 and 2014 .......................................................... 32
Figure 3.5. Keywords in the field of nanotechnology, 1996 - 2014 .................................. 37
Figure 3.6. Time-series of the developments of keywords in the field of nanotechnology, 1996, 2001, 2007 and 2014 ......................................................... 38
Figure 3.7. Time-series of the developments of keywords in the combined fields of biotechnology and nanotechnology, 1996 and 2014 ......................................................... 43
Figure 3.8. Convergence between the fields of biotechnology and nanotechnology, 1996 - 2014 ...... 46
Figure 3.9. Changes in the fields of biotechnology and nanotechnology, 1996 - 2014 ..................... 47
Executive Summary

The number of biotechnology and nanotechnology patents filed between 1990 and 1999 follows the trend in the increase of all patent numbers filed worldwide. This trend has been assigned to decision of many universities to commercialise their intellectual property. In the case of biotechnology, patent filing numbers have plateaued since 2000; the only exception to this was a three-year (i.e. 2000-2002) spur of patent filing activity by inventors from the United States, which has been assigned to an increased interest in applications in the field of genetics, and which terminated with the completion of the Human Genome Project in 2003. In the case of nanotechnology, however, the numbers of patents filed worldwide continued to increase after 2000.

In contrast to the field of biotechnology, where the top patenting economies all follow a similar trend in patenting behaviour, the trend in the number of nanotechnology patents filed by the one major patenting economy differs significantly from that of another: in this context, it is noteworthy that Japan, which in 1990 filed over 50% of all nanotechnology patents, since 1998 lost a significant part of its global nanotechnology patent share to a surge in patenting by Korea, the People’s Republic of China (hereafter “China”) and Chinese Taipei.

An analysis of patent applications by sub-area shows that the spread of biotechnology patents over eight sub-areas has been remarkably stable since 1990; the only notable change being the slowly growing population of the sub-area on “bioinformatics” since 1999, and the slight decrease in the area of “environmental biotechnology” since 1993. In the case of nanotechnology, however, the population of any of the nine sub-areas is subject to more changes than that of biotechnology. Assignments to the sub-areas of “information processing, storage or transmission” and “nano-magnetism” have dropped over the past 10 years, while an increasing number of patents is assigned to the sub-areas of “nano-optics” and “manufacture or treatment of nanostructures”.

The technology specialisation in sub-areas of biotechnology and nanotechnology shows some significant changes in some economies: the field of “environmental biotechnology” was dominated by Korea during the time period of 1998-2001 and 2002-2005. Since 2006, however, most patents in this sub-area are filed by Japan and China. With the exception of the EU28 countries, which exhibit a technology specialisation in “nanobiotechnology / nano-medicine”, no other economy shows a pronounced and specialisation over a long stretch of time, encompassing several of the measured four-year periods.

Text-mining analyses of scientific publications in journals assigned to the subject categories of biotechnology and nanotechnology provide an insight into the transformation of the science content landscapes of both technologies over the last 20 years: while the field of nanotechnology underwent a shift away from its original focus on metallurgical characterisation towards an application-oriented research on high-tech devices and tools, the field of biotechnology refocussed on a number of concepts that overlap with the field of nanotechnology. Time-series of 2D network maps of
biotechnology and nanotechnology keywords show that the area of overlap between the two fields is increasing, identifying an area of convergence of the two technologies. In this area, nanotechnology increasingly takes the role of an enabling technology to biotechnology, in that the latter applies devices and tools based on materials and concepts enabled by the former.
1. Introduction

Biotechnology and nanotechnology are each fields of science and technology in their own right. They also provide increasingly widely adopted tools for use in other fields of science and technology, including each other. This complex pattern of deployment signifies a growing convergence of traditional disciplines of science and engineering, and poses a formidable challenge to the creation and measurement of reliable indicators of the technologies’ development and impact.

While the trend towards so-called ‘converging technologies’ is increasingly felt by those involved with science and technologies, it has been difficult to characterise or describe. Data acquired for a set of indicators that is equally applicable to and comparable between a number of technologies, such as biotechnology and nanotechnology (and including their contributing disciplines of biology, chemistry, physics and engineering) could help describe and quantify the converging technologies.
2. Patent statistics of biotechnology and nanotechnology

In the absence of economic data on the number, nature and economic impact of bio- and nanotechnology-enabled processes and products, patent statistics provide a reasonable proxy to R&D output/innovation, and enable to observe activity in specific emerging fields (OECD, 2009[5]). Patent data contain information on the intended industrial use of inventions based on very detailed technology classification systems, the most common and widely used of which is the International Patent Classification (IPC) system. This system allows the analysis of trends in individual sub-areas of main application fields in both biotechnology and nanotechnology.

All patent statistics in this report rely on patent applications, of which some may have been granted already. Patents of the fields of biotechnology and nanotechnology were identified using the IPC system: one or several classification codes are attributed to a patent during its examination. The IPC classification is revised periodically to account for the emergence of new technologies and the evolution of existing ones. This may lead to the reclassification of patents into different classes and to the absence of specific classes. Biotechnology patents were identified using a list of IPC codes, which was last revised in 2016, as part of the review of the statistical definitions of biotechnology and nanotechnology. Nanotechnology Patents were identified using the patent tagging system Y01N, developed in 2003, and since then revised into a set of IPC Subclasses in 2011 (EPO, 2013[6]).

All patent counts are based on IP5 patent families, according to the earliest filing date and inventor's residence, using fractional counts.7

IP5 patent families correspond to patent families with at least two members, one of which filed within the Five IP offices (i.e. the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), The State Intellectual Office of the People’s Republic of China (SIPO), and the US patent and Trademark Office (USPTO)) (Dernis et al., 2015[7]).

2.1. Patent counts in biotechnology and nanotechnology

The overall development of the number of patents in biotechnology from 1990 to 2012 is provided in Figure 2.1; the graph indicates that the number of patents in biotechnology peaked between 2000 and 2002, and subsequently decrease by about 10%, before plateauing at a level of around 9 000 filings per year. Since 2008, the trend even seems to suggest a slight decrease in biotechnology patent filings.

The strong increase in biotechnology patents filed between 1990 and 1999 follows an overall trend in the increase of all patent numbers filed worldwide. Such a strong trend has also been observed for information and communication technology (ICT) patents. In 2014, Kers et al. assigned this trend to a dramatic increase in the intention of universities to commercialise their intellectual property.
Figure 2.1. Development in biotechnology patent applications, 1990-2012

Number of IP5 patent families by inventor’s country and priority date

Note: Patent counts are based on IP5 patent families, according to the earliest filing date, using fractional counts. Patents are allocated to technology domains on the basis of their International Patent Classification (IPC) codes.


Figure 2.1 illustrates that the strong increase in patents in the years 2000 to 2002 in the shape of a double-peak is largely accounted for by a corresponding increase in patent numbers in the United States only; the European Union (EU28) exhibits a much less pronounced double-peak during this time-span. Additional analysis reveals that this temporary steep increase is evenly carried by all nine application sub-areas, so that no specific application of biotechnology can be identified to have caused this sudden and short-lived patenting spur. Nevertheless, the steep increase in biotechnology patent applications during the years 2000 to 2002 has been assigned to an increased international interest in genetic patent applications (Kers et al., 2014[8]).

The observed drop in biotechnology patent applications numbers after 2002 is attributed to the completion of the Human Genome Project in 2003, as well as the increasing public debate on genetic engineering in general and the patentability of genes in particular (Kers et al., 2014[8]). The subsequent tightening of national rules for the granting of patents on genetic material is considered an explanation for the overall drop in genetic and biotechnology patent applications in recent years (van Beuzekom and Arundel, 2009[3]).

Figure 2.2 shows that China, too, experienced a sudden surge in patent numbers filings in 2000; other than in the case of the United States’ double-peaked increase in all sub-areas, however, the Chinese spur is based largely on a four- to five-fold growth in patent numbers in the application sub-areas of ‘Medical Biotechnology’, ‘Molecular Biotechnology’ and ‘Genetic Engineering’ only (see also the relevant discussion of ‘Patenting by Application Sub-Areas’ below). In 2014, Kers et al. found that this increase is caused by the patent filings of two companies only, and suggested that political change and economic development may be the main cause of this spike in patent filings in China. OECD analysis shows that the increase is based on a small set of Chinese companies who filed many patents in biotechnology related to genes or polypeptides encodings in 2000.
Figure 2.2. Development in biotechnology patent applications in the BRIICS, 1990-2012

Number of IP5 patent families by inventor’s country and priority date

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BRIICS (total)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Russian Federation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Africa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Patent counts are based on IP5 patent families, according to the earliest filing date, using fractional counts. Patents are allocated to technology domains on the basis of their International Patent Classification (IPC) codes.

Figure 2.2 furthermore illustrates that the total patent count for all BRIICS economies is heavily influenced by the patent filing statistics of China.

The development in the share in biotechnology patents since 1990 is provided in Figure 2.3; the share in patents lost by the US is mostly due to an increasing patent activity by Korea, China, Chinese Taipei and France, while the United Kingdom and Germany are losing patent share as well.

Comparison with the field of nanotechnology shows a somewhat different order of players and development of overall patent numbers (Figure 2.4 and Figure 2.5): numbers in nanotechnology patents have been rising almost exponentially until around 2000, and afterwards, the growth slowed, but was still steady until about 2009. The drop in nanotechnology patent filings since 2010 is a data truncation that arises from the fact that nanotechnology patents are identified by category tags, which are assigned to patents after a patent has been published.
**Figure 2.3. Share of biotechnology patents by main economies patenting in biotechnology, 1990-2012**

IP5 patent families of selected economies by inventor’s country and priority date

- United States
- Japan
- France
- Korea
- Germany
- United Kingdom
- China
- Canada
- Netherlands
- Chinese Taipei

*Note:* The top layer (in light blue) represents the rest of the world. Patent counts are based on IP5 patent families, according to the earliest filing date, using fractional counts. Patents are allocated to technology domains on the basis of their International Patent Classification (IPC) codes.


**Figure 2.4. Development in nanotechnology patent applications, 1990-2012**

Number of IP5 patent families by inventor’s country and priority date

*Note:* Patent counts are based on IP5 patent families, according to the earliest filing date, using fractional counts. Patents are allocated to technology domains on the basis of their International Patent Classification (IPC) codes. Data from 2012 are incomplete.

The most remarkable difference between biotechnology and nanotechnology patent developments is Japan’s leadership of the field of nanotechnology patent numbers; Japan is typically filing only half as many patents as the United States, and EU28 in biotechnology (Figure 2.1). In the case of nanotechnology, however, Japan continuously filed significantly more patents in nanotechnology than any other economy.

Nevertheless, Japan has been steadily losing its originally over 50% share of all nanotechnology patents, despite a nearly 5-fold increase in published Japanese patents between 1990 and 2008. The reason for the significant redistribution of patent share since around 1998 is a surge in nanotechnology patent filings by Korea, China and Chinese Taipei. Other than in the case of biotechnology, Germany’s and the United Kingdom’s share in nanotechnology patents ranks relatively low, and in the case of Germany, is even decreasing.

2.2. Revealed Technological Advantages in biotechnology and nanotechnology

The Revealed Technological Advantage (RTA) index is commonly used to compare patenting behaviours between economies, without necessarily going into the reasons behind the observed differences (OECD, 2009[5]). The RTA is defined as the share of an economy’s patents in a particular technology field relative to the share of total patents in that economy. The index is equal to zero when the economy has no patents in a given field; is equal to 1, when the economy’s share in the sector equals its share in all fields (no specialisation); and above 1 when a positive specialisation is observed.

Figure 2.6 shows selected economies’ normalised RTA (i.e. the reference is now ZERO, instead of 1) in order to compare economies’ specialisation in the two technologies over
the six individual three-year RTA values. The resulting graph highlights significant differences in specialisation between a selected number of economies:

- In the case of biotechnology, quite a few economies show a constant strong RTA over all of the three-year periods analysed since 1990: Denmark’s steadily high RTA in biotechnology (i.e. with a value between 1.7 and 2.6) certifies the economy’s strong specialisation in biotechnology. Similar observations can be made for Australia and New Zealand: these countries’ specialisation is very different from the one observed for China, where a single high RTA in 1998 – 2001 is caused by the patent surge previously discussed.

- In the case on nanotechnology, however, most countries show a low or even negative normalised RTA. These findings indicate that nanotechnology is a very fast moving and changing field, in which many economies file patents in short surges of innovation, and few economies specialise for long enough, in order to develop a patent portfolio in nanotechnology that would yield a steadily positive normalised RTA over a three-year period, as is calculated here.
Figure 2.6. Revealed Technological Advantage (RTA) index in biotechnology and nanotechnology, 1990-2013

Normalised index for selected economies.

Note: Patent counts are based on IP5 patent families, according to the earliest filing date. Patents are allocated to technology domains on the basis of their International Patent Classification (IPC) codes. Nanotechnology data from 2012 are incomplete. Figures for 2010-2013 are Secretariat estimates. Only economies and periods with a minimum number of patents are shown: biotechnology: minimum 100 per period, nanotechnology: minimum 50 per period.
2.3. Patent application sub-areas of biotechnology and nanotechnology

Both biotechnology and nanotechnology patents show a large diversity in their intended applications, ranging from process optimisation to entirely new devices. The classification into sub-areas of each technology (nine in the case of biotechnology and nine (plus one sub-area for ‘not elsewhere classified’) in the case of nanotechnology) thus provides a useful way to further break down the two technology fields with a view to identifying trends in specific industrial application areas.

- Figure 2.7 shows the share of biotechnology patents by application sub-area, illustrating how the sub-area of ‘Bioinformatics’ gains increasing importance since the late 1990s, as more patents are filed with an intended application in this area. Except for this new and growing sub-area, however, the spread of biotechnology patents over all nine sub-areas has been constant since 1990. The dip between 2000 and 2001 in the number of patents in the subcategory of genetic engineering reflects the trend observed and explained in the section “Patent Counts in Biotechnology and Nanotechnology”, discussed above.

Figure 2.7. Share of biotechnology patents by sub-areas, 1990-2012

Note: Patent counts are based on IP5 patent families, according to the earliest filing date, using fractional counts. Patents are allocated to technology domains on the basis of their International Patent Classification (IPC) codes.

- In the case of nanotechnology, by contrast, the development of application sub-areas over time is a lot more dynamic than in the case of biotechnology and shows that patent counts in any given sub-area can decrease and increase over time, without a clear developmental directionality (Figure 2.8).
2.4. Technological specialisation in specific sub-areas of biotechnology and nanotechnology

The Revealed Technological Advantage (RTA) for a specific application sub-area of a technology describes an economy’s specialisation in that sub-area compared to the entire field of the technology. It provides a quick indication of an economy’s focus within a larger technology field, and thereby allows an assessment of policy and infrastructure requirements.

The RTA is defined as an economy’s share of patents in a particular sub-area divided by the economy’s share in the entire technology field. The index is equal to zero when the economy holds no patent in a given sub-area; is equal to 1 when the economy’s share in the sub-area equals its share in the entire technology field (no specialisation); and above 1 when a positive specialisation is observed.

The composite figures (Figure 9 and Figure 10) below show the specialisation in sub-areas of biotechnology for selected economies:

- Figure 2.8 illustrates that before 2010, the EU28 had a slight RTA in “Environmental Biotechnology” (i.e. an RTA of 1.6 in 1990-1993 and an RTA of 1.5 in 1994-1997 (not shown), respectively); since 1998, the EU28 had no significant RTA in any biotechnology sub-area.
Figure 2.9. Selected specialisation indices in biotechnology sub-areas for EU28, Japan, USA and Germany, 1990-1993 and 2010-2012

Indices for selected economies

(a) EU28 (1990-1993) EU28 (2010-2012)

(b) Japan (1990-1993) Japan (2010-2012)

(c) USA (1990-1993) USA (2010-2012)

(d) Germany (1990-1993) Germany (2010-2012)

Note: Patent counts are based on IP5 patent families, according to the earliest filing date, using fractional counts. Patents are allocated to technology domains on the basis of their International Patent Classification (IPC) codes. Only economies with more than 500 patents over the period reviewed are included. The sub-area “Bioinformatics” has been included since 1998.

- Figure 2.9(b) shows that after the period 1990-1993, Japan developed a strong RTA in “Environmental Biotechnology”, which it has since then even enhanced (i.e. RTA (1994-1997): 1.7 (not shown); RTA (1998-2001): 2.1 (not shown); RTA (2010-2012): 2.8).
- The United States does not show any specialisation in a sub-area of biotechnology (see Figure 2.9(c)); in the case of ‘Environmental Biotechnology”, an index below 1 can be observed, indicating that the United States files relatively fewer patents in this sub-area than in the other sub-areas.
• Figure 2.9(d) shows that Germany had little specialisation in 1990-1993, but shows some specialisation in “Environmental Biotechnology” in 2010-2012.
• Far the most pronounced RTA was displayed by Korea: as already illustrated in Figure 2.1, Korea had filed few biotechnology patents until 1997; since 1998, however, biotechnology patents filed by Korean inventors increased year on year. Figure 2.10(a) shows that Korea made its entry as a statistically relevant biotechnology patenting economy with an RTA in “Environmental Biotechnology” of 5.4. Since then, the RTA dropped; in the last measured period (2010-2012), Korea’s RTA in “Environmental Biotechnology” was 1.4.
• China (like Korea) started filing a statistically relevant number of biotechnology patents in the period 1998-2001 (cf. Figure 2.2). As illustrated in Figure 2.10(b), China initially showed a slight specialisation in “Molecular Biotechnology” and “Genetic Engineering” (i.e. RTAs of 1.8 and 1.6, respectively, in 1998–2001), since 2006, China has a pronounced and growing RTA in “Environmental Engineering” (i.e. RTA (2006-2009): 2.5 (not shown); RTA (2010-2012): 2.6).
• Figure 2.10(c) illustrates that in 1998-2001, Australia had a pronounced RTA in “Agricultural Biotechnology” (i.e. an RTA of 2.0 in 1998-2001 and in 2006-2009 (not shown), respectively); more recently, however, the Netherlands have been dominating the specialisation in this sub-area, having reached an RTA of 2.8 in 2010-2012 (see Figure 2.10(d)).
Figure 2.10. Selected specialisation indices in biotechnology sub-areas for Korea, China, Australia and the Netherlands, 1998-2001 and 2010-2012

Indices for selected economies

Note: Patent counts are based on IP5 patent families, according to the earliest filing date, using fractional counts. Patents are allocated to technology domains on the basis of their International Patent Classification (IPC) codes. Only economies with more than 500 patents over the period reviewed are included. The sub-area “Bioinformatics” has been included since 1998.

Comparison between the composite RTA figures shown above identifies Japan’s strong specialisation in “Environmental Biotechnology” (Figure 2.9 and Figure 2.10); indeed, since the late 90s, Japan filed more patents in this sub-area than any other economy. Nevertheless, the overall patenting numbers in this sub-area are so low that even Japan’s specialisation in this sub-area represents around 1% of the economy’s patents in biotechnology only.

It needs to be noted that the United States does not show a pronounced specialisation in any of the eight biotechnology sub-areas.
In the case of nanotechnology, technological specialisation of patent inventor economies is a lot more dynamic than in the case of biotechnology, displaying many changes in technological specialisation between the selected economies and measured periods. This dynamism can in part be explained by the lower number of patents filed each year in nanotechnology (cf. Figure 2.4 and Figure 2.8), and the higher number of sub-areas (i.e. biotechnology is divided into eight sub-areas, while nanotechnology is divided into ten subareas).

The composite Figure 2.11 below shows the specialisation in sub-areas of nanotechnology for selected economies:

Figure 2.11(a) shows that the EU28 group exhibits a specialisation in “Nano-biotechnology / nano-medicine”, as illustrated in Figure 2.11(a), this specialisation was strongest from 1990 until 2001 (i.e. RTA (1990-1993): 2.0; RTA (1994-1997): 2.2 (not shown); RTA (1998-2001): 2.2 (not shown)), and has since then fallen (i.e. RTA (2006-2009): 1.6 (not shown); RTA (2010-2012): 1.2).

Figure 2.11(b) illustrates that Japan initially showed no specific specialisation (i.e. in 1990-1993, the RTAs are below 1.4 in all nanotechnology sub-areas), but has since 1998 developed a slight but steady specialisation in “Nano-magnetism” and “Nano-Optics”.

Similar to the field of biotechnology, the United States does not show any pronounced specialisation in a sub-area of nanotechnology. Nevertheless, Figure 2.11(c) shows a slight specialisation in the sub-areas of “Interacting, sensing or actuating” (i.e. RTA (1998-2001): 1.6 (not shown); RTA (2002-2005): 1.7 (not shown)) and “Nano-Biotechnology / nano-medicine” (RTA (1990-1993) and RTA (1994-1997): 1.5 (1994-1997) (not shown); RTA (1998-2001): 1.4 (not shown); RTA (2002-2005) and RTA (2006-2009): 1.5 (both not shown); RTA (2010-2012): 1.6).

Figure 2.11(d) illustrates that Korea shows a higher specialisation in the nanotechnology sub-area of “Information processing” than any other economy; in the period 1990-1993, Korea’s RTA in the sub-area was 1.7, but has since then dropped steadily (i.e. RTA (2002-2005): 1.3 (not shown); RTA (2006-2009): 1.1 (not shown); RTA (2010-2012): 0.6). Most recently, Korea has developed a specialisation in “Nano-structures manufacture/synthesis” (i.e. RTA (2006-2009): 1.6 (not shown); RTA (2010-2012): 2.1).
Figure 2.11. Selected specialisation indices in nanotechnology sub-areas for EU28, Japan, USA and Korea, 1990-1993, 1998-2001 and 2010-2012

Indices for selected economies

Note: Patent counts are based on IP5 patent families, according to the earliest filing date, using fractional counts. Patents are allocated to technology domains on the basis of their International Patent Classification (IPC) codes. Only economies with more than 500 patents over the period reviewed are included. The sub-area “Bioinformatics” has been included since 1998.
3. Developments in the contents of the fields of biotechnology and nanotechnology

This section of the report is concerned with the exploratory investigation of a potential new indicator of the content development in scientific fields: the indicator follows the longitudinal keyword developments in the scientific fields of biotechnology and nanotechnology over the past 20 years, in order to gain an insight into potential changes in the research content of the two fields of science, and maybe identify areas of convergence between them.

The analysis is based on a text-mining approach, in which the titles of biotechnology articles and nanotechnology articles are combined to give the required sample size (over a specific time period), and subjected to an algorithm that counts the occurrence (and co-occurrence) of each word. The most often occurring words are identified as the most important “keywords” of the given sample size. Visualised on a 2D network maps, the relative position and proximity of the keywords to each other provides insight into the co-occurrence, association and context of the keywords.

Biotechnology and nanotechnology articles were identified using the journal categorisation system developed by SCImago for the Elsevier Scopus® database (Elsevier B.V.) (SCimago, 2007). The SCImago Journal & Country ranking is based on the SCImago Journal Rank (SJR) indicator (Guerrero-Bote and Moya-Anegón, 2012). This categorisation was chosen, because it provides a subject category for both biotechnology and nanotechnology at the same level. Elsevier’s Scopus ASJC (All Science Journal Classification), on the other hand, recognises the field of biotechnology with an ASJC code, but not the field of nanotechnology.

Under the SCImago categorisation, individual periodicals and proceedings in the Scopus® database are identified by ISSN numbers, and categorised into 27 subject areas and 313 specific subject categories (Table 3.1). This categorisation is not mutually exclusive, however; a single title can be assigned to several subject categories.

- “Biotechnology” is a subject category under the subject area of “Biochemistry, Genetics and Molecular Biology”.
- “Nanoscience and Nanotechnology” is a subject category under the subject area of “Materials Science”.

The SCImago subject categorisation is conducted annually on the Scopus® database, so that every source title that features on the database is assigned to one or more subject categories.

One of the SCImago subject areas is specifically assigned to “Multidisciplinary”, containing only one subject category, namely “Multidisciplinary”. This means that biotechnology and nanotechnology publications in prestigious journals, such as “Science” and “Nature” are not captured in the subject categories of “Biotechnology” and “Nanoscience and Nanotechnology” used for the publication statistics presented here.
The relevant lists of journal titles consist to a majority of periodicals and a few proceedings; within these all types of publications are considered. In order to obtain a most complete list of all titles that constitute the SCImago subject categories of “Biotechnology” and “Nanoscience and Nanotechnology”, the titles listed in the annual SCImago journal subject categorisation, which had been done since 1999, were accumulated (1999-2015).

- The list of journals forming the subject category of “Biotechnology” consists of 274 individual titles.
- The list of journals forming the subject category of “Nanoscience & Nanotechnology” consists of 95 individual titles.
- 11 individual titles feature both in the subject category list of “Biotechnology” and “Nanoscience and Nanotechnology”.

### Table 3.1. Selected Subject Areas and Subject Categories of the SCImago Journal & Country Ranking

<table>
<thead>
<tr>
<th>Biochemistry, Genetics and Molecular Biology</th>
<th>Materials Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aging</td>
<td>Biomaterials</td>
</tr>
<tr>
<td>Biochemistry</td>
<td>Ceramics and Composites</td>
</tr>
<tr>
<td>Biochemistry, Genetics and Molecular Biology (miscellaneous)</td>
<td>Electronic, Optical and Magnetic Materials</td>
</tr>
<tr>
<td>Biophysics</td>
<td>Materials Chemistry</td>
</tr>
<tr>
<td>Biotechnology</td>
<td>Materials Science (miscellaneous)</td>
</tr>
<tr>
<td>Cancer Research</td>
<td>Metals and Alloys</td>
</tr>
<tr>
<td>Cell Biology</td>
<td>Nanoscience and Nanotechnology</td>
</tr>
<tr>
<td>Clinical Biochemistry</td>
<td>Polymers and Plastics</td>
</tr>
<tr>
<td>Developmental Biology</td>
<td>Surfaces, Coatings and Films</td>
</tr>
<tr>
<td>Endocrinology</td>
<td></td>
</tr>
<tr>
<td>Genetics</td>
<td></td>
</tr>
<tr>
<td>Molecular Biology</td>
<td></td>
</tr>
<tr>
<td>Molecular Medicine</td>
<td></td>
</tr>
<tr>
<td>Physiology</td>
<td></td>
</tr>
<tr>
<td>Structural Biology</td>
<td></td>
</tr>
</tbody>
</table>

*Note: The Subject Categories ‘Biotechnology’ and ‘Nanoscience and Nanotechnology’ are indicated in grey-shaded cells.*

The SCImago subject categorisation is conducted annually on the Scopus® database, so that every source title that features on the database is assigned to one or more subject categories.8

Text-mining analyses are based on Scopus Custom Data, Elsevier, Version 12.2015, October 2016.

For each year between 1996 and 2014, the articles published in the journals categorised as (a) Subject Category = Biotechnology, and (b) Subject Category = Nanoscience & Nanotechnology, according to the SJR indicator were identified, and the 2 000 most highly-cited documents (i.e. those articles that received the highest number of citations within the given year and subject category) were selected for the analysis.

Time-series of the most often used words in the titles of articles in both biotechnology and nanotechnology journals were created using the methodology and text-mining tool described above: The titles of the 2 000 most often cited articles of biotechnology and
nanotechnology, respectively, per year between 1996 and 2014 were analysed to identify the most often occurring words (i.e. hereafter referred to as “keywords”).

Some of the keywords identified in this analysis are broad and generic (e.g. “property”, “characterisation”, “engineering”, etc.) and could be used in many different technology fields. Nevertheless, in the context of developments within a delineated technology field, changes in their use and contextual appearance indicate changes in the field’s focus of interest. For example, an emerging scientific discipline will commence with investigations and “characterization(s)” of the underlying scientific phenomena and resulting “properties” of a new material, while mature science and technology fields will build on the basic knowledge previously achieved and apply this to the targeted “synthesis” of new materials, such as “nanoparticles” or “graphene”, the analysis of more complex systems, such as the “cell”.

3.1. Developments and changes in keywords of biotechnology and nanotechnology

Figure 3.1 illustrates the development in the occurrence of biotechnology keywords over the past 20 years: in 2002, the keyword “nanoparticle” entered the list of the 10 most often used biotechnology keywords, and increased in importance over the following years. The keyword “activity” on the other hand, which indicates an analytical interest in the properties and behaviour of biomaterials, organisms or molecules, was increasingly less used since 2003, and dropped off the top 10 keyword list entirely in 2008 and 2013. The occurrences of the keywords “bacterium’ and “gene” also steadily dropped over the past 20 years, while that of the keywords “cell” and “detection” increased.
Figure 3.1. Changing occurrence of the most important keywords in biotechnology literature, 1996–2014

Share of the 10 most important biotechnology keywords, based on normalised counts of their occurrences in the titles of scientific articles published in biotechnology journals.

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator.

Source: OECD calculations based on Scopus Customs Data, Elsevier, Version 12.2015, October; text-mining: VOSviewer, version 6.1.3.

Figure 3.2 illustrates the development of the top 10 most often occurring keywords in the field of nanotechnology: in contrast to the field of biotechnology, nanotechnology appears to have undergone some significant changes in the past 20 years, manifesting itself in the loss of keywords and the introduction of new keywords:

- In 1996, the keyword “alloy”, had the second highest occurrence (after “engineering”); by 2009, however, this keyword had dropped entirely off the list of top 10 nanotechnology keywords, and the occurrence of the word “engineering” had decreased significantly.
- In 1999, the keyword “nanoparticle” entered the list of the top 10 nanotechnology keywords, and its occurrence grew steadily; since 2007, it has been the most often occurring word in the titles of nanotechnology articles.
- In 2007, the keyword “graphene” first appeared on the list of the top 10 words in the titles of nanotechnology articles; since then, it has become one of the most often used words in nanotechnology literature.
Figure 3.2. Changing occurrence of the most important keywords in nanotechnology literature, 1996–2014

Share of the 10 most important nanotechnology keywords, based on normalised counts of their occurrences in the titles of scientific articles published in nanotechnology journals

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator.
Source: OECD calculations based on Scopus Customs Data, Elsevier, Version 12.2015, October; text-mining: VOSviewer, version 6.1.3.

3.2. Developments and changes in the content and delineation of biotechnology and nanotechnology

2D network maps were generated, illustrating the relative occurrence of each keyword through the size of their circle size and their label, and the linkages between the keywords through their position and proximity to each other. The 2D network map thereby provides insight into the (co-)occurrence and context of each keyword within a technology fields.

Time-series of 2D network maps were created by first combining the titles of the 2 000 most often cited articles of a technology field (i.e. either biotechnology or nanotechnology) per year over all years (i.e. 1996–2014), in order to create a single sample set for the relevant technology field. The VOSviewer text-mining programme was applied to this sample set, resulting in a 2D network plot of the most often used words in the scientific articles of the relevant technology field. The x,y-position of each word on this map was used to create an anchoring file for the annual 2D network plots of the time-series.

Figure 3.3 shows the 2D network map of the most often occurring words in the titles of biotechnology articles, accumulated over the 1996-2014 period. In the case of biotechnology, 72 words made the threshold of occurring ≥100 times in the 38 000 accumulated titles (1996-2014).
Figure 3.3. Keywords in the field of biotechnology, 1996-2014

2D network map of the occurrences of and linkages and clustering between the most important words in the title of biotechnology articles in 1996-2014

*Note:* Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.

Time-series of 2D network maps are useful visualisation of the developments of the content and boundaries, as well as the potential shift of focus in each technology field. A time-series furthermore allows conclusions about the existing and potentially increasing overlap of the two fields with regard to each other, thus identifying areas of convergence.

Figure 3.4 provides 2D network maps of keywords for the field of biotechnology in the years 1996 (Figure 3.4(a)), 2003 (Figure 3.4(b)), 2010 Figure 3.4(c)) and 2014 (Figure 3.4(d)). Comparison between these network maps over time shows that the field of biotechnology appears to have largely remained unchanged in its overall shape and content. Since 2003, however, the field seems to have focused on far fewer important keywords than before: while the most important biotechnology keywords until 2003 were “gene”, “characterization”, “analysis”, “engineering”, “protein”, “bacterium”, and “cell” (by order of descending occurrence), by 2014 this list had been reduced to “detection”, “cell”, “nanoparticle”, and “analysis” (by order of descending occurrence). This observation is conform with the above mentioned analysis of developments and changes in the top10 keywords used in biotechnology.

In 2003, the keyword “nanoparticle” entered the map of biotechnology keywords, and immediately advanced to one of the three most important concepts of biotechnology; this keyword is furthermore located in close proximity to ‘drug deliver’, because the former is often utilised as a vector in the latter in biomedical and toxicology research.

The concept of “pcr” (i.e. ‘polymerase chain reaction’, an important breakthrough in molecular biology invented in the 1980s, which soon became a routine technique that was applied throughout the field of biological research and -technology) disappeared from the biotechnology keyword map in 2007, and by 2010 the closely related concepts of “isolation” and “quantification” also disappeared from the biotechnology landscape, in line with a gradually decreasing importance of the activity keyword “characterization”.
Figure 3.4. Time-series of the developments of keywords in the field of biotechnology, 1996, 2003, 2010 and 2014

(a) 2D network map of the occurrences of and linkages and clustering between the most important words in the title of biotechnology articles in 1996

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.

(b) 2D network map of the occurrences of and linkages and clustering between the most important words in the title of biotechnology articles in 2003

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.

(c) 2D network map of the occurrences of and linkages and clustering between the most important words in the title of biotechnology articles in 2010

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.
(d) 2D network map of the occurrences of and linkages and clustering between the most important words in the title of biotechnology articles in 2014

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence. 
Figure 3.5 shows the 2D network map of the most often occurring words in the titles of nanotechnology articles, accumulated over the 1996-2014 period: 66 words made the threshold of ≥100 occurrences in the accumulated 38,000 titles.\textsuperscript{14,15} Comparison with Figure 3.3 shows that in the case of nanotechnology, this number of keywords is lower than the 72 words that made the threshold in the case of biotechnology; it is a first indicator of the greater diversity of nanotechnology and the shift that occurred in the field between 1996 and 2014, as illustrated and explained in the following text.

The time-series of nanotechnology keyword developments is provided in the composite Figure 3.6; it shows that the field underwent a continuous shift in research content foci, illustrated by change in the shape of the entire network map of keywords: during the years 1996 to 2001 the scientific study of metallurgy, such as that of the “microstructure” of an “alloy” or a “steel” and their “property” dominated nanotechnological R&D (Figure 3.6(a) and Figure 3.6(b)).

In 2007, “graphene” appeared on the nanotechnology keyword map as a new “material”-associated concept that lies in close proximity to concepts of complex applications and devices, such as the “lithium ion batteries” and “dye sensitised solar cells” (Figure 3.6(c)).

By 2014, the former metallurgy concepts had been all but abandoned and the main focus of the field of nanotechnology had shifted to the “nanoparticle”, the “synthesis” (of “carbon nanotube(s)”), and the “material” “graphene”, as well as to the application of these materials in specific devices, such as “electrode(s)”, “diode(s)”, and “dye sensitised solar cell(s)” (Figure 3.6(d)).

In 2004, the keyword “gold nanoparticle” entered the map in close coordination with the concepts of “biosensor”-base detection, adjacent to the “nanoparticle” and the “cell”. Around this field that describes the application of nanotechnology-enabled materials and devices in life-science applications, an increasing convergence of the two fields of biotechnology and nanotechnology occurs from around 2003/2004 onwards.
Figure 3.5. Keywords in the field of nanotechnology, 1996 - 2014

2D network map of the occurrences of and linkages and clustering between the most important words in the title of nanotechnology articles in 1996

*Note:* Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.

Figure 3.6. Time-series of the developments of keywords in the field of nanotechnology, 1996, 2001, 2007 and 2014

(a) 2D network map of the occurrences of and linkages and clustering between the most important words in the title of nanotechnology articles in 1996

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.
(b) 2D network map of the occurrences of and linkages and clustering between the most important words in the title of nanotechnology articles in 2001

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.
(c) 2D network map of the occurrences of and linkages and clustering between the most important words in the title of nanotechnology articles in 2007

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.
(d) 2D network map of the occurrences of and linkages and clustering between the most important words in the title of nanotechnology articles in 2014

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.
A time-series of 2D network maps of the combined fields of biotechnology and nanotechnology was created, in order to provide an insight into the existing and potentially increasing overlap of the two fields with regard to each other, thus identifying areas of convergence. The time-series was generated by first combining the titles of the 2,000 most often cited articles of a given year of both technology fields, whilst removing the duplicated titles from the list. An anchoring map was created by accumulating the resulting data sets of annual article titles over all 19 years of analysis, and subsequently applying the VOSviewer text-mining programme to the resulting sample set, in order to create a single 2D network map of the most often occurring words in the titles of biotechnology and nanotechnology articles. The x,y-position of each word on this map was used to create an anchoring file for the annual 2D network plots of the time-series.

Figure 3.7 shows the start-point (i.e. 1996; Figure 3.7(a)) and the end-point (i.e. 2014; Figure 3.7(b)) of the time-series of the combined fields of biotechnology and nanotechnology. The figure illustrates the gradual shift of both fields towards increasing overlap and convergence.
Figure 3.7. Time-series of the developments of keywords in the combined fields of biotechnology and nanotechnology, 1996 and 2014

(a) 2D network map of the occurrences of and linkages and clustering between the most important words in the title of biotechnology and nanotechnology articles in 1996

*Note:* Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.

(b) 2D network map of the occurrences of and linkages and clustering between the most important words in the title of biotechnology and nanotechnology articles in 2014

*Note:* Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.

Figure 3.8 shows an annotated network map of the most often occurring words in the titles of the combined fields of biotechnology and nanotechnology, accumulated over 19 years between 1996 and 2014. The figure highlights the area of keywords that stem predominantly from one of the two technology fields, indicating an area of increasing convergence between the two fields. Comparison of Figure 3.8 with the 1996 and the 2014 maps of Figure 3.8 above show that between 1996 and 2014 the fields of biotechnology and nanotechnology did indeed shift closer together.

Figure 3.9 summarises the observed convergence within combined fields of biotechnology and nanotechnology by illustrating the directional changed that each technology field underwent between 1996 and 2014:

- In the case of **biotechnology**, the technology field maintained its overall outline, but underwent a shift in focus away from “analysis” and “characterisation” towards “cell”, “nanoparticle” and “detection”. These observations are in line with the findings of the section “Developments and changes in keywords of biotechnology and nanotechnology” above (cf. Figure 3.1).
- In the case of **nanotechnology**, the entire map of nanotechnology keywords changed shape and shifted between 1996 and 2014: while in 1996 the most often occurring words in the titles of nanotechnology articles were of metallurgical nature, such as “microstructure” and “alloy”, by 2014, these words have disappeared from the list and new words, such as “graphene” and “nanoparticle” show a high occurrence. These observations strengthen the findings described in the section “Developments and changes in keywords of biotechnology and nanotechnology” above (cf. Figure 3.2).

Overall, the text-mining analyses show that the field of nanotechnology is more diverse than that of biotechnology; this is evident from the fact that far fewer keywords in the title of nanotechnology documents make the threshold of ≥100 occurrences in 2 000 titles: in the late 90s, only around 40 nanotechnology keywords make this threshold, compared to around 65 biotechnology keywords. This difference seems to disappear in later years, however: since 2012, the field of biotechnology seems to be united by a set of fewer than 40 keywords that occur at least 100 times per year.
Figure 3.8. Convergence between the fields of biotechnology and nanotechnology, 1996 - 2014

Annotated 2D network maps of the occurrences of and linkages and clustering between the most important words in the titles of biotechnology and nanotechnology articles in 1996-2014 (scaling: large circle/label = high occurrence)

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.

Figure 3.9. Changes in the fields of biotechnology and nanotechnology, 1996 - 2014

Annotated 2D network maps of the occurrences of and linkages and clustering between the most important words in the titles of biotechnology and nanotechnology articles in 1996-2014 (scaling: large circle/label = high occurrence)

Note: Technology fields were defined as subject categories according to the SCImago Journal Rank (SJR) indicator; scaling: large circle/label = high occurrence.

4. Conclusions

Analysis of both traditional and novel statistics and indicators have revealed both similarities in the developments of the two technologies fields, as well as significant differences in their nature.

Although patents data may be considered as an imperfect measure of R&D output, they provide a good proxy of the inventive activity of countries in technology areas. In future work, it could be useful to look at patent applicants in addition to the patent data and analysis on inventors presented in this report.

Novel analytical tools, such as text-mining, have provided results that are commensurate with and supportive of the results obtained from analysing traditional statistics and indicators.

This report and its underlying data collection and analysis mark the starting point to an ongoing analysis of longitudinal statistics for both technology fields. The insights obtained through time-series of patent statistics allow the identification and interpretation of trends in the development of both technologies on the global level and on that of individual economies, thus contributing important cornerstones for impact and policy assessments.

Future work could try to analyse connections between the publication and patenting behaviour of economies, in order to establish potential connections or parallels.

Both biotechnology and nanotechnology are heavily impacted by an increasing globalisation, which manifests itself in the rising number of patent inventions and publications in economies such as China, Korea and Chinese Taipei.

The time-series analyses established in this report is recommended to be continued, in order to provide support to the assessments of impact, policy and technology, as well as to potential foresight exercises.
References


Notes

3. All patent data were compiled using the Intellectual Property Database of the OECD, STI Micro-data Lab: [http://oe.cd/ipstats](http://oe.cd/ipstats).
4. Text-mining was conducted using the freely available VOSviewer programme (version 6.1.5) ([http://www.vosviewer.com/](http://www.vosviewer.com/)). Keyword occurrence were normalised to the overall sample size of the number of individual titles used for each text-mining analysis. Individual word-exclusion and word-transformation tables (i.e. thesaurus tables) were created and checked manually for each longitudinal time-series.
5. The VOSviewer (version VOSviewer 1.6.5) programme was set automatically assign clusters, using the “association strength” method to normalise the strength of the links between the co-occurring words; other programme settings are: 1. Layout: attraction: 2, repulsion: 1, use default values; 2. Clustering: resolution: 1.00, min. cluster size: 1, merge small clusters.
6. [https://www.scopus.com/](https://www.scopus.com/)
7. In some cases, a single journal may have two or three ISSN numbers.
8. One of the SCImago subject areas is specifically assigned to “Multidisciplinary”, containing only one subject category, namely “Multidisciplinary”. This means that biotechnology and nanotechnology publications in prestigious journals, such as “Science” and “Nature” are not captured in the subject categories of “Biotechnology” and “Nanoscience and Nanotechnology” used for the publication statistics presented here.
9. A minimum threshold of \( \geq 5 \) keyword occurrences (at 100% relevance) in the 2000 most cited articles per year per technology field was applied.
10. Text-mining was conducted using the freely available VOSviewer programme (version 6.1.5) ([http://www.vosviewer.com/](http://www.vosviewer.com/)). Keyword occurrence were normalised to the overall sample size of the number of individual titles used for each text-mining analysis. Individual word-exclusion and word-transformation tables (i.e. thesaurus tables) were created and checked manually for each longitudinal time-series.
11. A minimum threshold of \( \geq 100 \) keyword occurrences in the accumulated sample set of the 38 000 most cited articles (i.e. 19 years of 2 000 titles per year) per technology field was applied.
12. The visualisation setting of the VOSviewer programme may render the labels of some words invisible, in order to avoid overlaps.
13. The VOSviewer (version VOSviewer 1.6.5) programme was set to automatically assign clusters, using the “association strength” method to normalise the strength of the links between the co-occurring words; other programme settings are: 1. Layout: attraction: 2, repulsion: 1, use default values; 2. Clustering: resolution: 1.00, min. cluster size: 1, merge small clusters.
14. A minimum threshold of \( \geq 100 \) keyword occurrences in the accumulated sample set of the 38 000 most cited articles (i.e. 19 years of 2000 titles per year) per technology field was applied.
The visualisation setting of the VOSviewer programme may render the labels of some words invisible, in order to avoid overlaps.

A minimum threshold of ≥150 keyword occurrences in the resulting accumulated sample set of the 74,128 most cited articles in the combined fields of biotechnology and nanotechnology was applied.

It needs to be noted that the word “cell” may have different meanings in the field of nanotechnology from that of the field of biotechnology. The data set of the titles of nanotechnology articles was therefore checked manually, in order to analyse the meaning of the word “cell” used in these articles. Manual verification confirms that all occurrences of the term “solar cell” are exclusively counted towards the appearance of the designated keyword “solar cell” on the map; the same is true for the term “dye sensitized solar cell”. The term “cell” covers all remaining meanings of the word, over 95% of these concern the biological cell, while less than 5% may concern material grain cells, fuel cells or similar materials and engineering concepts.