Welcome to the 31st issue of Research Trends.

Issue 31 of Research Trends examines the changing landscape of scholarly communication and scientific scholarship. The scientific world, much like every other aspect of modern life, has been vastly influenced by globalization and technological progress which, in turn, impacts scientific and scholarly exchanges. This impact can be seen through language usage, migration patterns and emerging scientific communities among many others.

In this issue we bring together a collection of articles that touch upon the future of scientific scholarship and communication from different perspectives: the ever increasing importance of English as the language of written communication; new ways to standardize and acknowledge contributorship to research articles; how researchers from scientifically emerging world regions enter the international research arena; and how new international migration patterns may lead to new collaboration networks.

In addition, this issue reports back on two important Research Trends seminars, both held in Washington D.C. in October, one focusing on research migration and brain circulation, and the second on the practical ways in which research evaluation methodologies are used. The collection of thoroughly written articles in this issue clearly illustrate what was underlined by several speakers at the latter seminars, namely, that bibliometrics, or, more generally, informetrics, can and should be used to communicate scientific patterns and trends that go beyond the evaluation of institutions or individual researchers.

We hope you enjoy this issue. Please share your thoughts and feedback with us! You can do this in the comments section following each article on our website (www.researchtrends.com) or by sending us an email (researchtrends@elsevier.com).

Kind regards,

Henk F. Moed
Editor-in-Chief
Fixing authorship - towards a practical model of contributorship

In this contribution, Mike Taylor and Gudmundur A. Thorisson discuss the problems surrounding authorship in research today, and how these can be resolved in this digital age.

The Language of (Future) scientific communication

This article presents a short study examining the use of languages other than English in scientific communication.

FORCE11 gains momentum

Anita de Waard and Maryann Martone discuss the background and aims of FORCE11, dedicated to advancing scholarly communication and e-scholarship.

International scientific migration analysis generates new insights

In this article Gali Halevi and Dr. Henk F. Moed describe researchers’ migration trends and co-authorship patterns, and some of the similarities and differences between the two.

The Rise of Latin American science

Research Trends presents a bibliometric investigation by Sarah Huggett on the rise of Latin American science.

Reporting Back: Research Mobility and Brain Circulation

Gali Halevi reports on a Research Trends seminar held at George Washington University on research mobility.

Reporting Back: Research Evaluation in Practice

A report, by Gali Halevi, on a seminar held at the National Geographic Society on the practical applications of research evaluation methodologies.

Did you know?

...about the growth in online education?
Section 1: Research Assessment

Fixing authorship - towards a practical model of contributorship

Introduction
As we near the completion of the metamorphosis of paper-based scholarly publishing to a medium entirely based on the Internet, so there is increasing need to enrich the environment with a connected network, unfettered by the legacy of putting ink onto paper. One of the more recent areas to come under consideration is issues and concepts of authorship, and how these can be represented in a wholly digital world. For legal and copyright reasons, the concept of 'an author' of a scholarly work is likely to persist for some time. However, the idea that a simple list of authors is the optimum way of recording scholarly achievement has reached the end of its shelf life. It's time to move on.

Anyone who is connected with scholarly publishing knows that there are a variety of tasks that are covered and obscured by the term "authorship", and there are vital research tasks that are not considered to be worthy of the term. Moreover, there are many grey areas: for example, 'guest' authorship - where names appear in author lists of people who have had little or no impact on the research work - and 'ghost' authorship - where legitimate authors do not appear on the author list for reasons of expediency or politics.

Clearly, there cannot be just one resolution for authorship-related problems. However, the study of contributorship - and the development of a standard infrastructure to support more nuanced relationships between researcher and published output - promises to solve the logistical issues, and to illuminate those that have an ethical basis. A prominent example of work in this area is the recent International Workshop on Contributorship and Scholarly Attribution (IWCSA), in which we participated and which recently published its results (1).

Authorship broken, needs fixing
Current definitions of authorship only cover a very limited series of relationships that a person can have with a published article. Typical author lists tend to only include authors and/or editors, with other contributions and relationships being inconsistently indicated via text in an acknowledgements section.

This binomial approach - essentially a relic from the print age - to recognizing contributions to a published scholarly work has many flaws. The Harvard Workshop recognized nine specific issues which are listed in Table 1.

Many readers will be familiar with some or even all of these issues as authors or editors. Here we want to highlight and elaborate on what we consider the most prominent ones.

Varied authorship conventions across disciplines
It often comes as a surprise to find that different disciplines vary in the significance of author order and role. Take, for example, the diverse ways in which the same author order of a fictional paper written by Smith, Taylor and Thorisson might be interpreted depending on discipline (see Table 2).

Increasing number of authors on articles
High Energy Physics (HEP) is well-known for long author lists on research papers, with over 3,000 authors credited in recent extreme cases. This is in part because of the complexity and scale of HEP research, but also because HEP publications tend to give equal weighting to researchers and engineers alike. Clearly, the traditional model of the author as the writer of the work is not being applied in this discipline [2].

---

Mike Taylor
orcid.org/0000-0002-8534-5985
Gudmundur A. Thorisson
orcid.org/0000-0001-5635-1860
Equally, having 1000+ authors on a single paper presents novel logistical problems of managing a non-trivial amount of publication metadata - merely getting all the names and affiliations correct is a significant challenge. In fields other than HEP, there is also a clear trend towards an increased number of authors per published paper. For example, the Wellcome Trust reports that the number of authors on its genetics papers rose from around 10 to nearly 29 between 2004 and 2010. Furthermore, many standard ways of assessing scholarly impact will share the value amongst the authors in an entirely arbitrary manner. This leads to the so-called "dilution effect", whereby even a well-cited paper makes little or no contribution to the metrics for individual authors because credit is "diluted" across the large number of authors.

Inadequate definitions of authorship

There is no universal definition of what is meant by research authorship: the closest that exists are a set of rules drawn up by the International Committee of Medical Journal Editors (ICMJE) (3). These rules have been adapted and used by a number of journals over the last several years, although even the ICMJE itself recognizes that they are outdated (Christine Laine, Editor of Annals of Internal Medicine, reported at IWCSAI).

Inability to identify individual contributions

With any multi-author work, there will be a breakdown of tasks that the individuals listed as authors have contributed to the work. Traditional author lists do not allow for any credit below this level. Many journals now allow (or even require) contributorship statements at the end of the article, but these are rarely in any kind of standardized form that can be processed in automated fashion to inform calculation of impact, expertise or standing. This lack of granularity can lead to the case where a senior researcher who has had little or no influence on a paper can be credited with "proper" authorship, whereas a computer programmer who made a significant contribution via the construction of key algorithms is perhaps not credited at all.

### Table 1: Problems caused by existing authorship practice (Harvard Workshop)

<table>
<thead>
<tr>
<th>Problem identified by Workshop</th>
<th>Resolution approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Varied authorship conventions across disciplines</td>
<td>-</td>
</tr>
<tr>
<td>Increasing number of authors on articles</td>
<td>-</td>
</tr>
<tr>
<td>Inadequate definitions of authorship</td>
<td>-</td>
</tr>
<tr>
<td>Inability to identify individual contributions</td>
<td>-</td>
</tr>
<tr>
<td>Damaging effect of authorship disputes</td>
<td>-</td>
</tr>
<tr>
<td>Current metrics are inadequate to capture and include new forms of scholarship and effort</td>
<td>Altmetrics (e.g., altmetric.com, altmetrics.org, impactstory.org)</td>
</tr>
<tr>
<td>Inability of funders to track the outputs of their funding</td>
<td>Fundref (<a href="http://www.crossref.org/fundref/index.html">http://www.crossref.org/fundref/index.html</a>)</td>
</tr>
<tr>
<td>Name ambiguity leads to misattribution of credit and accountability</td>
<td>ORCID (<a href="http://www.orcid.org">www.orcid.org</a>)</td>
</tr>
<tr>
<td>Aggregation of attribution information from a large number of sources</td>
<td>ORCID (<a href="http://www.orcid.org">www.orcid.org</a>), etc.</td>
</tr>
</tbody>
</table>

### Table 2: Varied authorship conventions across disciplines, illustrated using a fictional paper authored by Smith, Taylor and Thorisson

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Authorship Convention</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Energy Physics</td>
<td>Author list is in alphabetic order, no precedence can be interpreted. Names may include engineers as well as researchers.</td>
</tr>
<tr>
<td>Economics, some fields within Social Sciences</td>
<td>Author list is in alphabetic order, no precedence can be interpreted.</td>
</tr>
<tr>
<td>Life Sciences</td>
<td>Smith the postdoc did most of the experimental work, but Thorisson was the principal investigator who led the scientific direction of the work. The alphabetical order is coincidental.</td>
</tr>
<tr>
<td>&quot;Standard&quot; order</td>
<td>Smith is the senior researcher who did most of the work. Taylor was subordinate to Smith, Thorisson is subordinate to Taylor. The alphabetical order is coincidental.</td>
</tr>
</tbody>
</table>
**Sidebar 1:** Contributorship example

"Smith, Taylor and Thorisson" is a fictional paper. A wider contributorship definition includes Ms Bercow, a Ph.D student, who was heavily involved in the experiment, but who was excluded as none of her tasks traditionally merit authorship. In comparison, Professor Smith had full authorship status, despite not having had significant engagement with this specific paper. Each individual's contribution is linked to their ORCID record.

Bercow (Ph.D student, not previously credited with authorship)
- data collection
- ran the experiment
- wrote algorithm
- created artwork
- literature review

Smith (Laboratory Head)
- lab leadership
- reviewed paper

Taylor (Researcher)
- intellectual conception
- data analytics
- wrote the paper
- submitted the paper
- obtained funding

Thorisson (Senior Researcher)
- intellectual conception
- created experimental methodology
- data analytics
- wrote the paper

**Damaging effect of authorship disputes**

The lack of clarity of authorship claims and credit has led to a growth in authorship disputes and a number of scandals. A detailed and standardized method of declaring contributions is likely to put an end to all but the most egregious of such disputes. The problems revealed by an analysis of author/article relationships fall into two broad categories: logistical (in other words, technical) and ethical. However, these are not conveniently discrete categories: an inability to precisely define the relationship leads to a position whereby a research team is obliged to force classification upon its members. Given that authorship is the principal means of recognizing academic achievement, this is not without weight.

**Contributorship**

We hope that one of the major outcomes of this field of work will be an evidence-based system of classifying relationships between researcher and a published work. Moreover, we hope that this taxonomy will facilitate codification of relationships that go beyond traditional authorship, thus removing the difficult decisions that can arise when compiling an author list. For example, by explicitly allowing “data collection” or “algorithm creation” as a type of contribution, it would be possible to formally attribute credit to members of the team that a strict adherence to authorship conventions (such as they are) would likely ignore, whilst not conflating the precise nature of the researcher’s contribution with intellectual leadership. In the same vein, specifying “Head of research team” or “Principal investigator” would facilitate distinguishing a senior member’s relationship with the work from those who also made intellectual contributions (see Sidebar 1).

Clearly, the answer to this problem goes beyond the creation of a standard - there needs to be an infrastructure for storing these complex relationships, tools to create them and maintain them, and ways of displaying them. Most importantly, the benefits of fully recording these relationships must outweigh (and be seen to outweigh) the cost of the additional complexity and work required (i.e. beyond what is currently the norm).

Software can certainly help in this effort (although the idea of determining who-did-what with a list of 1000+ researchers is overwhelming!) and there have been some very good examples of simple, spreadsheet-based tools in recent proof-of-principle projects. However, the task of apportioning responsibilities (and rewards) can start earlier - perhaps within research tools such as Mendeley.

**Help is coming**

Many of the issues highlighted above are being tackled by a diverse community of agencies and approaches, many of which came together for the IWCSA workshop. Here we want to highlight a particularly important one: the Open Researcher & Contributor ID initiative (ORCID: http://about.orcid.org). Launched in mid-October 2012, the registry service operated by ORCID enables researchers to create a public identity and obtain a persistent personal identifier, and to maintain a centralized record of their scholarly activities (4), (5).

Whilst the basic idea of an online "author profile" is not unique or innovative in itself, several key attributes differentiate the new service from the myriad free and commercial services in this space. First, it is backed by a non-profit, community-based organization with participation from commercial publishers, academic institutions, research libraries, funding agencies and many others. Second, major stakeholders in the ORCID community are committed to building software applications and platforms that will build on and integrate with the central ORCID service for automatically linking scholars and their published works.

At the time of writing, the ORCID service is limited in functionality and is experiencing some early growing pains, but the service is improving over time and with the strong support of the community. Despite these initial teething troubles, several integrations built by ORCID’s launch partners are already operational and more will come online in the next several months.
So what is ORCID’s relevance to the attribution challenges outlined above? Although the first-generation service is functionally limited, the core system has been built to support future developments and definitions that go beyond basic author or editor roles. These can potentially include richer contributorship statements such as the examples already given above. It follows that ORCID can serve as a central index or discovery hub in which to look up not merely the base contributor-work relationship, but also the nuances of that relationship if more detailed information is available.

Conclusions
Definitions are softening: in the new world of online digital publishing, “articles” are more than words on paper, metrics are more than citation counts, usage is more than subscriptions - and authors are more than just writers. The concept of authorship is rooted in our culture and in our minds, and that principle will not go away. But the idea of contributorship offers a richer set of definitions that enable our contributions to human knowledge to be recorded more precisely, if only we are willing to embrace it, and if the tools and infrastructure are developed that allow us to capture this information whilst not increasing administrative burden.

Conflict of interest statement
The authors have both been active contributors to ORCID in the past three years. As of October 2012, one of them (GAT) is employed by ORCID part time to work on the EU-funded ODIN project [http://odin-project.eu].

Contributorship statement
The authors contributed equally to the drafting of this article.

About the authors
Gudmundur ‘Mummi’ Thorisson is an academic and consultant interested in scientific communication, in particular as this relates to open access to and use/reuse of research data in the life sciences. He has been involved in various projects relating to identity & unique identifiers in research and scholarly communication, most recently the ORCID initiative. Through his previous work in the GEN2PHEN project [http://www.gen2phen.org] he has also contributed to several database projects in the biomedical research domain, notably GWAS Central [http://www.gwascentral.org]. Gudmundur holds a PhD from the University of Leicester in the United Kingdom and worked there as a post-doctoral researcher after graduating in 2010. He currently works part time for ORCID on the ODIN project [http://odin-project.eu], whilst also working in a research support role at the Institute of Life and Environmental Sciences (http://luvs.hi.is), University of Iceland, Reykjavik where he is now based.

Personal website: http://gthorisson.name

ORCID profile:
http://orcid.org/0000-0001-5635-1860

Mike Taylor is research specialist in Elsevier Labs and the newest member of the Research Trends Editorial board. His current areas of work include altmetrics, contributorship, research networks, the future of scholarly communications and other identity issues. He has worked in various capacities within the ORCID initiative. Previous to joining Elsevier Labs, Mike worked in various technology and publishing groups within Elsevier.

Website: http://labs.elsevier.com

ORCID profile:
http://orcid.org/0000-0002-8534-5985

References:
2. Aaij, R et al (enormous list of authors) (2012) “Measurement of the ratio of branching fractions B (B 0 → K * 0 γ) / B (B 0 → φ) and the direct CP asymmetry in B 0 → K 0 γ”, Nuclear Physics B, Vol. 867, No. 1, pp. 1-18.
Section 2: Country Trends

The Language of (Future) scientific communication

Dr Daphne van Weijen

English is generally considered to be the lingua franca of the scientific community. For example, roughly 80% of all the journals indexed in Scopus are published in English. The adoption of English as the universal language of science is due in part to historical political and economic factors which favored English over other potential candidate languages such as Chinese, French, German, Russian, or Spanish (1), (2), (3). Indeed, German was actually the favored language in scholarly communication for the first part of the 20th century (4). However, although English is now clearly established as the main language of international scientific communication, researchers continue to publish their work in other languages than English as well. Furthermore, research suggests that the extent to which researchers still publish in their native language, as opposed to English, differs across the disciplines. They seem to be more likely to publish in languages other than English within the Social Sciences, Applied Sciences and Humanities, than in the natural, theoretical and hard sciences (1), (2). This article reports on a short study using Scopus data to determine (a) whether the use of languages other than English for scientific communication is increasing or decreasing, and (b) in which subject fields researchers publish most when publishing in their native languages instead of in English.

The preferred language of publication

In an earlier issue of Research Trends, we published a brief article on the use of English as the international language of science from 1996 to 2007 (3). Results of that study indicated that researchers were more likely to publish their work in English than in their native language in most of the Western European countries included in the sample. The ratio for English to Dutch and English to Italian publications was particularly high, compared to those of the other countries in the study (German, France, Spain and the Russian Federation). However, please note that Scopus covers non-English language journals only if they include English article titles and abstracts. We decided to replicate this analysis, to determine whether this trend has continued in these countries over the past four years.

As in the earlier study (3) published in 2008, the ratios of the number of journal articles published in English and in each country’s official language are presented in Figure 1. We chose to extend the analysis to include Brazil and China in addition to the 6 countries included in the original analysis, as these are considered rising research economies. This is confirmed by the fact that the compound annual growth rate (CAGR) for articles indexed in Scopus between 1996 and 2011 from Brazil was 13% and China 19%, which is far greater than the 3 to 5% CAGR that is usually expected.

![Ratio English to local language journal articles](image)

**Figure 1**: Ratio of the number of journal articles published by researchers in English to those in the official language of eight different countries, 1996–2011. Source: Scopus.
Figure 1 shows that, in line with the original study, the use of English has continued to rise strongly in the Netherlands, Italy and the Russian Federation over the past four years. It has also increased somewhat in Germany, but remained relatively stable in France, Spain and China. However, in Brazil, the ratio between the use of English and Portuguese is clearly decreasing, although this might be due in part to an increase in the coverage of Brazilian journals published in Portuguese instead of English in Scopus. However, overall, the use of English clearly continues to increase over time.

Subject specific use of English?

The next question is whether there are subject fields in which researchers still publish regularly in their own language instead of in English. To answer this question, a general search was carried out in Scopus to determine the number of articles published in each of the selected languages between 1996 and 2011. The languages included in the search were the same as those presented in Figure 1, with the addition of English, so a comparison could be made between English and the other languages.

Table 1 provides an overview of the percentage of articles published in the four main categories per language, as a percentage of the total publication output in that language from 1996 to 2011.

The results indicate that researchers publishing in English, Chinese or Russian tend to publish most in fields related to the ‘harder’ Physical and Life Sciences, such as Physics, Engineering and Materials Science. On the other hand, researchers who choose to publish in Dutch, French, Italian, Portuguese or Spanish tend to publish their work most in fields related to the ‘softer’ sciences, such as the Health Sciences, Social Sciences, Psychology and Arts and Humanities. This ranges from almost 80 percent for the Netherlands and Italy to roughly 60 percent for Germany and Portugal. Although these ranges are similar across countries, there is a high level of variation in the actual fields within these main categories. For example more than half of all Dutch language publications are related to Health Sciences, which includes Medicine, Dentistry, Nursing and Veterinary Science, while in Italian nearly 41 percent of all publications are related to Social Sciences, Arts and Humanities.

Overall, these results appear to confirm that researchers publishing in languages other than English tend to do so somewhat more in the softer disciplines than in the harder ones (1, 2). Although English clearly continues to be the preferred language of scientific communication, there are still plenty of disciplines within which researchers continue to publish in their native language as well.

References:
Section 3: Expert Opinion

FORCE11 gains momentum creating the future of research communications and e-scholarship

Scientists, scholars, publishers, and librarians in many locations and disciplines are developing methods and tools to improve the process of creation, reviewing and/or editing of scholarly content, working on technologies and techniques to interpret, visualize, or connect (scientific) knowledge more effectively. They are formulating concepts, tools, standards, and techniques for sharing multimodal research data. These developments are currently taking place in disparate and disconnected domains, including Computational Linguistics, Bioinformatics, Information Science, the semantic web and data technologies in general, Social Sciences, and computer-human interface studies. It is apparent to publishers and scholars alike that the future holds radical and disruptive changes in both the nature and form of the transmission of scholarly reports. What has largely been missing, however, is a forum where these diverse groups can discuss how they are working to effect these changes, and develop a common platform to communicate, collaborate and co-develop new architectures, models, and modes of working; in short – to invent the future as a distributed collective.

FORCE11 is an international group representing scholars from various disciplines, librarians, archivists, publishers, and research funders, developing an understanding of the problems and potential surrounding scholarly communication by means of information technology. The activities of this group grew out of a series of workshops held in 2011, Beyond the PDF (held in January 2011 in San Diego) and the Dagstuhl Workshop for the Future of Research Communications, the outcomes of which were reflected in the research agenda presented in the FORCE11 Manifesto (1).

The FORCE11 Manifesto lays out the vision and motivation for FORCE11 along with a potential agenda for focusing activities of FORCE11. In this article, we will first summarize the key proposals from this paper and then discuss our current and future activities.

Key Proposals from the manifesto: Define new publishable objects:

To date, online versions of ‘scholarly outputs’ have tended to replicate print forms, rather than exploit the additional functionalities afforded by the digital terrain. We believe that digital publishing of enhanced papers will enable more effective scholarly communication.

Our vision entails creating a new, enriched form of scholarly publication that enables the creation and management of relationships between knowledge, claims and data. In this vision, the journal article or research paper is rapidly being replaced as the standard unit of currency by which knowledge is exchanged: it becomes but one among many forms. In the most generic sense, the new form of knowledge exchange centers on the research object - a container for a number of related digital objects - for example a paper with associated datasets, workflows, software packages, etc., that are all the products of a research investigation and that together encapsulate some new understanding. Publishing of research objects is not necessarily publishing as we know it today, achieved by the same mechanisms as used for traditional scholarly articles. It consists of providing free and open access to the component parts of the research object, which may or may not have been individually reviewed by others either pre- or post-publication.

There is a temporal aspect to research and the scholarly lifecycle that also needs to be recorded, either within research objects or between research objects, and that should also be capable of being reproduced. This means the creation of a knowledge infrastructure that allows the sharing of computationally executable components, such as workflows, computer code and statistical calculations, as scientifically valid content components; and an infrastructure that allows these components to be made accessible, reviewed, referenced and attributed.
Collate innovative publishing tools:
Developing the tools to support these new modes, if undertaken from scratch, would be an immense undertaking. Thus, where possible, existing tools should be adapted and integrated within a newly opened and increasingly integrated infrastructure. This change is likely to occur gradually through a series of incremental steps, most of which will not be driven by the technology. Rather, the technology should respond to the recognized requirements of scientists for improved dissemination, reproducibility, recognition, etc. Efforts at archiving, retrieving and citing digital research objects in standardized ways should be closely linked with open data and open-source software publication approaches, and should converge on common standards and practices. Citations to datasets and other digital research objects within publications should be treated on a par with the current treatment of bibliographic citations. Citations to these in the text should be made with a standard reference mark (in-text reference pointer) and the full reference should be given in the reference list of the publication, using a resolvable globally unique identifier (URL, DOI, HDL). Additionally, a formal semantic representation of the metadata into semantic data standards such as OWL and RDF, describing these research objects, their provenance, their relationships to and citations of one another etc., would be very useful and now achievable.

Treat data as a first-class object:
We have to develop best practices for depositing research datasets in repositories that enable linking to relevant documents, and that have high compliance levels driven by appropriate incentives, resources and policies. In addition, for scientific domains, the new forms of publication must facilitate reproducibility of results, which means, at least for in silico research, the ability to preserve and re-perform executable workflows or services. This will require the ability to re-construct the context in which these objects were executed, which may well contain or reference other executable objects as well as data objects that may evolve through time. In this way, the content of communications about research will follow the same evolutionary path that we have seen for general web content: a move from the static to the increasingly dynamic.

Collectively develop new business models:
Current business models for scholarly publication face significant disruption due to many factors: the growth in open access, the advent of alternative publication platforms that exploit new technologies for inexpensive communication and information exchange over the internet, a widening view of what constitutes a publishable research object (e.g. data, workflows), and the challenges of curating, linking and preserving the wider world of digital research objects. Furthermore, it is anticipated that the overall funds dedicated to scholarly communication may well become more restricted in future, at least on a per researcher basis. Both the major customers (research libraries) and brokers (currently, publishers) have an interest in being an active party in shaping the transition to new, sustainable business models, to ensure that the transition is a smooth one.

In a collaboration involving scholars, publishers, libraries, funding agencies, and academic institutions, we need to develop models that can enable this exciting future to develop, while offering sustainable forms of existence for the constituent parties, although perhaps not in their present states. To be financially viable, new communication modes will need to demonstrate tangible value to both producers and consumers. To be sustainable, the cost recovery streams will need to be aligned to perceived value. The changes we envisage pave the way for a revolution in the manner in which research is carried out and communicated, leading to significant improvements in scholarly productivity and quality, and enhanced transparency. In collaboratively reinventing science publishing, we hope to increase the public trust in and access to the value and outputs of science, and draw new participants into our endeavors; quite possibly the greatest challenge we face within Science as well as the Arts and Humanities.

Explore new metrics of impact:
To obtain the benefits that networked knowledge promises, we have to put in place reward systems that encourage scholars and researchers to participate and contribute. We need to acknowledge the fact that notions such as journal impact factor are mere surrogates for measuring the true impact of scholarship, and are increasingly irrelevant in a world of disaggregated knowledge units of vastly varying granularity (2). We need to derive new mechanisms that allow us more accurately to measure true contributions to the ongoing enterprise of augmenting the world’s store of knowledge. Measuring impact is complex because it depends on context, on purpose, and on audience. It can have different effects for different individuals. Similarly, a communication can have different degrees and even polarities of effect. For example, a research paper might be simplified and published by newspapers to make headline news with great societal impact, but be roundly criticized or even ignored by academic colleagues. To address these issues, better mechanisms of measurement need to be put in place, that allow for different types of impact and influence. A multi-dimensional measurement instrument would be useful. It needs to be customizable for specific situations and individuals and it must be easy to use both for the individual academic and for the reviewer or decision-maker.
Current Work and Next Steps

This agenda is ambitious, but progress - at least on some fronts - has been rapid. In the spring of 2012, a one-year grant was awarded by the Alfred P. Sloan Foundation to the FORCE11 group, led by Phil Bourne of UCSD, to achieve three goals:

- Development of a web platform to allow virtual communication and tool building for accelerating change in science publishing and increasing the community that participates in this process;
- Stimulate collaboration on the creation of a series of exemplars to further explore the key points proposed in the FORCE11 Manifesto;
- Organize another workshop to bring together an ever-growing community of tool builders, scientists, publishers, librarians, funders, and other interested parties to discuss these matters.

Six months into this process, we are actively driving this agenda (lead by Executive Director Maryann Martone) and are pushing this agenda on all three fronts:

- What started as a simple website has now evolved into a distributed, Drupal-based collection of modules, containing blogs, access to a wide catalog of tools, websites and other resources, a calendar and set of links, and several other virtual community components.
- The next workshop, ‘Beyond the PDF 2’, is scheduled for March 18-20, 2013, in Amsterdam, The Netherlands (see http://force11.org/beyondthepdf2 for more details). The goal is to make this a ‘future-centric workshop’ that allows active virtual participation through webinars, online demos, and real-time social media interactions.
- We are developing concepts for exemplars to showcase the wide variety that is currently ongoing inside and around FORCE11 themes.

From the time the Manifesto was produced to now, new tools have been developed and existing ones matured. Exciting new developments appear daily and we hope that many of these will be in evidence in the next Beyond the PDF conference. But we admit that replacing a system that is so intertwined with the evaluation and advancement of academia will not be easy or swift.

We wish to invite anyone interested to join us in this endeavor by signing up as members on http://force11.org, attending the workshop in Holland, either virtually or in person, or by starting a discussion and looking for partners to define new projects that address the issues we have identified or other related matters pertaining to the future of research communications and e-Scholarship. If we get this right, the potential is immense. We greatly look forward to hearing your comments and ideas, and welcome you to join us, by signing up as a member at http://force11.org, or contacting us directly.

Anita de Waard, Elsevier Labs, FORCE11 Executive Board Member – a.dewaard@elsevier.com

Maryann E. Martone, UCSD/NIF, FORCE11 Executive Director – maryann@ncmir.ucsd.edu

About the authors

Anita de Waard has a background in experimental physics, which she studied in Leiden and Moscow. She joined Elsevier as a physics publisher in 1988 and has been working as Disruptive Technology Director within the Labs division of Elsevier since 1997. Her work focuses on establishing active research collaborations with key academic institutes in Europe and the US and as such, she has co-organised the ‘Beyond the PDF’ and ‘Force11’ workshops and instigated Elsevier’s Grand Challenge and the Executable Papers challenge. Her research interests include implementing standards and cross-disciplinary frameworks for sharing annotations and content. Next to that, she conducts research in collaboration with the University of Utrecht, pertaining to a discourse analysis on key rhetorical components in scientific text.

Maryann Martone received her BA from Wellesley College in biological psychology and her Ph. D. in neuroscience in 1990 from the University of California, San Diego, where she is currently a Professor in the Department of Neuroscience. She is the principal investigator of the Neuroinformatics Framework project, a national project to establish a uniform resource description framework for neuroscience. Her recent work has focused on building ontologies for neuroscience for data integration. She just completed her tenure as the US scientific representative to the International Neuroinformatics Coordinating Facility (INCF), where she still heads to program on ontologies. Dr. Martone recently joined FORCE11, an organization dedicated to advancing scholarly communication and e-scholarship, as Executive Director. She serves as Co-Editor in Chief of Brain and Behavior, a new open access journal for brain-related research.

References:

Section 4: Behind the data

International scientific migration analysis generates new insights

Gali Halevi
MLS, PhD and
Dr. Henk F. Moed

Introduction
Scientific networks, collaboration and exchange have been the center of attention in numerous research articles and conferences’ discussions. For example, publications on the topic of “brain drain” have grown from 34 in 2000 to over 100 in 2011. The main reason for the increased interest in these topics has been the premise that these types of exchanges benefit scientific progress in that they foster innovation, and enhance and enable the flow of ideas between scientists in different institutions (1), (2), (3). In addition to the actual growth of science and scientific activity, there has been much effort to show that such progress benefits the economy through a line of investigation tying basic research to patents production.

Bibliometrics took a main methodological role in the studies of co-authorship and its results as indicators of collaborative trends by using affiliation information embedded in the bibliographic data of publications. In addition to the ability to track and sketch scientific collaborations between institutions, the availability of author profiles (now also through ORCID) and their affiliation information in Scopus™ has also made possible the tracking of scientific migration from country to country (4), (5), (6). Such information is of immense value to our ability to study research migration and use it as a way to inform science policy and track the formation of research excellence centers as they draw domestic and international talent to their doors.

This article describes researchers’ migration trends between 17 countries (see Table 2) and sketches some of its major trends. In addition, it looks at co-authorship patterns and describes the similarities and differences between these two phenomena in order to examine the unique patterns that both these lines of investigation offer and the ways in which each can be used as a way to shed some light on the formation of science excellence in different areas of the world.

The model
In order to study migration patterns, we have defined a specific model for the analysis, in which the move of researchers from one country to another can be more easily tracked. Since bibliometric methods are used, the connection between the theoretical construct and the bibliometric one is specified (see Table 1).

<table>
<thead>
<tr>
<th>Theoretical Concept / Interpretation</th>
<th>Bibliometric Constructs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher</td>
<td>Scopus Author ID</td>
</tr>
<tr>
<td>Active Researcher</td>
<td>Publishing year</td>
</tr>
<tr>
<td>Currently Active Researcher</td>
<td>Publishing in 2011-2012</td>
</tr>
<tr>
<td>Researcher starting a scientific career during years 2001-2002</td>
<td>First publication appears in 2001-2002</td>
</tr>
<tr>
<td>“Young” researcher in 2011</td>
<td>First publication year &gt;2000</td>
</tr>
<tr>
<td>Migrating Researcher (from country A to B)</td>
<td>Publishing author’s “work” country changes from A to B</td>
</tr>
</tbody>
</table>

Table 1: Conceptual premises and their bibliometric constructs
We collected the research output of 17 countries, among which 10 are considered growing countries (noted in red) and 7 are considered as established (noted in blue), from different regions in the world (see Table 2). For each country, the research output for 2000-2012 was collected. In order to trace the movements of researchers from one country to another, we used the unique Author ID offered by Scopus™ as a way to identify individual authors. In Scopus™, the affiliations associated with an author through their publications are kept and become a part of the unique author profile constructed within Scopus™. This allows for an analysis of migration, as one can identify in which institution and country an author has published. Moreover, the fact that the affiliation is tracked per author allows for a comparison between co-authorship and migration, and enables the distinction between the two as separate indicators of areas of collaboration vs. mobility.

Results: migration towards USA
Using the synchronous approach, analyzing the 2011 publications published by authors from a particular study country and including authors who started their careers from 2001 to 2010, we were able to trace the strengths of migration between various countries. For example, in Figure 1, there are three levels of migration trends to the USA. The strongest migration levels can be seen from countries listed in the inner circle, such as China, Canada, India, UK, Australia, and others, as denoted in the red lines closest to the center (within the green circle). Moving further away from the center and denoting less migration to the USA are countries such as Iran, Mexico, Singapore, Turkey, Ireland, Poland, and others. Still, one can see clearly that there is a significant amount of migration from these countries, as the middle circle denotes stronger migration than the outer circle, which includes countries such as Malaysia, Pakistan, Hong Kong, Romania and so forth.

Migration versus co-authorship
The connections between geographical areas of collaboration and migration can be clearly viewed by plotting the higher co-authorship / migration ratio countries on a network map. Figure 2 shows links between pairs of countries on the basis of the ratio of the percentage of authors migrating from one country to another and the percentage of co-authorships between the two countries. It only shows links for which this ratio exceeds 1.2. In other words, the map shows which pairs of countries demonstrate a migration relationship that is at least 20% stronger than expected on the basis of their level of co-authorship. In this map, notice the role that China plays as a hub for migration and collaboration between Singapore and Taiwan, connecting them to the USA and the UK. Again, in this map one can see the major countries scientifically engaged with the USA and also attracted to it in terms of migration, such as China, India, Brazil, Japan, UK, but also Iran, Turkey, Thailand, Romania, Bangladesh, and others. Pockets of migration-collaboration can also be seen between Malaysia - Nigeria and Iran, Romania - Belgium and Hungary, Italy - Switzerland and Argentina, Iran - Australia, Azerbaijan, Netherlands, UK, Canada, France, Japan and USA.

Discussion
Our study of 17 countries has shown that there is a difference between co-authorship and migration trends. From the data available it is apparent that common language and geographical proximity drive international migration more strongly than it drives co-authorships. In addition, it seems that political tensions do not present a barrier to collaboration and migration when it comes to scientific publications. This can be seen in the relatively high ratio of co-authorship and migration between Iran and the USA, India and Pakistan. There are some interesting patterns in the types of migrations emerging from this line of investigation. Some countries tend to show more temporary migration patterns as researchers move to a different country to complete an academic degree or residency but return to their origin country to continue their career and subsequent publications. This type of migration supports the development of the country’s professional skills levels and infrastructure and this type of exchange seems to be increasing. Furthermore, declining patterns of researchers leaving their country on a permanent basis can also be found at the opposite side of the spectrum.
Conclusions

Using a bibliometric approach to analyze affiliations within articles and the ability to systematically attribute them to unique authors’ profiles enables the study of migration and co-authorships trends. Research migration analysis clearly has different patterns than co-authorship’s and generates new insights into the global scientific network, as it can potentially create a breeding ground for future international collaboration. Both ‘brain drain’ and ‘building up scientific infrastructure’ are visible in the data, but cannot always be separated.

Caution must be applied when analyzing authors’ profiles in the way described in this article, as they do sometimes contain errors that could distort the results. However, it should be noted that such error in this study is minor since relative indicators based on large numbers are insensitive to errors in author profiles.

References:
Section 5: Country Trends

The Rise of Latin American science

Sarah Huggett
MPhil

Latin America is a vibrant, multicultural region, with an estimate of nearly 600 million inhabitants and combined GDP of $6.27 trillion USD PPP. Several of its respective governments have been trying to increase their nation’s international visibility, some of them through growing emphasis on Science, Medicine, & Technology. But how does Latin America fare bibliometrically? Research Trends investigates...

Latin America as a region

From 2000 to 2010, Latin America has seen high growth of more than 9% per year in scholarly output, resulting in a nearly 70% increase in its share of world papers over the same period, to reach just under 4.4% of the world’s annual output of scholarly papers in 2010. Latin American research is growing fast and becoming more visible on a global scale. And this is not the only bibliometrically observed improvement to Latin America’s scholarly output over the last few years: Latin America’s relative citation impact, albeit still under world average, has been improving by 1.6% per year from 2000-2010, from about half of world average in 2000 to more than four fifths in 2010 (see Figure 1).

In which subject fields do Latin America’s strengths lie?

In certain subject areas, Latin America’s share of world scholarly articles is even more sizeable. For Dentistry, it is nearly 10%, and for Agricultural & Biological Sciences, nearly 11%, while for Veterinary it reaches an even higher 12%. These prolific areas have been increasing in relative citation impact from 2000-2010, from 1% per year for Dentistry (very close to the world average, which was 0.97 in 2010) to 3% per year for Agricultural & Biological Sciences (see Figure 2).

Which Latin American countries are leading the rise?

Latin America is composed of numerous countries of various sizes, each with their own particular attitude and programs towards R&D; as a consequence, the scientific output of each country varies wildly. For instance, while researchers based in Honduras published 58 papers in 2010, researchers based in Brazil published nearly 38,000, reaching 2.3% in 2010. The only other two Latin American countries with more than 0.5% of total world scholarly papers in 2010 are Mexico with just under 0.7% and Argentina with just above 0.5%. These countries have also been improving their citation impact: Brazil by 1.3% per year from 2000 to 2010 to reach 0.75 in 2010, Mexico by 1.8% to reach 0.81 in 2010, and Argentina by a higher 2.5% per year, nearly reaching the world average at 0.9 in 2010 (see Figure 3).
Figure 1: Latin America's annual share of total scholarly papers and 5-year relative citation impact, for which 2010 refers to 2006-2010 scholarly papers and their 2006-2010 citations with world average at 1. Source: Scopus.

Figure 2: Annual share of total scholarly papers and 5-year relative citation impact, for which 2010 refers to 2006-2010 scholarly papers and their 2006-2010 citations with world average at 1, for subject areas with paper share equal to or higher than 9.9% in 2010. Source: Scopus.

Figure 3: Annual share of total scholarly papers and 5-year relative citation impact, for which 2010 refers to 2006-2010 scholarly papers and their 2006-2010 citations with world average at 1, for countries with paper share higher than 0.5% in 2010. Source: Scopus.
Brazil: emphasis on Health Sciences

The sheer size of Brazil in terms of population, and recent investments in Brazilian R&D at both national and international levels, can explain the prolificacy of Brazilian science. Indeed, in 2010, Brazil accounts for more than half of Latin America’s output in scholarly papers (52.7%). In some areas related to Health Sciences, Brazil’s scholarly paper share actually soars to around 70% of Latin America’s output, although their citation impact remains inferior to that of Latin America as a whole. Researchers based in Brazil publish, for instance, nearly 71% of Latin America’s papers in Dentistry, but their research only reaches 80% of the global Latin American research in that area (see Figure 4).

Mexico: focus on Physical Sciences

Mexico’s investment in R&D may appear modest compared to Brazil’s, but in 2010 this country contributed 16% of the overall Latin American research output. Physical Sciences appear to be a priority for Mexico, as illustrated by the Large Millimeter Telescope. In some areas of Physical Sciences, including Physics and Astronomy, it is reaching more than 22% of Latin America’s scholarly papers; however in none of these areas does its citation impact equal that of Latin America as a whole (see Figure 5).

Argentina: well distributed outputs

Argentina’s R&D expenditures may appear limited, yet in 2010 it managed to publish nearly 12% of Latin America’s scholarly paper output. Interestingly, its highest article shares of Latin America’s output are in varied areas spanning from the Physical and Life Sciences to the Arts & Humanities. In Earth and Planetary Science, it publishes over 17% of Latin America’s papers in 2010, although these only reach three quarters of the citation impact in that area for Latin America. In Arts and Humanities, Chemical Engineering, Earth and Planetary Science, and Immunology and Microbiology, it not only reaches over 15% of Latin America’s research, but in these areas Argentina’s citation impact is above that of Latin America as a whole (see Figure 6).

Scholarly paper share:
Share of the world’s output of scholarly papers published in a given year.

5-year relative impact:
Relative measure of citation impact (number of citations divided by number of papers) for which 2010 refers to 2006-2010 scholarly papers and their 2006-2010 citations.

The Future of Latin American Science?

As a region Latin America has become a visible actor in global research through increases in both article share and relative citation impact. The diversity of Latin American countries is reflected through the various R&D strategies adopted by the different countries composing it. As demonstrated by Brazil, size does matter, and international investments can make a real difference to a country’s scholarly output. Mexico shows how a focus in a particular area can increase output. Both countries are, however, beaten to the post in terms of relative impact by Argentina, which manages to reach above Latin America’s average in specific and varied subject areas. All three countries illustrate that different priorities and strategies can lead to different, yet successful outcomes – perhaps this is the way forward for Latin America and the patchwork of countries composing it.
Figure 4: Brazil's annual share of Latin America scholarly papers and 5-year relative citation impact, for which 2010 refers to 2006-2010 scholarly papers and their 2006-2010 citations with Latin America's average at 1, for subject areas with paper share equal to or higher than 65% in 2010. Source: Scopus.

Figure 5: Mexico's annual share of Latin America scholarly papers and 5-year relative citation impact, for which 2010 refers to 2006-2010 scholarly papers and their 2006-2010 citations with Latin America's average at 1, for subject areas with paper share equal to or higher than 22% in 2010. Source: Scopus.

Figure 6: Argentina's annual share of Latin America scholarly papers and 5-year relative citation impact, for which 2010 refers to 2006-2010 scholarly papers and their 2006-2010 citations with Latin America's average at 1, for subject areas with paper share equal to or higher than 15.5% in 2010. Source: Scopus.
On October 9th Research Trends organized a seminar that focused on the current international and local trends in research mobility and its implications on fostering innovation from a Government, Academia and Industry perspective. The seminar attracted deans, policy analysts, researchers, directors, and graduate students, who gathered at George Washington University to hear and discuss the subject with prominent speakers in the field.

Dr. Henk Moed from Elsevier, who conducted extensive research on the subjects of both co-authorship and mobility trends, presented the similarities and differences between the two as can be learned from bibliographic data. Some of his findings included the discovery that political tensions inhibit neither co-authorship nor researcher mobility, as could be found in increased migration and collaborations between them. In addition, his research found close ties between research mobility, geographical proximity and language similarity. This finding opened a discussion regarding the attraction of researchers to centers of excellence and whether that was the motivating factor.

Mr. Ben Wildavsky, the author of the Great Brain Race, answered some of the issues raised during the discussion by emphasizing the fact that research mobility is driven by worldwide phenomena such as the rise in local academic ranking systems and the race to create world class universities that will attract and retain talent. He added that researchers will migrate only when they have good reason to do so, such as going to an exceptional research institute that seeks their expertise or a learning opportunity that can advance their careers. In his presentation, Mr. Wildavsky gave some specific examples of such trends, including Asian university ranking systems and researcher career paths.

Dr. Mark Regets, Project Officer, Science & Engineering Indicators Program at The National Science Foundation, approached the topic from the government R&D point of view, looking at the global capacity for science and technology and showing its rapid growth in most parts of the world. In his presentation Dr. Regets presented data that demonstrated the increased and more complex flows of students, workers, and finances, the increased regional S&T collaboration and links between regions driven by a global labor market for certain research skills.

The seminar’s presentations and discussion concluded with the understanding that the move of scholars from one country to another has far-reaching implications for economics, scientific innovation and progress on both local and international levels. Such mobility influences the creation of entrepreneurial networks that lead to economic growth and scientific discoveries. Student exchange, immigration, pursuing career opportunities and other reasons motivating talent to move from country to country are all contributing to the mobile nature of research and the formation of collaborative networks. This phenomenon has been studied in different ways such as bibliometric research looking at authors’ affiliations and co-authorships, R&D and immigration patterns analysis as well as funding and productivity. This area has gained attention, and research has seen titles that include topics such as research mobility, brain circulation and brain drain among the few found in this context.

In order to better direct research mobility and enable migration as a part of the research policy agenda, there is a need to combine different disciplinary methodologies, both qualitative and quantitative. Combining bibliometrics, economics and social sciences can lead to a better understanding of the subject and enable decision makers to create better policy to support talent attraction and retention that will benefit the country as a whole.
Section 7
Reporting Back

Research Evaluation in Practice, National Geographic Society, October 17th 2012, Washington DC
http://www.researchtrends.com/research-trends-seminar/

Gali Halevi
MLS, PhD

Over a hundred people gathered at the beautiful Grosvenor Auditorium at the National Geographic Society to participate in a day-long seminar on practical applications of research evaluation methodologies. The seminar included diverse perspectives on research evaluation and its implications on funding allocations in industry, government and academic settings. The day opened with a keynote speech by Debra Perez, Assistant Vice President for Research and Evaluation at the Robert Wood Johnson Foundation. As a representative of one of the largest corporate grant giving foundations in America, Dr. Perez focused on the foundation’s research and evaluation work concentrating on public health and health care services. Dr. Perez shared statistics and facts concerning issues of race and economic status relating to health and health care in the Americas, and spoke of the foundation’s funded programs in these areas working towards equality and quality of services for underprivileged and minorities.

A methodological approach to research evaluation was presented by Dr. Henk Moed of Elsevier, who discussed the multifaceted nature of research evaluation. In his presentation, Dr. Moed presented the Multi-Dimensional Research Assessment Matrix, whereby the motivation, purpose and methodologies used for evaluation are taken into consideration. The main premise of the matrix is that one has to consider the why, what, when and how, and choose the correct method for each scenario, before applying any methodology to evaluate the impact of research or a researcher.

An international perspective on the subject was presented by Dr. Marc Luwel from the Hercules Foundation in Belgium. In his talk, Dr. Luwel described the Flemish approach to performance-based (research) funding and its evolution over time. The main motivation for the development of evidence-based research funding was the need to promote excellence and be able to better manage universities working with diminishing resources. To address these challenges, Dr. Luwel described the Flemish development of multi-level indicators, including input-output, JIF, CWTS-crown index, H-index, Review panels and departmental ratings which were aggregated at university level to allocate block funding. These indicators were used in a funding formula that was re-visited and evaluated over the years in order to be able to address changing issues. Following the introductory presentation, Dr. Luwel presented a full case study of the Flemish approach to research evaluation, which was described in detail.

A science of science policy perspective was provided by Dr. Rebecca Rosen from the American Institute of Research (AIR) and a former National Science Foundation (NSF) staff member. Dr. Rosen gave an expansive overview of the work done by AIR and NSF with regards to collecting, analyzing and disseminating science-related data to assist research evaluation and science policy decision makers to reach conclusions in a timely and effective manner. Dr. Rosen gave specific examples of how NSF and AIR approach the data infrastructure challenges and the tools and methodologies built to address them. In addition, Dr. Rosen covered AIR and NSF efforts in the US, France, and Australia to integrate existing administrative, programmatic, and results databases into data platforms that feed novel portfolio visualization tools.
A unique and powerful social science perspective on evaluation was given by Dr. Abraham Wandersman from the University of South Carolina, Columbia. Dr. Wandersman, professor of Psychology, presented the Quality Implementation Tool, developed in order to address the need for an evaluative methodology that stems from empowerment and is goal- and quality-oriented. Dr. Wandersman gave examples of the model's use in practice as well as an empowerment evaluation example using the Tool. The framework presented by Dr. Wandersman was co-developed by the Center for Disease Control CDC staff and university researchers to bridge the research-practice gap by integrating research-to-practice models with community-centered/practice-centered models.

The day was concluded by Dr. John Francis, the vice president for research, conservation, and exploration at the National Geographic Society (NGS). Dr. Francis described the various research funding grants given by NGS to scientific research and exploration. In addition to the grants NGS provides for basic and field research, Dr. Francis also discussed the various programs run by NGS in schools, colleges and its citizens-participation programs, which aim to connect people with nature and raise awareness for environmental issues around the world. Dr. Francis focused on the various types of grants offered by NGS and the manner by which each one is evaluated in order to support exploration and discovery, natural and cultural conservation, and groundbreaking scientific fieldwork, all aimed at learning about and protecting our planet.

This seminar offered a comprehensive view of research evaluation and grant making and brought together diverse perspectives from industry, academia and government. Moreover, each presentation during the seminar covered not only different approaches to evaluation but also different practices showcasing topics such as health care, data infrastructure, bibliometrics, algorithms, psychology, and nature, and how each uses evaluative methodologies in practice.
Did you know that more than 6.1 million students in the United States took at least one online class during the fall 2010 term? That’s a 10.1% increase over the year before. Furthermore, online enrolment accounted for roughly a third of the total enrolment in that term.

For more on this topic, see:


Notes: