IDENTIFIERS AND OPEN SCIENCE: ELEMENTS FOR A SOCIO-TECHNICAL ANALYSIS OF ORCID

GAËLLE BÉQUET

PhD in Information Science
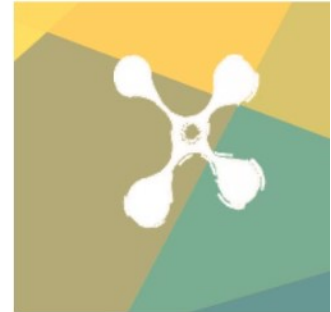Director of the ISSN International Centre, France
gaelle.bequet@issn.org
https://orcid.org/0000-0001-8299-9724

IDENTIFICADORES E CIÊNCIA ABERTA: ELEMENTOS PARA UMA ANÁLISE SÓCIO-TÉCNICA DA ORCID

Resumo

Objetivo: descrever e analisar o ambiente técnico e social de um identificador usado globalmente na comunidade de pesquisa preocupada com a implementação da ciência aberta. **Design │ Metodologia │ abordagem:** este artigo descreve o processo de criação de um identificador através das lentes da Social Construction Of Technology (SCOT) e da Actor Network Theory (ANT). Esta pesquisa usa fontes primárias disponíveis nos sites Orcid e ISNI, bem como literatura cinza fornecida por um bibliotecário envolvido no desenvolvimento do ISNI. Fontes secundárias, e. g. artigos revisados por pares também são citados. **Resultados:** os grupos sociais que apoiam o Orcid são identificados e suas ações são descritas. O modelo de desenvolvimento da Orcid é analisado. A relação entre Orcid e o advento de sistemas de monitoramento da produção científica pública é estabelecida. **Originalidade │ Valor:** identificadores persistentes são considerados uma infraestrutura chave da ciência aberta. Várias partes interessadas estão envolvidas em sua implementação. É importante entender qual agência essas partes interessadas têm nos sistemas técnicos para fazê-los atender aos seus objetivos.

Palavra-chave: Identificadores, Ciência aberta, Construção social da tecnologia (SCOT), Teoria da rede de atores (ANT), Padrões de informação e documentação

IDENTIFIERS AND OPEN SCIENCE: ELEMENTS FOR A SOCIO-TECHNICAL ANALYSIS OF ORCID

Abstract

Goal: to describe and analyze the technical and social environment of an identifier globally used in the research community concerned with the implementation of open science. **Design│Methodology│Approach:** this article describes the creation process
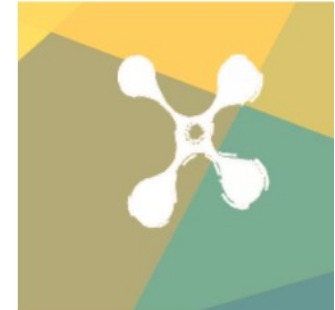
of an identifier through the lens of the Social Construction Of Technology (SCOT) and the Actor Network Theory (ANT). This research uses primary sources available on the Orcid and ISNI websites as well as grey literature provided by a librarian involved in the development of ISNI. Secondary sources, e. g. peer-reviewed articles, are also quoted. **Results:** the social groups supporting Orcid are identified and their actions are described. Orcid's development model is analyzed. The relation between Orcid and the advent of systems monitoring public scientific production is established. **Originality│Value:** persistent identifiers are considered as a key infrastructure of open science. Various stakeholders are involved in their implementation. It is important to understand which agency these stakeholders have on technical systems to make them serve their objectives.

Keywords: Identifiers, Open science, Social Construction of Technology (SCOT), Actor Network Theory (ANT), Standards in information and documentation
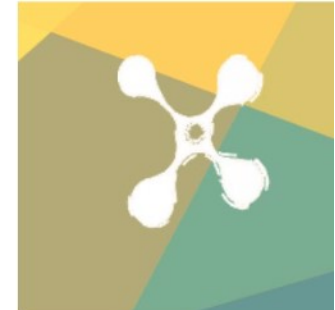
1. Introduction

An international identifier cannot be reduced to a number with a few metadata. It is a social construction, as shown by the cyclical standardization process of the International Organization for Standardization and its various validation stages, and the establishment of registration authorities responsible for the dissemination of an identifier through the signing of agreements. An international identifier is also a technical construction because the registration authority develops and maintains databases and services that support the effective implementation of the identifier in its official scope.

How to define an international identifier? Developed by Subcommittee 9 of Technical Committee 46 Information and Documentation of the International Standard Organization, the technical specification ISO/TS 22943:2022 on identification principles proposes definitions of several concepts related to the identification of documents or other objects at an international level. An identifier is defined as a "sequence of characters that uniquely denotes a referent. Identifiers may be used to specify the referent. In some cases, they may be substituted for the referent or may be used to retrieve the referent or its metadata." This specification establishes the overarching characteristics, e. g. uniqueness, persistence, granularity, metadata quality or accessibility, that an identifier must have to be considered reliable by its user community, e. g. archives, libraries and museums, publishers, authors and researchers and online content providers.

The first identifiers with an international scope were created in the mid-1970s, with the International Standard Book Number (1970) and the International Standard Serial Number (1975) being the precursors in this field. ISBN and ISSN operate in a very similar way: a central authority coordinates a network of national agencies that assign identifiers

and transmit data to a centralized database. However, the content of these databases differs: for the ISBN, the centralized database lists the publishers of the member countries to which an ISBN prefix has been assigned; for the ISSN, it lists the identifiers of the serial publications themselves described by metadata. Other international identifiers were subsequently standardized by ISO on the same model of specialization by type of analog documents. Examples include the International Standard Music Number (1993) for printed music and the International Standard Audiovisual Number (2002) for audiovisual content.

With the advent of the Internet, the progressive digitization of analog documents and the exponential production of natively digital documents, new identifiers have appeared such as the Digital Object Identifier, standardized by the ISO in 2012 and based on the Handle technology developed since 1995 by the Corporation for National Research Initiatives (CNRI) in the USA. DOIs are thus assigned to different types of digital resources by the producers themselves who have DOI prefixes. These producers are responsible for creating and updating metadata as well as maintaining and declaring to a central agency, e. g. Crossref or Datacite, the valid link between the DOI and the server address (URL) where the digital resource is hosted. The information resources are thus distributed while the access mechanism is centralized. Following the evolution of the identification of different analog documents, DOIs are now assigned to different types of digital documents, e. g., journal articles, research data, and even audiovisual content, just as ISBN and ISSN are now issued to digital books and journals as specified in their respective ISO standards.

These identifiers are commonly used in scientific research for the identification and description of research results, e. g. for print and digital journals (ISSN) and for articles (DOI). Scientific publishers, both commercial and non-profit, commonly use these
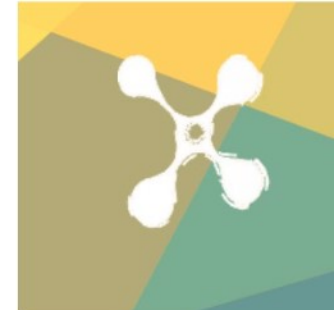
identifiers in their internal production tools and in their transactions with content providers, libraries, and researchers, as shown by the cooperation established in 2017 between the ISSN network and Springer-Nature to identify monographic series and journals with an ISSN at an early stage, i. e. even before they are published. Beyond the identification of documents, it is the entire research ecosystem that has become the scope of identifiers with the introduction of International Standard Name Identifier (2012) and Research Activity Identifier (2022).

Contemporary global scientific production has acquired great visibility thanks to the Internet, where millions of documents produced by millions of researchers employed by as many organizations are accessible free of charge or for a fee. Between 2000 and 2019, the number of researchers has increased from an average of 6 per 1000 employees to 9 per 1000 in OECD countries. In India, the number of researchers grew from 110 to 255 per 1 million between 2000 and 2017. Because of this staggering growth and the difficulty of publishing in increasingly selective and expensive scientific journals, some researchers have questioned the established model whereby research is funded by government authorities while a majority of its results are made available by commercial publishers. This critique was articulated in the Budapest Open Access Initiative Declaration (2001) which called for the creation of Open Access journals and institutional open science repositories. The Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003) reinforced this international momentum in support of open science.

In 2021, Unesco defined open science as "an inclusive construct that combines various movements and practices aiming to make multilingual scientific knowledge openly available, accessible and reusable for everyone, to increase scientific collaborations and sharing of information for the benefits of science and society, and to
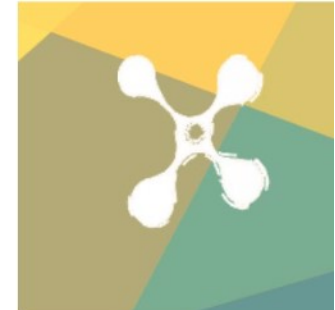
open the processes of scientific knowledge creation, evaluation and communication to societal actors beyond the traditional scientific community." According to the Recommendation on Open Science adopted at the 41st Unesco General Conference, the success of open science requires the establishment of documentary infrastructures including journals, articles, academic repositories, archives and scientific data associated with bibliometric and scientometric tools, all of which are based on unique perennial identifiers but also on the mobilization of a range of social groups, "scientists and societal actors [. . .] policy makers and practitioners, entrepreneurs and community members, giving all stakeholders a voice in developing research that is compatible with their concerns, needs and aspirations." Unesco thus wishes to promote participatory science and involve citizens in its production.

Based on the hypothesis that identifiers and open science are socio-technical constructs, we will present in a first part the analytical tools of social constructivism, insisting on their relevance for the study of the implementation of identifiers in the process of opening science. In the second part, we will analyze through the public documentation of the Orcid consortium, archives and secondary scientific sources, the current deployment process of Orcid for the identification of researchers and its impact in the construction of a specific model of open science. The third part will be devoted to a critical evaluation of this model.

2. Methodological procedures

This research uses primary sources available on the Orcid and ISNI websites as well as grey literature provided by a librarian involved in the development of ISNI. Secondary sources, e. g. peer-reviewed articles, are also quoted.

The methodological approach adopted in this article is that of social constructivism used in the field of Science and Technology Studies (STS). In The Social Construction of Reality (1966) Peter Berger and Thomas Luckmann postulate that stable social institutions emerge from multiple and diverse interactions between individuals. These institutions are then objectified by other individuals in whose eyes they appear as unalterable realities.

The STS have declined this concept to study scientific facts and technology. Laboratory Life: The Social Construction of Scientific Facts (Bruno Latour and Steve Woolgar, 1979) is an ethnographic description of daily life in a biology laboratory. In their book, the authors present an ethnographic account, based on note-taking, of the discovery process by describing the tribe of researchers, their interactions, their beliefs and the rituals practiced in the laboratory.

In Aramis ou l'Amour des techniques (1992), Bruno Latour describes a posteriori the failure of a French automated train project. For this investigation, Latour uses the interview reports, the documents retracing the different phases of the project commented and analyzed by the actors of the project. He is thus in line with Erving Goffmann's symbolic interactionism, which postulates that all social actors act in relation to others within social frameworks. However, he deploys from symbolic interactionism a new vision of social frameworks which are not stable realities but evolving frameworks constructed by the subjectivities of the actors.

Susan Leigh Star and James R. Griesemer in Institutional Ecology, "Translations" and Boundary Objects (1989) promote an ecological analysis of the creation of an institution, i.e. a natural history museum in California, focusing on its environment and on the social groups, be they researchers, amateurs or collectors, who participated in this project. For
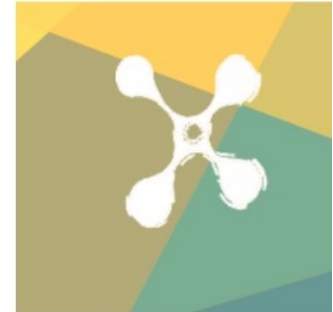
their research, Star and Griesemer use archival documents, correspondence and interviews with some of the actors of the project.

The author of this article uses the same methodology as the authors of social constructivism cited above to analyze the concomitant creation and development of two information systems specialized in the identification of physical persons at the international level. The author describes the interactions between the social groups involved in the competitive process of deployment of these two identifiers. To do so, she uses primary sources available on the ORCID and ISNI websites, video recordings of conferences as well as grey literature provided by a librarian involved in the development of ISNI. Secondary sources, e.g. peer-reviewed articles, are also quoted. The author of the article experimented with this methodology in her doctoral dissertation, Trois bibliothèques européennes face à Google: Aux origines de la bibliothèque numérique (2014).

3. Results

3. 1. For a mobilization of the concepts of social constructivism

In 1984, Trevor Pinch and Wiebe Bijker published a seminal paper that postulates that technical tools are embedded in a social environment. (Pinch & Bijker, 1984). "Relevant social groups" are involved at all stages of artifact development. To study a technical artifact, it is not enough to describe its components, it is also necessary to identify the communities involved and to observe their attitudes towards the artifact. The artifact is represented in different ways by these communities, which gives it what the two authors call "interpretative flexibility". During the development of an artifact, the social

groups agree on a "technological framework" as the outcome of negotiations that lead to the stabilization of the technical artifact. The latter is described as a "black box".

In the same line of thought, Bruno Latour (2007) postulates that a technical artifact catalyzes a social network made up of different communities, e. g. entrepreneurs, commercial or industrial organizations, administrations, clients, professional users, who elaborate and defend their own representations of the artifact. These social groups clash during controversies around a technical object, as for example during the constitution of a digital library (Béquet, 2009). "Each element of a technical arrangement depends on a choice of world, proposes a certain definition of the user, involves a policy of patents and standards, engages an industrial strategy." (Latour, 2003).

Madeleine Akrich (1993) posits that an artifact is not a "black box" but rather a "boundary object" that separates and coordinates several social groups. This technical artefact also has the particularity of carrying within itself the conception of the user imagined by the creator of the object. This can be seen in the documents that describe the project and those that are exchanged within the socio-technical network.

Patrice Flichy (2003) defines the concept of socio-technical framework which associates the framework of functioning (knowledge, know-how, techniques) and the framework of use (project team, investors, testers, users). Three moments can be distinguished in the elaboration of an artefact. The first moment is that of the "suitcase object" on which all the social groups will project their representations. It is characterized by the production of a utopian documentation: the social groups imagine an artifact which will meet all their needs apart from any technical consideration. Among all these ideal scenarios, the groups make choices and enter into negotiation or conflict around a "boundary object" that is still customizable. Then comes the moment of the stabilization of the socio-technical framework and the technological locking of the innovation. The

cooperation between the different groups is then formalized through financial circuits, contracts and agreements. The artifact can continue to evolve through the deployment of new services, but without calling into question the agreement initially reached between the groups.

Following them, other researchers have applied the theory of social constructivism to information systems that have specific characteristics compared to other technical artifacts.

> There has been a growing consensus that information technology is both shaping of, and shaped, by its working environment. [...] Firstly, an information system is a particularly flexible type of technical artifact [...] Secondly, the concept of 'closure' is unlikely to materialize in the context of an information system, as modifications and revisions are likely to be effected at regular intervals throughout its working life. (Doherty et al., 2006).

These authors assert that the functionalities of an information system are framed by "enforcing constraints", i. e. obligatory passages for users such as access to a service via the use of identifiers or the compulsory entry of data, and by "proscribing constraints", i. e. functionalities that do not exist or are unusable in the information system. They believe that these constraints limit the interpretative flexibility of the artifact by the social groups involved.

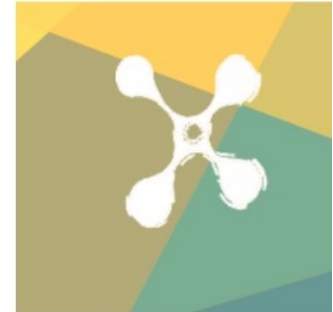In the context of scientific research, identifiers are thus manifested by sequences of characters associated with a referent. It is important to bear in mind that these identifiers are underpinned by complex identification systems which are "structures that support the assignment, management and governance of identifiers within a community. The identification system may comprise a combination of the following:

- the identifier standard specifying an identifier, its associated metadata, and assignment mechanism,

- the management systems for collecting, storing and sharing the identifiers and associated metadata,

- a mechanism for resolving an identifier to retrieve its metadata and, in some cases, a referent,

- its governance systems,

- the social network that supports the use of the identifier." (ISO 2022).

Each identifier is therefore associated with one or more standards, metadata, an attribution mechanism and a database. It functions thanks to governance rules and thanks to the communities that produce and use this identifier. The identifier can therefore be described as a fully-fledged information system.

The article "The social construction of self-sovereign identity: An extended model of interpretive flexibility" is an example of this socio-technical approach to an emerging identifier. Its authors are interested in an innovation concerning a digital identification system based on blockchain technology. This allows an individual to directly manage his interactions with online services while preserving his personal data which are stored on a solution chosen by the individual (computer, telephone, personal cloud). The term Self-Sovereign Identity (SSI) is used to name this innovation which combines several inventions, namely Decentralized Identifiers (DIDs), Verifiable Credentials (VCs) and a Public Key Infrastructure (PKI). This innovation could eventually supplant the digital identities issued to an individual by third-party organizations, e. g. Facebook, Google, governments. Indeed, these third-party organizations use databases that are either not very compatible with the protection of personal data or are centralized and vulnerable to cyber attacks. Based on interviews with representatives of companies, government, non-

profit organizations and academia and through the analysis of the documentation of SSI solution providers, the authors separate the technical aspects, the social dimensions and the institutional constraints of SSI. In addition to the functional boundaries of SSI and its social interpretive flexibility, they show that there is an organizational dimension in the interpretation of a technology (IT legacy issues, modernization of citizen identification management, Anglo-Saxon distrust of any state centralization of information) as well as environmental factors (European General Data Protection Regulation on personal data).

Drawing on the theoretical approaches listed above and this article, the following is an analysis of the deployment of the Orcid identifier.

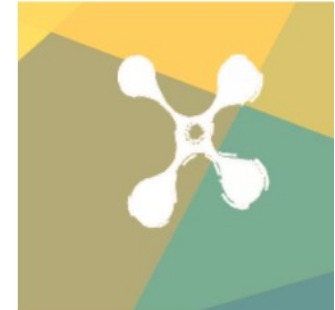## 3.2. Deployment of the Orcid identifier for researchers (2009-2022)

3.2. 1. What is Orcid?

Orcid stands for Open Researcher and Contributor ID. The USA-based Orcid Inc. was founded in 2010 with an aim to "build a central registry of unique identifiers for individual researchers to foster scholarly communication and research." The free Orcid registry, which went live in 2012, collects all Orcid identifiers. Orcid Inc. became a nonprofit organization in April 2013[1], thus exempt from income taxes under U. S. law.

What is the problem that Orcid proposed to solve in 2011?

> Orcid Inc. aims to solve the author/contributor name ambiguity problem in scholarly communications by creating a central registry of unique identifiers for individual researchers and an open and transparent linking mechanism between Orcid and other current author ID schemes. Orcid will focus on currently active

---

[1] Retrieved October 30, 2022, from https://projects.propublica.org/nonprofits/organizations/275142743
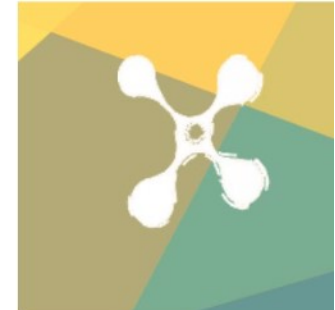
researchers. Orcid will be a hybrid system of self-and organization-asserted identity. Data will come from individuals and organizations. (Bilder, 2011, p.2).

The development of higher education, as illustrated by the Universities Worldwide[2] website, which listed 6,864 universities in 178 countries in June 2004 and 9,826 in 209 countries in August 2022, and the associated growth in the number of researchers, has favored the emergence of research projects conducted by large, transnational and multidisciplinary teams funded by a wide variety of organizations. Associating researchers with the scientific institutions that employ them and the funding agencies that support their projects, differentiating them in publications by reducing the probability of homonymity linked to the growth of the sector and the internationalization of scientific research are becoming strategic objectives for global research governance. Orcid is therefore positioned as a new identifier with an international scope focused on active researchers, fitting into an existing set of researcher identifiers (Herther, 2010) such as arXiv Author Id created by the ArXiv academic repository, Digital Author Identifier by the Netherlands libraries, universities and Surf, Author Id by Scopus/Elsevier and Researcher Id by Thomson Reuters. These proprietary identifiers are either limited to databases or limited to a territory. From the outset, Orcid has had a universal vocation and wished to establish interoperability mechanisms with existing identifiers. Among these, the International Standard Party Identifier (ISPI), which began to be standardized by the ISO in 2006, is positioned in the same niche as Orcid.

Orcid aims to establish a centralized registry of identifiers and metadata qualifying individuals involved in scientific research, e. g. names, biography, employment,

---

[2] Universities Worldwide, Retrieved October 30, 2022, from https://univ.cc/

education and qualifications. Orcid operates through a set of governance rules. It relies on a community of users and members who use this identifier. We can therefore describe Orcid as a fully-fledged information system and review its emergence and development using concepts from socio-technical analysis.

3.2.2. The social groups behind Orcid

In December 2009, the journal Nature[3] announced the creation of an author ID system backed by 23 organizations, including Thomson Reuters, Nature Publishing Group, Elsevier, ProQuest, Springer, CrossRef, the British Library and the Wellcome Trust. Thanks to Orcid, we will be able to distinguish the "multitudinous Dr Smiths and Professor Wangs" and associate them with their scientific productions. At its inception, Orcid was created to address mainly an audience of Anglo-Saxon and Asian researchers. The creation of a "digital curriculum vitae" was a complementary functionality to the simple identification, which confers on Orcid the status of a "suitcase object". Thomson Reuters played a leading role in this initiative by giving the Orcid consortium, which acquired a legal personality in 2010, a copy of the code of its identifier system, namely Researcher ID, used by the Web of Science database. CrossRef, which manages DOIs for scientific publications, has also been another critical partner. Indeed, this organization worked on a very similar ContributorID project starting in 2007. It then joined the Orcid project in 2009, dropped ContributorID and created a bridge between Orcid and DOI that became effective in 2015.

As of 2011, 265 institutions, i. e. publishers and research institutions, had already expressed their support for Orcid Inc. (Orcid, 2011). The revenue to sustain the organization was estimated to be over $2 million. This amount was not reached until

---

[3] Credit where credit is due, *Nature*, *462*(7275), 17 December 2009

2015[4]. At the outset, Orcid Inc. was therefore a structurally loss-making organization that relied on external contributions, such as the provision of CrossRef staff to assist in the development of the Orcid system (Bilder, 2011). A significant investment was devoted to the review of the computer code delivered by Thomson Reuters. Funding was also being provided by the National Institute of Health. In February 2011, the Mellon Foundation awarded a grant to MIT to create the Orcid business model. This business model was mixed, combining free access for researchers creating and managing their Orcid record and institutional membership for scientific organizations, consortia and publishers wishing to use the database and interface it via APIs with their own information systems[5]. Orcid was presented as an initiative of the research community supported by publishers.

3.2. 3. ISNI, a precursor identifier to Orcid

Orcid has several similarities to ISNI. Beginning in October 2006, three years before the Nature article on Orcid, work began on standardizing a new identifier under the leadership of the International Confederation of Societies of Authors and Composers (CISAC) (ISO, 2006). The ISO working group included experts commissioned by the Bibliothèque Nationale de France, the Federation of Italian Publishers (AEI), Bowker/ProQuest, OCLC, Nielsen Book Data, the British Library, the Conference of European National Libraries (CENL), the International DOI Foundation and the ISSN International Centre. This working group also included participants representing ISAN International Agency, the Civil Society for the Administration of Performers' Rights, ISBN International Agency, the Swiss Society for the Rights of Authors of Musical Works

---

[4] Retrieved October 30, 2022, from https://projects.propublica.org/nonprofits/organizations/275142743

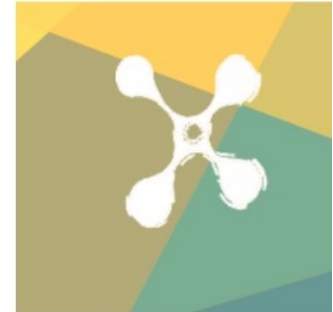[5] Retrieved October 30, 2022, from https://info.orcid.org/membership/

(SUISA), and the International Federation of Reproduction Rights Organizations (IFFRO).

CISAC was leading this standardization project within ISO Technical Committee 46 to develop what was called at that time an International Standard Party Identifier (ISPI) based on the CISAC proprietary identifier, i. e. the Interested Party Information (IPI). This IPI system is the global unique identification of a right holder acting across multiple creation classes (musical work, literary work, work of art), assuming different roles (musical creator, film director, author of fine art), and owning rights (performing right, reproduction right, radio broadcast right). Through ISO standardization, CISAC's IPI system claimed to become a universal and interoperable identifier. The IPI sought to identify "parties", i. e. individuals (natural person), organizations (legal person), groups of individuals (e. g. music band) and organizations (e. g. associations). Eventually, the ISO working group agreed that this identifier within the creative industries ecosystem was intended for individuals who might work under different public identities and use pseudonyms. These various identities required as many identifiers and therefore did not define a "party" but a "name". The ISO 27729 (ISO, 2012) standard published in March 2012 created the International Standard Name Identifier and the term "party" disappeared from the title.

The ISNI Registration Authority was officially endorsed by ISO in 2012 and OCLC (Leiden) became the operator of the identifier assignment system. The ISNI information system consisted of a database based on the PICA software, acquired by OCLC in 2007, in which were uploaded records of natural and legal person authorities pooled by national libraries in Virtual International Authority File[6] (VIAF) to which ISNIs are assigned. The

---

[6] Retrieved October 30, 2022, from https://viaf.org

ISNI database quickly reached a critical size thanks to the contribution of VIAF data. However, it was based on an aging system that did not allow for immediate APIs for interfacing with partner information systems. The team based in Leiden was limited to a few people supported by a small IT team.

3.2. 4. The relationship between Orcid and ISNI: cooperation, complementarity or competition?

The relationship between ISNI and Orcid could be described as a low-key controversy. Technically, the Orcid system was originally based on the one developed by Thomson Reuters for ResearcherID. Orcid claimed to create links with other identifiers in the academic domain. As proof of this desire to cooperate, the Orcid identifier has adopted the same syntax as ISNI from its very beginning: "The Orcid iD is an https URI with a 16-digit number that is compatible with the ISO Standard (ISO 27729), also known as the International Standard Name Identifier (ISNI), e. g. https://orcid.org/0000-0001-2345-6789."[7] Orcid thus started issuing identifiers similar to ISNIs without being officially accredited as a registration agency by ISNI International Agency. The discourse of the two organizations was ambivalent, oscillating between emphasizing complementarity and asserting independence.

Consultation meetings have been held since March 2012 between the two agencies responsible for implementing the two identifier systems in order to work towards their interoperability. This consultation led to the signing of a Memorandum of Understanding[8] between Orcid and ISNI in January 2014. In this document, Orcid positioned itself as an

---

[7] Retrieved October 30, 2022, from https://support.orcid.org/hc/en-us/articles/360006897674-Structure-of-the-Orcid-Identifier

[8] Retrieved October 30, 2022, from https://info.orcid.org/wp-content/uploads/2020/06/Orcid-ISNI-MOU_20140204_Signed.pdf

open and interdisciplinary identifier created by the research community. It differs from ISNI, which is presented as an identifier with a more institutional character because it was created by ISO and supported by national libraries and international rights management organizations. Orcid would thus be a bottom-up, community-based initiative while ISNI would be a top-down, institution-led one. There is, however, a kinship between the two systems, a guarantee of potential interoperability, since ISNI IA has allocated a block of two million identifiers to Orcid for its own use: "ISNI has reserved a block of identifiers for use by Orcid, so there will be no overlaps in assignments. This range of identifiers is defined between 0000-0001-5000-0007 and 0000-0003-5000-0001."[9] The memorandum of understanding also stated that the two systems would work on tools for linking their identifiers and sharing metadata. Technical cooperation and joint communication efforts were also announced. In 2013, Orcid planned to launch an affiliation module using ISNI and Ringgold Organization Identifiers[10]
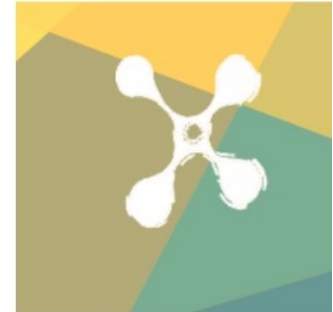
How was this relationship between ISNI and Orcid perceived by the main stakeholders, namely researchers? In 2015, a report by five European research organizations (Knowledge Exchange, 2015) compared ISNI and Orcid and assessed the use of these identifiers in five European countries, i. e. Denmark, Finland, Germany, Netherlands, United Kingdom. In fact, the two identifiers coexisted in national identification systems that set up reconciliation mechanisms themselves, e. g. Integrated Authority File[11] in Germany and Narcis[12] in the Netherlands.

---

[9] Retrieved October 30, 2022, from https://support.orcid.org/hc/en-us/articles/360006973973-What-is-the-relationship-between-ISNI-and-Orcid-

[10] Retrieved October 30, 2022, from https://info.orcid.org/fr/orcid-plans-to-launch-affiliation-module-using-isni-and-ringgold-organization-identifiers/

[11] Retrieved October 30, 2022, from GND - Homepage

[12] Retrieved October 30, 2022, from https://www.narcis.nl/

Has this memorandum of understanding concretely improved cooperation between the two agencies? The two identifiers have continued to develop in parallel, with each information system retaining its autonomy, as the official Orcid communication shows: "Because Orcid and ISNI have different purposes and serve different communities both organizations are necessary. The organizations will each hold different data, have different privacy and ownership rules for data, have different business models and offer different services. Most importantly of all, ISNI and Orcid will be identifying different things for different communities."[13] In 2022, a query in both systems shows that active researchers may have two separate identifiers. However, a reconciliation mechanism has recently been implemented allowing researchers to link their identifiers, but no specific communication has been made on this subject by the two organizations and no guidelines are currently available.

3.2. 5. The expansion of the Orcid system

Orcid is intended for a limited, albeit global, audience, i. e. researchers, their employers and funders. In order to enhance its open and community-based approach, Orcid favors the direct creation of records by researchers who retain control of the data associated with their Orcid. However, Orcid does not disdain the contribution of research organizations (Bilder, 2011) which can provide their researchers' data. This raises legal issues, which we will discuss in a third section. In comparison, ISNI initially relied on data provided by libraries and associations managing intellectual property rights in a controlled identification process. The model evolved in 2021 with the creation of a web

---

[13] Retrieved October 30, 2022, from https://support.orcid.org/hc/en-us/articles/360006973973-What-is-the-relationship-between-ISNI-and-Orcid
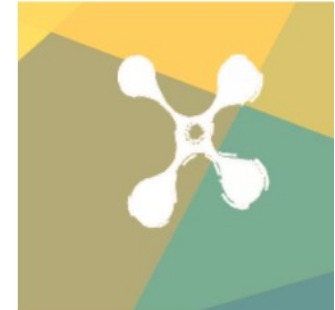
interface by the British Library allowing individuals and organizations identified by ISNI to modify their metadata, including adding their Orcid[14].

To better understand the expansion and the model of the Orcid information system, we must now look at the social groups that support it. Let us first consider the major scientific publishers involved in the system from the beginning. Elsevier and Springer-Nature are multinational companies with a large portfolio of publications. Their interest is to promote the use of identifiers integrated to their internal production chains, from manuscript submission to publication and peer review. They can compel researchers to create an Orcid by making it mandatory for any manuscript submission. For example, in 2016, 112 publishers and publications issued an open letter[15] to require authors to use Orcids. "The move gives Orcid a big boost in a long competition to set an ID standard." (Bohannon, 2016, P.213) Emerald, Sage, Taylor & Francis and Wiley are members of Orcid as of 2022. About 100 publishers support Orcid, the majority of which (54) are located in the UK and the USA.

Research funding agencies are another community involved early on in Orcid because they want to improve the management of the funds granted and make the results of the projects they fund visible through the publications of the recipient researchers. In 2016, the Austrian Science Fund FWF stated: "In its workflows, FWF has to ensure a clearly identifiable link between grants and grant-holders. Orcid provides such unique identifiers. FWF is obligated to document its investments in basic research. Orcid enables researchers to create records of their publications and a variety of other

---

[14] Retrieved October 30, 2022, from https://isni.bl.uk/
[15] Retrieved October 30, 2022, from https://info.orcid.org/requiring-orcid-in-publications/

forms of scientific output. The identifiability of their work is a fundamental prerequisite for any assessment of the impacts of science policy."[16]

Learned societies and university consortia are also represented in the Orcid membership[17]. In addition, Orcid has received significant support from the European Commission in two projects. Orcid EU received 739,851 euros in the THOR project (2015-2017)[18]. Orcid US participated in the FREYA project (2017-2020)[19] to "extend a robust environment for Persistent Identifiers (PIDs) into a core component of European and global research e-infrastructures. [. . .] This will provide an essential building block of the European Open Science Cloud (EOSC)." The FREYA project aimed to achieve interoperability between several identifiers, including Orcid, CrossRef and DataCite DOIs, and emerging identifiers such as ROR and RAID, through a PID Graph enabling innovative services to be built for the academic sector as we will see later.

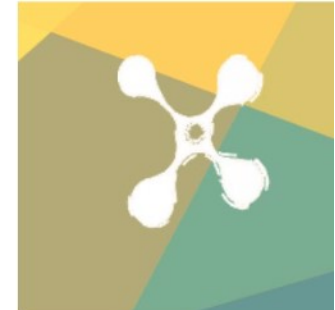### 3.2. 6. Orcid in the construction of open science in Europe

Thanks to the European Commission funding allocated to Orcid and its partners, which has enabled initiatives such as the PID Forum or the annual PIDapalooza conference, the importance of persistent identifiers (PIDs) in building open science has been highlighted. An independent expert report commissioned by the EOSC Executive Board in October 2020 developed a "Persistent Identifier (PID) policy written for senior decision makers within potential EOSC service and infrastructure providers, and of interest to all EOSC stakeholders. It defines a set of expectations about what persistent

---

[16] Retrieved October 30, 2022, from https://www.fwf.ac.at/en/news-and-media-relations/news/detail/nid/20151116-2155/

[17] Retrieved October 30, 2022, from https://orcid.org/members

[18] Retrieved October 30, 2022, from https://cordis.europa.eu/project/id/654039

[19] Retrieved October 30, 2022, from https://cordis.europa.eu/project/id/777523

identifiers will be used to support a functioning environment of FAIR research."[20] This was followed in December 2020 by a report entitled PID Architecture for the EOSC[21]. While Orcid is cited in these two studies, some limitations of the identifier are also highlighted: "Orcid uses the http://orcid.org global resolver, but neither Orcid PID can be resolved by other global resolvers, like the DOI one, nor can the Orcid resolver resolve any other PIDs. A more generic and universal solution would be beneficial here"[22]; "Resolution in the sense of virtualisation and redirection to the data is not necessary, since the data is maintained in the same monolithic system. Resolution in this case is simply providing the content behind the identifier in the ORCiD database."[23]

Why is Orcid so involved in the process of building open science? As we have seen above, the constitution of its database of identifiers was initially based on data sharing with databases managed by publishers. The latter then put pressure on authors to create their Orcid. In a second phase, Orcid convinced funding agencies to force project sponsors to provide an Orcid in their applications for funding. Finally, this constraint was exercised via national research organizations through plans for open science, for example in France (College Europe and International, 2019) and Germany. [24] By being assigned to a majority of researchers at a global level, Orcid becomes a pivotal identifier that makes it possible to link these researchers to their open access production (thanks in particular to interoperability with the DOIs of CrossRef and DataCite), whether it is published by commercial publishers (gold open access) or made available on academic

---

[20] Retrieved October 30, 2022, from https://op.europa.eu/fr/publication-detail/-/publication/35c5ca10-1417-11eb-b57e-01aa75ed71a1

[21] Retrieved October 30, 2022, from https://op.europa.eu/en/publication-detail/-/publication/3136c3e6-4f07-11eb-b59f-01aa75ed71a1/language-en/format-PDF/source-184010810

[22] Ibid. p. 5

[23] Ibid. p.25.

[24] Retrieved October 30, 2022, from https://www.orcid-de.org/

repositories (green open access). Based on this linking of data held by the providers of identifiers, which are non-profit and community-based organizations, many free or paying services can be developed, such as those offered by Digital Science[25] around the analysis of scientific production and its impact, or the implementation of indicators for the evaluation of research, as in the United Kingdom. [26]

3. 3. A critique of the Orcid system

The Orcid identifier is recommended, and sometimes even imposed, by a range of actors, e. g. multinational publishers, research funding agencies, public research organizations and academic repositories.[27] Its development is also favored by its association with national identifiers allowing to link metadata or to share them[28]. This "success story" of an identifier created about ten years ago has generated a lot of support, as seen above, and many criticisms detailed below.

3. 3. 1. Identification under constraint

In order to quickly reach a critical mass of identifiers, Orcid has enlisted in its network organizations that have the capacity to compel researchers to request an Orcid, whereas its initial model emphasized the freedom of researchers to create an open identifier. This researcher freedom remains a central element in Orcid's communication.[29]
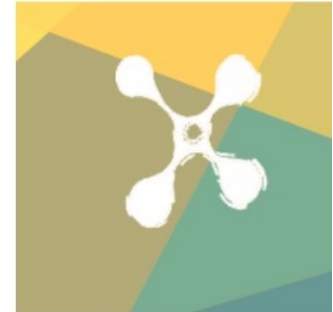
---

[25] Retrieved October 30, 2022, from https://www.dimensions.ai/blog/digital-science-announces-dimensions-integration-of-orcid/

[26] Retrieved October 30, 2022, from https://www.ukri.org/wp-content/uploads/2021/12/RE-151221-TheMetricTideFullReport2015.pdf

[27] Retrieved October 30, 2022, from https://ngr.coar-repositories.org/behaviour/identification-of-users/

[28] Retrieved October 30, 2022, from https://abes.fr/reseaux-idref-orcid/le-reseau/

[29] Retrieved October 30, 2022, from https://info.orcid.org/researcher-faq/#easy-faq-4991

However, the constraint exerted by the publisher, which imposes identification via Orcid of a researcher in its manuscript deposit system, can be analyzed as an impediment to the freedom of research (Choras & Jaroszewska-Choras, 2020), notably because it contravenes Article 27 of the Universal Declaration of Human Rights stating that "everyone has the right freely [. . .] to share in scientific advancement and its benefits."[30] One could also imagine that other information systems use Orcid as an access identifier (Knowledge Exchange, 2015) for the use of their services. In France, this forced identification could be perceived as an obstacle to freedom of research because it limits freedom of expression (Mouton, 2008). Publishers imposing the use of Orcid de facto discriminate between researchers by blocking the submission of articles without this identifier.

3. 3.2. Sensitive personal data

The Orcid identifier is associated with personal metadata that is subject to specific regulation in the European Union, i. e. the General Data Protection Regulation (GDPR)[31]. Although Orcid is a US registered non-profit organization, it is subject to the GDPR which has extra-territorial scope so that services outside the EU that process data of people inside the EU have to comply with it[32]. However, any dispute over metadata held in the Orcid system would be adjudicated in the US, which de facto discriminates against researchers from other countries.

The GDPR requires, among other things, that an individual give consent to the processing of his or her personal data. In the case of bulk Orcid creations by a research

---

[30] Retrieved October 30, 2022, from https://www.ohchr.org/en/human-rights/universal-declaration/translations/english

[31] Retrieved October 30, 2022, from https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-Organizations/principles-gdpr

[32] Retrieved October 30, 2022, from https://gdpr.eu/compliance-checklist-us-companies/

organization, the individual's consent is not automatically acquired. For this reason, Orcid has implemented a process called *create on demand* to "enable the member organization to explicitly request permissions from a researcher (opt in) to read and write to that individual's Orcid record."[33]

The GDPR also includes the right to be forgotten. A researcher should be able to deactivate their Orcid account, which is provided for in the system. However, this Orcid continues to be used in the databases of publishers or research organizations that may have retrieved the Orcid and associated metadata.[34] This metadata can also be used for profiling a researcher or to promote or limit access to funding opportunities, either because of areas of expertise or because of negative evaluations of certain academic work.
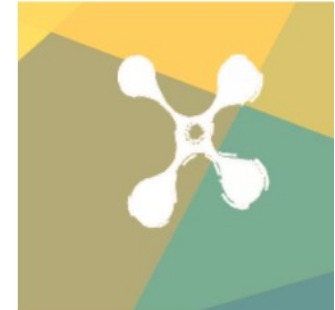
3. 3. 3. Is Orcid really the unique identifier for researchers?

In 2016, the publisher of a scientific journal questioned Orcid's ability to create and maintain a unique identifier for a given researcher. Indeed, the practice of "ghost authorship" is widespread in some regions, with PhD students writing papers for their supervisors, who publish them under their names and with their Orcid. Several authors may also share the same Orcid (Leopold, 2016). And as we have seen above, a researcher may have several identifiers between which reconciliation mechanisms are still non-existent or poorly documented.

This coexistence between multiple academic author identifiers is evident in surveys of researchers. In 2017, a study (Tran & Lyon, 2017) of 335 researchers at Stony Brook

---

[33] Retrieved October 30, 2022, from https://info.orcid.org/the-importance-of-opt-in

[34] Retrieved October 30, 2022, from https://support.orcid.org/hc/en-us/articles/360006973813-Deactivating-an-Orcid-account
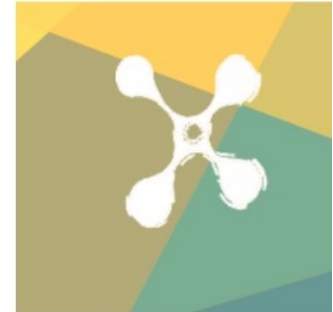
University, USA, showed that 49 were using Orcid, 29 Author ID (Scopus), 25 Researcher ID (Thomson Reuters) and 20 arXiv ID. 198 researchers did not use any ID. A 2020 study (Boudry & Durand-Barthez, 2020) of 1,047 researchers at the University of Caen, France, showed that 179 used Orcid and 152 used Researcher ID. These two studies agree on the fact that a minority of researchers do have an identifier as long as no binding policy has been put in place. They also agree that researchers are more likely to create profiles on academic social networking platforms such as ResearchGate and Academia.edu, or even on more general platforms such as LinkedIn or Facebook, than on Orcid or ResearcherID.

3. 3. 4. Orcid as an instrument of the platformization of science?

Some researchers (Mirowski, 2018) have recently described open science as a deliberate policy of subverting traditional science, which is accused of having failed for several reasons. The promoters of open science believe that citizens are increasingly questioning scientific results that are no longer considered as established and rational facts. Science would be an activity carried out by a quasi-aristocratic fraction of the population without sufficiently involving the average citizen. The productivity of researchers would be decreasing while the rate of retraction of articles would be increasing, proving that scientific journals, based on peer review, are not reliable. These journals should therefore be replaced by online content and service platforms that combine preprint publication with post-publication peer review and altmetrics.

The way in which open science is materialized in information systems and infrastructures can be described as platformization.

The term 'platform' as defined today in a digital context, now includes giving people and companies 'a platform' in the figurative and political sense, as well as the infrastructure through which they can sell products and services, share data and content, express themselves, and connect with other people. [...] Platformization is used to describe a company transitioning from a business selling products to one managing direct transactions between two or more actors in a platform-mediated network. (Andrews, 2020, p.265).

In the scientific research ecosystem, the best-known content and service platforms are Academia.edu, ResearchGate and Mendeley[35], all three funded by venture capital (Boudry & Durand-Barthez, 2020). The value production of these platforms is based on building communities of active researchers who share their results, initiate collaborations, analyze research data and establish metrics. Commercial publishers have come to consider that value production no longer resides solely in the publication of research results, but in the ability to control this production from end to end and to create data analysis tools that allow research policies to be adjusted. Elsevier markets Pure, "a Research Information Management System (RIMS) or Current Research Information System (CRIS) designed to be simple and turnkey. Deep integration into the Research Intelligence portfolio and external Open Access (OA) databases and Open Data repositories enables actionable analytics across sources for enhanced decision making and evidence-based execution of research strategy."[36]

How can Orcid be considered as a central element of the platformization of research? Orcid is building a monopolistic database at the international level dedicated to the identification of researchers. Thanks to the data collected from researchers and their

---

[35] Mendeley was bought by Elsevier in 2013.
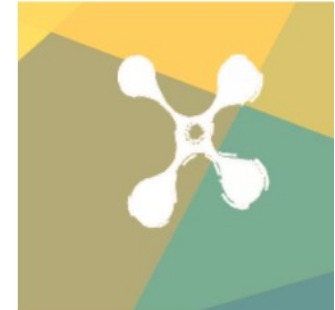[36] Retrieved October 30, 2022, from https://www.elsevier.com/solutions/pure

employers, as we have seen above, it is possible to construct new metrics for analyzing the behavior of individuals, such as their geographic mobility, even if "Orcid was not designed for that purpose but its database of 3 million scientists may become a vital tool for charting their flow around the world" (Bohannon, 2017, p. 690). Orcid can also aggregate information on the policies of research organizations, notably through the recruitment they carry out. In 2016, Elsevier's Pure system became interoperable with Orcid for researcher identification and with Ringgold for organization identification[37]. Orcid also links a researcher to their scientific outputs through a partnership with CrossRef and DOI retrieval. This information is fed into the Pure system allowing a subscribing organization to plan and implement its research policy in a single platform that controls the flow of data between different sources, including open databases like Orcid. In the open science ecosystem, an identifier like Orcid is essential to aggregate dispersed information and thus contribute to the platformization of open science.

4. Conclusions

Identifiers are essential tools for the development and monitoring of open science that will multiply the venues on which researchers disseminate their results. The current information system model for an identifier is based on a dumb sequence of characters associated with metadata, all managed in a central database with data exchange and linking mechanisms. Beyond the identification of documents, the academic domain is a place of emergence of identifiers associated with a given type of referent. The identification of researchers is thus the scope of Orcid, which was assigned to more than

---

[37] Retrieved October 30, 2022, from https://www.elsevier.com/connect/orcid-pure-a-happy-marriage-that-makes-researchers-lives-easier
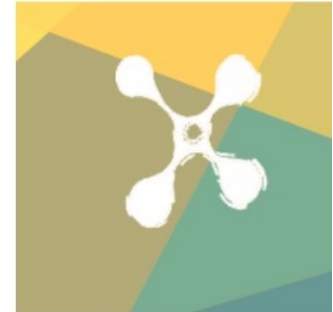
14 million researchers worldwide (2022). This study focuses on the social groups involved in the creation and development of an identifier for open science and their interactions. A more precise description of the technical tools underlying ISNI and Orcid could be a useful complement to it.

Author identification in academia is expected to expand dramatically in the future notably because the Office of Science and Technology Policy, USA, has recently published the memorandum *Ensuring Free, Immediate, and Equitable Access to Federally Funded Research*[38] requesting federally funded research outputs to come with digital persistent identifiers for all author and co-author names. At the same time, ISNI International Agency and Orcid Inc. have reached an agreement for a new block of 100 million identifiers to be assigned as Orcids. As a result, researchers can have two identifiers managed independently even if a linking tool exists. Clarifying the link between ISNI and Orcid would be beneficial to the research ecosystem.

The magnitude of the Orcid enterprise at the international level implies that some form of coercion is exerted on researchers and scientific organizations by publishers, research funding agencies, and government authorities. Orcid cannot be considered as an initiative of the research community, which results in relatively slow adoption in the absence of incentives or even coercion. Orcid is a pivot identifier for CRIS platforms which are intended to become the go-to place for any scientific endeavor, from fundraising to open access publication. Orcid is thus directly contributing to the current revolution in scientific research towards a science that is certainly more open but paradoxically more controlled.

---

[38] Retrieved from https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf

References

Akrich, M. (1993). *Les objets techniques et leurs utilisateurs, de la conception à l'action.* Retrieved September 7, 2022, from https://hal.archives-ouvertes.fr/halshs-00081731v1

Andrews, P. C. S. (2020). The platformization of open. In M. P. Eve & J. Gray (Eds.), *Reassembling scholarly communications: Histories, infrastructures, and global politics of open access.* MIT Press.

Béquet, G. (2009, May 25-30). *Digital library as a controversy: Gallica vs Google.* 9th Conference Libraries in the Digital Age [Paper presentation], Dubrovnik, Croatia. pp. 116-123. Retrieved September 7, 2022, from https://hal.archives-ouvertes.fr/hal-00718385/file/Digital_library_as_a_controversy.pdf

Béquet, G. (2015). *Trois bibliothèques européennes face à Google: Aux origines de la bibliothèque numérique (1990-2010).* Publications de l'Écoles des Chartes.

Berger, P., & Luckmann, T. (1966). *The social construction of reality. Doubleday.*

Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003). Retrieved September 7, 2022, from https://openaccess.mpg.de/Berlin-Declaration

Bilder, G. (2011). *Disambiguation without de-duplication: Modeling authority and trust in the Orcid system.* Retrieved September 7, 2022, from https://www.crossref.org/wp/labs/whitepapers/disambiguation-deduplication-wp-v4.pdf

Bohannon, J. (2016). Publishers embrace scheme to end name confusion. *Science, 351*(6270), p.213.

Bohannon, J. (2017). Restless minds: analysis of a vast set of public CVs reveals the world's most migratory scientists. *Science, 356*(6339), pp. 690-692.

Boudry, C., & Durand-Barthez, M. (2020). Use of author identifier services (Orcid, ResearcherID) and academic social networks (Academia.edu, ResearchGate) by the researchers of the University of Caen, Normandy (France): A case study. PLoS ONE Public Library of Science, 15(9), pp. e0238583. Retrieved September 7, 2022, from https://hal. archives-ouvertes.fr/hal-02936042

Budapest Open Access Initiative Declaration (2001). Retrieved September 7, 2022, from https://www.budapestopenaccessinitiative.org/

Choras, M., & Jaroszewska-Choras, D. (2020). The scrutinizing look on the impending proliferation of mandatory Orcid use from the perspective of data protection, privacy and freedom of science. *Interdisciplinary Science Reviews, 45*(4), pp. 492-507.

College Europe and International (2019). *Open credentials for open science*. Retrieved September 7, 2022, from https://hal-lara.archives-ouvertes.fr/hal-03640303/

Doherty, N. F., Coombs, C. R., & Loan-Clarke, J. (2006). *A re-conceptualization of the interpretive flexibility of information technologies: Redressing the balance between the social and the technical*. Loughborough's Institutional Repository. https://doi.org/10.1057/palgrave.ejis.3000653

Flichy, P. (2003). *L'innovation technique: Récents développements en sciences sociales vers une nouvelle théorie de l'innovation*. La Découverte.

Herther, N. (2010). Who's on first? Name disambiguation today. *Searcher: The Magazine for Database Professionals, 18(8).*

ISO International Organization for Standardization (2022). *ISO/TS 22943:2022 Information and documentation: Principles of identification*. Retrieved September 7, 2022, from https://www.iso.org/standard/83121.html

ISO International Organization for Standardization (2022). *ISO/TS 27729:2012 Information and documentation: International Standard Name Identifier (ISNI)*. Retrieved September 7, 2022, from https://www.iso.org/standard/44292.html

Knowledge Exchange (2015). *Researcher identifiers: national approaches to Orcid and ISNI implementation*. Retrieved September 7, 2022, from https://www.knowledge-exchange.info/event/researcher-identifiers

Latour, B. (1992). *Aramis ou l'amour des techniques*. Prix Roberval.

Latour, B. (2003). L'impossible métier de l'innovation technique (pp. 9-26). In P. Mustar & H. Penan *(Eds.), Encyclopédie de l'innovation*. Economica.
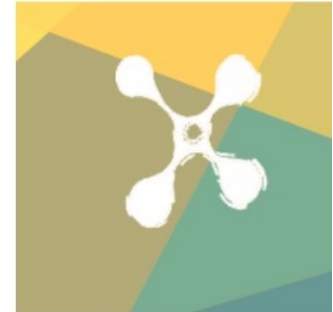
Latour, B. (2007). *Changer de société, refaire de la sociologie*. La Découverte.

Latour, B., & Woolgar, S. (1979). *Laboratory life: The social construction of scientific facts.* Sage

Leopold S. S. (2016). Editorial: Orcid is a wonderful (but not required) tool for authors. *Clin Orthop Relat Res.*, *474*(5), p. 1083-1085. https://doi.org/10.1007/s11999-016-4760-0

Mirowski, P. (2018). The future(s) of open science. S*ocial Studies of Science*, *48*(2), pp. 171-203.

Mouton, S. (2008). Les fondements constitutionnels de la liberté de recherche. In J. Larrieu, *Qu'en est-il du droit de la recherche?* Presses de l'Université Toulouse 1 Capitole. Retrieved September 7, 2022, from https://books.openedition.org/putc/2498?lang=fr

Orcid (2011). *Outreach meeting*. Retrieved September 7, 2022, from http://cdsweb.cern.ch/record/1384382?ln=sv

Orcid (2014). Orcid and ISNI: strategic partner MoU. Retrieved September 7, 2022, from https://info.orcid.org/wp-content/uploads/2020/06/Orcid-ISNI-MOU_20140204_Signed.pdf

Pinch, T., & Bijker, W. E. (1984). The social construction of facts and artifacts, or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science*, *14*(3). https://doi.org/10.1177/030631284014003004

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, translations and boundary objects. *Social Studies of Science, 19(3),* 387-420.

Tran, C. Y., & Lyon, J. A. (2017). Faculty use of author identifiers and researcher networking tools. *College & Research Libraries*, 78(2), 171. https://doi.org/10.5860/crl.78.2.171

Unesco (2021). Recommendation on open science. Retrieved September 7, 2022, from https://unesdoc.unesco.org/ark:/48223/pf0000379949

Acknowledgment