

ORCID growth and field-wise dynamics of adoption: A case study of the Toulouse scientific area

Marie-Dominique Heusse , and Guillaume Cabanac 



Marie-Dominique Heusse



Guillaume Cabanac

IRIT UMR 5505 CNRS, Université Paul Sabatier,
Toulouse, France

ORCID:

M. Heusse: [0000-0002-1537-465X](https://orcid.org/0000-0002-1537-465X)

G. Cabanac: [0000-0003-3060-6241](https://orcid.org/0000-0003-3060-6241)

*Corresponding author: Guillaume Cabanac, IRIT UMR
5505 CNRS, Université Paul Sabatier, Toulouse, France.
E-mail: guillaume.cabanac@univ-tlse3.fr

Abstract

Research-focused information systems harvest and promote the scientific output of researchers. Disambiguating author identities is key when disentangling homonyms to avoid merging several persons' records. ORCID offers an identifier to link one's identity, affiliations and bibliography. While funding agencies and scholarly publishers promote ORCID, little is known about its adoption rate. We introduce a method to quantify ORCID adoption according to researchers' discipline and occupation in a higher-education organization. We semi-automatically matched the 6,607 staff members affiliated to the 145 labs of the Toulouse scientific area with the 7.3 million profiles at orcid.org. The observed ORCID adoption of 41.8% comes with discipline-wise disparities. Unexpectedly, only 48.3% of all profiles listed at least one work and profiles with no works might just have been created to get an identifier. Those 'empty' profiles are of little interest for the entity disambiguation task. To our knowledge, this is the first study of ORCID adoption at the scale of a multidisciplinary scientific metropole. This method is replicable and future studies can target other cases to contrast the dynamics of ORCID adoption worldwide.

Keywords: bibliography, disambiguation, multidisciplinary, ORCID, researcher identifier

INTRODUCTION

Scholarly publications feature the names of the contributing authors in the by-line of the articles. With an increasing number of scholarly works published each year, the number of homonyms is growing too:

This article is an extended version of a conference paper given in French (Heusse & Cabanac, 2020, Adoption de l'identifiant chercheur ORCID : le cas des universités toulousaines. In: *INFORSID'20: 38e congrès de l'Informatique des Organisations et Systèmes d'Information et de Décision*). It covers a longer time period, includes the data for one overlooked lab, includes an analysis of adoption by sex and features a revised discussion.

Of the more than 6 million authors in a major journal citations and abstracts database, more than two-thirds of them share a last name and single initial with another author, and an ambiguous name in the same database refers on average to eight people. (Sabine, 2014).

Such ambiguity has several implications, a critical one being identity theft. The recent Surgisphere scandal a.k.a. LancetGate stressed this issue of homonymy detection (Piller, 2020). A US-based tenured faculty padded his CV with two-thirds of publications not written by him but by homonyms.

Bibliometric studies also suffer from the ‘namesake’ problem and Harzing (2015) stressed the case of author Y. Wang publishing nine papers a day as per Web of Science data. In reality, this identity amalgamates thousands of academics whose production is considered as one.

Research organizations worldwide use current research information systems (CRISs) to collect, analyse, and share their research output (Fabre et al., 2021; Sivertsen, 2019). They rely on document identifiers (e.g., DOI) and author identifiers, such as VIAF, ISNI and later ORCID in October 2012.¹

ORCID stands for ‘Open Researcher and Contributor ID’ and is operated by a non-profit organization that mints identifiers for authors of scholarly works (Haak et al., 2012). As of May 2021,² there were 11 million ORCIDs created, a number greater than 7.8 million researchers, a UNESCO estimate of the worldwide population of researchers (Soete et al., 2015, p. 33). This discrepancy may partly come from non-researchers creating ORCID profiles, such as support staff and librarians. Organizations and publishers also adopted and promoted ORCID: funding agencies and editorial managers request an ORCID to submit a proposal, a manuscript or a report during peer-review (Haak et al., 2018; Hanson et al., 2016). ORCID has also become a key component for open archives (Brown et al., 2016).

ORCID PROFILE: CREATION, UPDATE AND VISIBILITY SETTINGS

There were two ways to create an ORCID profile. On the one hand, up until 2016 an institution could create an ID for its employees, such as the University of Colorado.³ As of March 2021, ORCID listed 11 million profiles and 1,230 member organizations, 80% of which were of type ‘research institute’.⁴ There were six French institutional members (four research institutes and two publishers) at the end of 2019. Meanwhile many French institutions joined a consortium agreement with ORCID in December 2019.⁵ As of March 2021, ORCID listed 44 French members (40 research institutes, 2 repository/profile organizations, 1 funder, 1 intergovernmental economic organization), 91% of which being research institutes. To the best of our knowledge, none of the French members created ORCIDs on behalf of their staff, though.

On the other hand, authors are free to sign up with their email and obtain an ID. They are asked to pick a visibility setting for their profile: public, restricted to trusted parties only or

Key points

- The ORCID adoption measured at the scale of a French multidisciplinary scientific metropole was 41.8% of its 6,607 staff members.
- The ORCID adoption rate varies throughout the year: it peaked in October when applications to the French national research agency were due.
- The ORCID adoption was higher in science, technology, engineering and mathematics (STEM) versus social sciences and humanities (SSH). The yearly adoption rate differed among disciplines of both STEM and SSH.
- Senior and junior faculty and researchers adopted ORCID at a higher rate compared with postdocs, although postdocs are more likely to use an up-to-date ORCID profile for applications.
- The works section including publications in ORCID profiles are often empty or incomplete. In the studied population, only one-third of the 15 Highly Cited Researchers provided data for the works section.

private (Fig. 1). Several sections make an ORCID profile, as shown in Fig. 2: *Person identifiers*, *Employment*, *Education and qualifications*, *Invited positions and distinctions*, *Membership and service*, *Funding* and *Works*. These are completed by the profile owner and trusted parties upon permission. Other sections are updated by trusted parties only with the owner’s explicit permission: *Peer review* and *Research resources*.⁶ Profile owners can set the visibility of each item in all sections of their profiles. For instance, a publicly available profile may contain items that are available to trusted parties only or even visible to the owner only (e.g., some confidential works).

ORCID operates no control over the contents of the profiles. As a result, many profiles are void of information: all sections appear to be empty, and even the owner’s identity is undisclosed (we elaborate on this in Section 5). There is no way for readers to know if owners entered no data at all or if owners chose to mask some entries (picking a visibility different from ‘Everyone’). While some authors prefer not to disclose information legitimately, Teixeira da Silva (2021) reported some abusive profile creations by papers mills to fool publishers at paper submission time.

ORCID owners can contribute data themselves or use built-in import features. For some sections, such as *Funding* and *Works*, owners provide an identifier (e.g., DOI, PubMed ID, ArXiv ID) or a search query and the system fetches all relevant metadata. Haak

1. See <http://viaf.org>, www.isni.org, and <https://orcid.org>

2. See the live statistics <https://orcid.org/statistics>. The latest annual dump was released in October 2020 <https://doi.org/10.23640/07243.13066970.v1>.

3. <https://www.colorado.edu/fishttps://info.orcid.org/universities-now-creating-orcid-ids-for-their-researchers-and-scholars/>

4. <https://orcid.org/members> and <https://orcid.org/statistics>.

5. <https://www.couperin.org/services-et-prospective/orcid-doaj-sparc-europe-scoss/orcid>

6. See <https://support.orcid.org/hc/articles/360006971333> and <https://support.orcid.org/hc/articles/360011433613>.

Visibility settings

Your ORCID iD connects with your ORCID record that can contain links to your research activities, affiliations, awards, other versions of your name, and more. You control this content and who can see it. [More information on visibility settings.](#)

By default, who should be able to see information added to your ORCID Record?




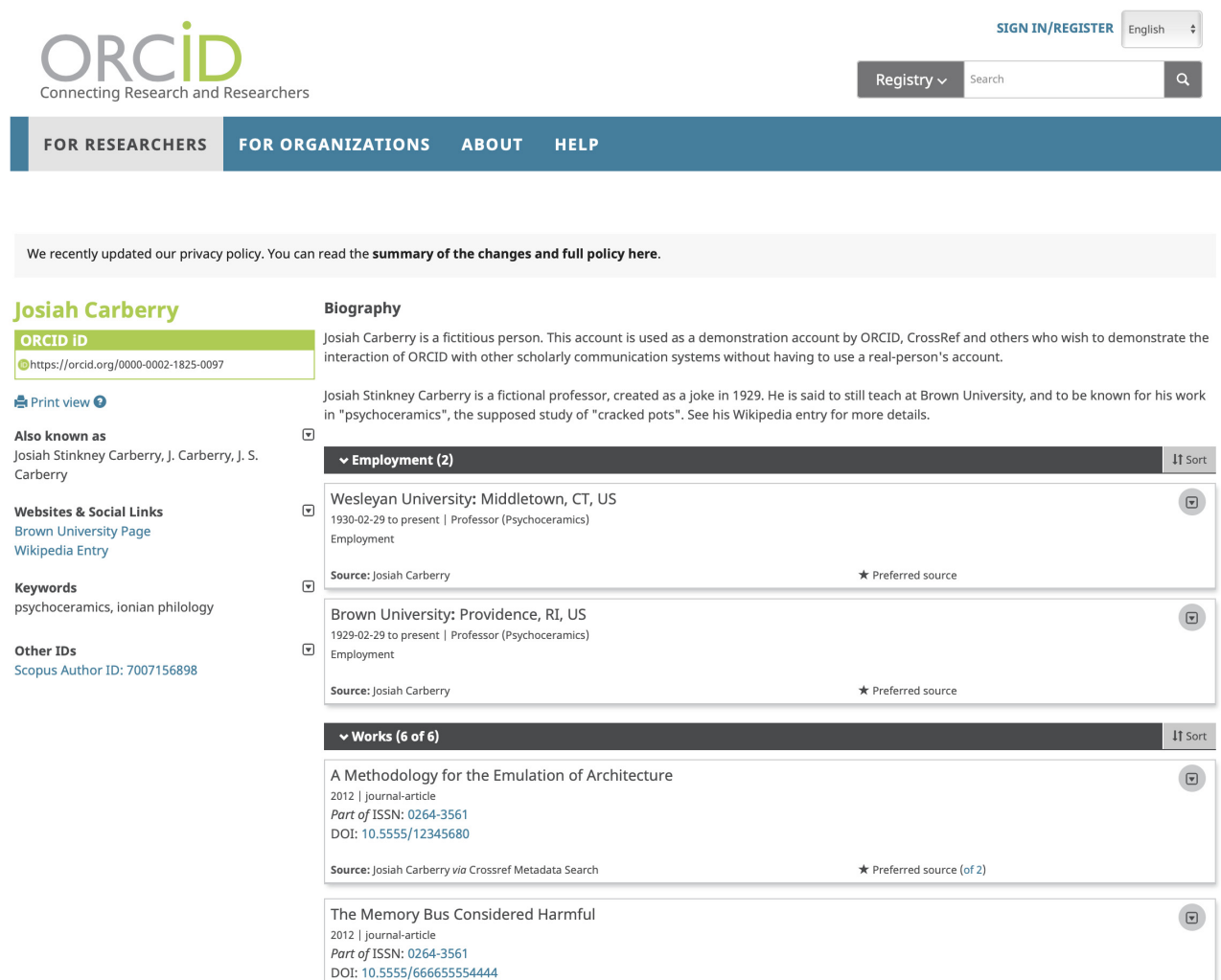
- ☐  **Everyone** (87% of users choose this)
- ☐  **Trusted parties** (5% of users choose this)
- ☐  **Only me** (8% of users choose this)

FIGURE 1 Visibility settings proposed while creating a new ORCID profile at orcid.org.



The screenshot shows the ORCID website interface. At the top, there's a navigation bar with "FOR RESEARCHERS", "FOR ORGANIZATIONS", "ABOUT", and "HELP". Below this, a banner mentions a privacy policy update. The main content area features a demo profile for "Josiah Carberry". The profile includes an ORCID iD, a biography, and sections for "Employment" and "Works".

ORCID iD
<https://orcid.org/0000-0002-1825-0097>

Also known as
 Josiah Stinkney Carberry, J. Carberry, J. S. Carberry

Websites & Social Links
[Brown University Page](#)
[Wikipedia Entry](#)

Keywords
 psychoceramics, Ionian philology

Other IDs
 Scopus Author ID: 7007156898

Biography
 Josiah Carberry is a fictitious person. This account is used as a demonstration account by ORCID, CrossRef and others who wish to demonstrate the interaction of ORCID with other scholarly communication systems without having to use a real-person's account.

Josiah Stinkney Carberry is a fictional professor, created as a joke in 1929. He is said to still teach at Brown University, and to be known for his work in "psychoceramics", the supposed study of "cracked pots". See his Wikipedia entry for more details.

Employment (2)

- Wesleyan University: Middletown, CT, US**
 1930-02-29 to present | Professor (Psychoceramics)
 Employment
 Source: Josiah Carberry (Preferred source)
- Brown University: Providence, RI, US**
 1929-02-29 to present | Professor (Psychoceramics)
 Employment
 Source: Josiah Carberry (Preferred source)

Works (6 of 6)

- A Methodology for the Emulation of Architecture**
 2012 | journal-article
 Part of ISSN: 0264-3561
 DOI: 10.5555/12345680
 Source: Josiah Carberry via Crossref Metadata Search (Preferred source (of 2))
- The Memory Bus Considered Harmful**
 2012 | journal-article
 Part of ISSN: 0264-3561
 DOI: 10.5555/666655554444

FIGURE 2 ORCID demo profile for the fake author Josiah Carberry.

et al. (2019, p. 15) reported that 67% of the 49 million works in ORCID profiles were added through an API provided by Scopus, ResearcherId, PubMed Central and CrossRef mostly. According to an OECD survey of corresponding authors of Scopus-indexed peer-reviewed journals, 68% of the 6,476 corresponding authors

from a representative sample of territories and fields of science had an ORCID (Bello & Galindo-Rueda, 2020, p. 74). This figure might overestimate the adoption of ORCID, as it concerns selective venues and selected scientists (corresponding authors) who are likely to have created an ORCID during the submission process.

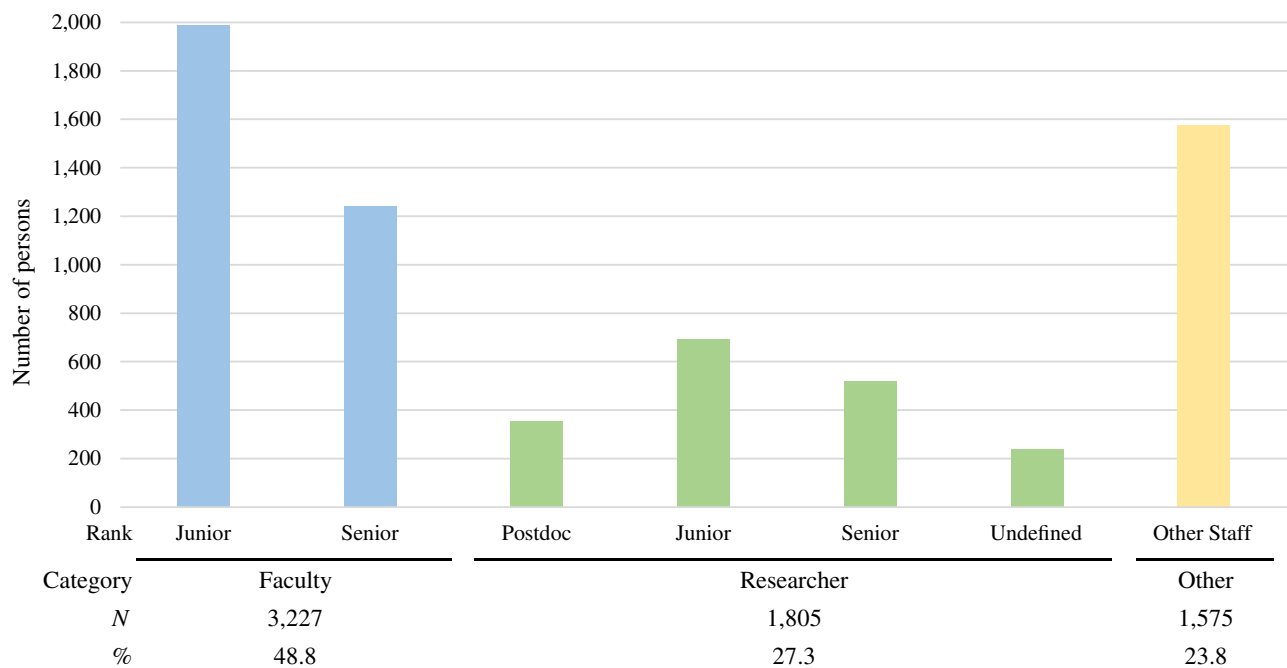


FIGURE 3 Share of the 6,607 persons by job category and rank.

To overcome these biases, we designed a study focusing on a large scientific area in France (Heusse & Cabanac, 2020) gathering 6,471 research staff (a sample of equivalent size compared to the OECD study). We found an adoption rate of about 40% in the Toulouse area, to be compared with the 17% in Caen University, a smaller site accounting for 1,047 research staff members (Boudry & Durand-Barthez, 2020). To the best of our knowledge, there is no other estimate of ORCID adoption in the general research population of a scientific metropole. The present paper extends our previous study published in French (Heusse & Cabanac, 2020) to report the ORCID adoption in the Toulouse scientific area for the period 2012–2020.

CASE STUDY: THE TOULOUSE SCIENTIFIC AREA

This article aims to assess the dynamics of ORCID adoption in one of the leading academic site in France.

About 20 years ago, European countries have harmonized the structure of training at university. The organization of research, however, is not uniform across Europe. This article focusing on a scientific area in France, let us briefly summarize the higher-education landscape in France (Angermüller, 2017; Chevaillier, 2001; Grossetti et al., 2020). This context is important to bear in mind when interpreting the results presented in this paper.

France had 115,308 teaching and research staff (Meuric, 2020, p. 91) in 2017 working as tenured civil servants or contract workers:

- Academics employed by universities and higher schools (*grandes écoles*) share their agenda between research and teaching

(192 hours a year of face-to-face teaching). Nine-tenth are tenured staff (Meuric, 2020, p. 100).

- Full-time research staff employed by national institutes (by decreasing number of researchers: CNRS, INRAE, INSERM, IRD, INRIA and a few smaller institutes with less than 200 researchers) and to a lesser extent by universities and higher schools, dedicate most of their time to research. They are not expected to teach yet some do teach a few hours a year.

Whatever their employer, all staff are affiliated to a public research laboratory (*laboratoire de recherche*) which provides office space and research facilities (Chevaillier, 2001, p. 57). A laboratory is not usually run by a single PI (as in some countries like in the US) but gather many teams (*équipes*) and host from 30 to 500+ members.

The Toulouse scientific area is third after Paris and Lyon in terms of scientific output (Grossetti et al., 2020). More than 100,000 students a year are trained in all disciplines, 45% of which are taught in a masters or doctoral programme. More than 9,000 researchers whose disciplinary breakdown is detailed in Section 4.1 work in 145 public research laboratories. The Toulouse scientific area is known best for its contributions in health, astronomy, universe science and economics.

METHOD AND DATA

This section introduces the data collection protocol we used. We detail how the demographic data retrieved from the institutions were matched to ORCID profiles and manually validated. The

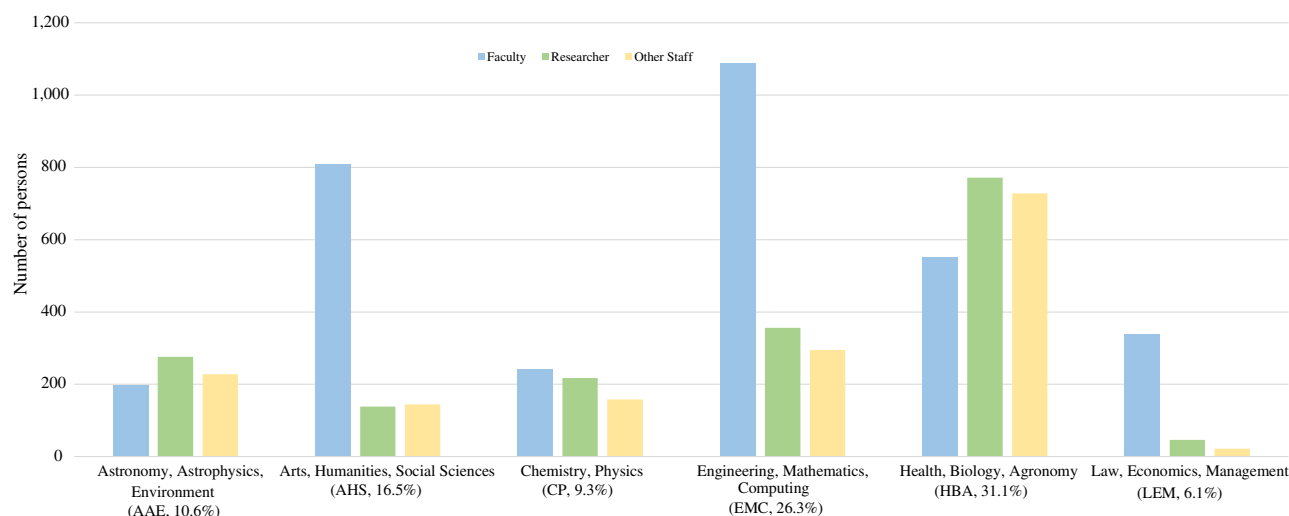


FIGURE 4 Share of the 6,607 persons by discipline and position.

resulting data set supports the longitudinal analysis of ORCID adoption at the scale of a French scientific conurbation. To the best of our knowledge, this is the first study of ORCID adoption at this geographic scale.

Demographics and bibliographic data collection

Demographic data of the 6,607 persons affiliated to the research centres in the Toulouse scientific area were collected in a previous study (Heusse, 2016). This staff registry tabulates 5,029 faculty and researchers as well as 1,578 support staff involved in research activities: engineers, technicians, clinical research fellows. The 4,284 PhD students enrolled in universities and schools (*écoles*) were not included. Each individual included in the registry is characterized by his/her identify (first name, last name), year of birth, sex, institution, job category, rank, laboratory and scientific domain. The staff registry lists individuals (Fig. 3) with the following job categories:

- Faculty (*enseignants-chercheurs*) teach and conduct research activities as junior faculty (*maître de conférences*) or senior faculty (*professeur des universités*).
- Researchers are involved in research and can teach without it being mandatory. Four ranks were considered: postdocs (*post-doctorant*), junior (*chargé de recherche*), senior (*directeur de recherche*) and undefined (e.g., visitors).
- Other Staff refers to support personnel, such as engineers (*ingénieur d'étude* and *ingénieur de recherche*), high-school teachers affiliated to university (*enseignant du second degré*), hospital practitioners (*praticien hospitalier*).

Demographic data were collated by each research structure (145 laboratories). We merged these data into the staff registry of the Toulouse scientific area and tagged each person with the discipline of the lab he/she belongs to. Each lab is associated to one of the six

following discipline groups delineated by the research council of the Toulouse scientific area to reflect its major scientific highlights:

- AAE: Astronomy, astrophysics, environment
- AHS: Arts, humanities, social sciences
- CP: Chemistry, physics
- EMC: Engineering, mathematics, computing
- HBA: Health, biology, agronomy
- LEM: Law, economics, management

Figure 4 shows how the 6,607 distribute among job categories and ranks. The AHS and LEM disciplines have a larger number of faculty compared with researchers and other staff. The situation is reversed in AAE and HBA where most of the personnel is no faculty but researchers and other staff. The large number of other staff in HBA corresponds to clinical research personnel who are active in teaching hospitals (*Centre Hospitalier Universitaire*). The sex ratio is imbalanced across disciplines (Fig. 5), with more males in all disciplines but Art, Humanities, Social Sciences and Health, Biology, Agronomy.

We aimed to retrieve the ORCID profile of the 6,607 individuals by querying the ORCID public API⁷ with their first name and last name. Results were overabundant for some identities such as 'Philippe Durand' that yielded 4,534 profiles. The top-ranked profiles were perfect homonyms and the remaining were partial matches such as "Philippe S. Durand" and "Romain Durand."⁸ We considered the top 20 results per query only. Each retrieved profile was downloaded as an XML file⁹

7. <https://orcid.org/organizations/integrators/API>

8. We noticed that some identities (as listed by the laboratories) differed from the identities authors use in their publications. Consider the case birth name vs. maiden name, for instance. As a result, we did not use phrasal search (e.g., "Philippe Durand" with quotes) to query ORCID profiles.

9. <https://members.orcid.org/api/tutorial/reading-xml>

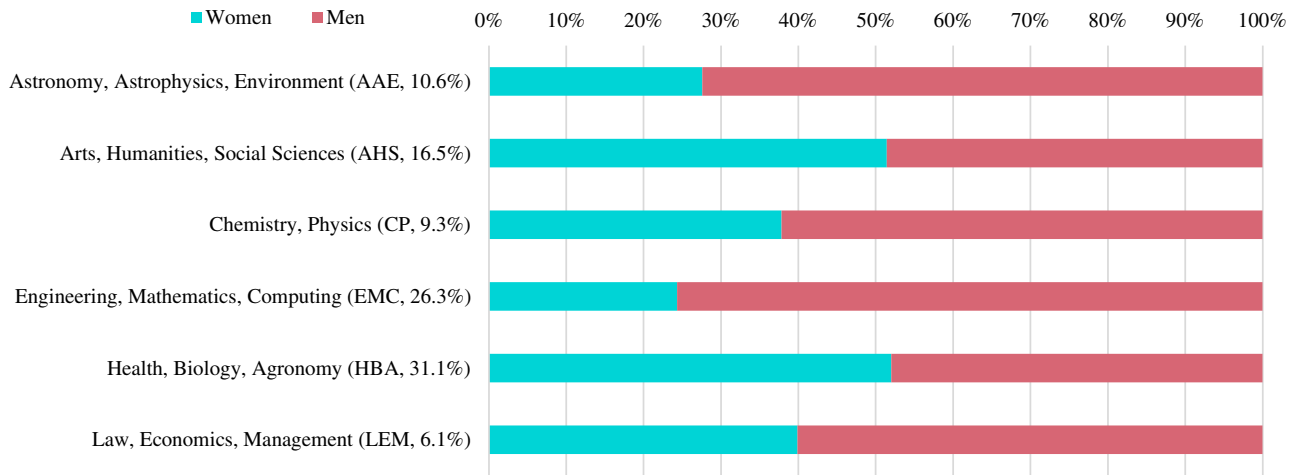


FIGURE 5 Share of the 6,607 persons by discipline and sex.

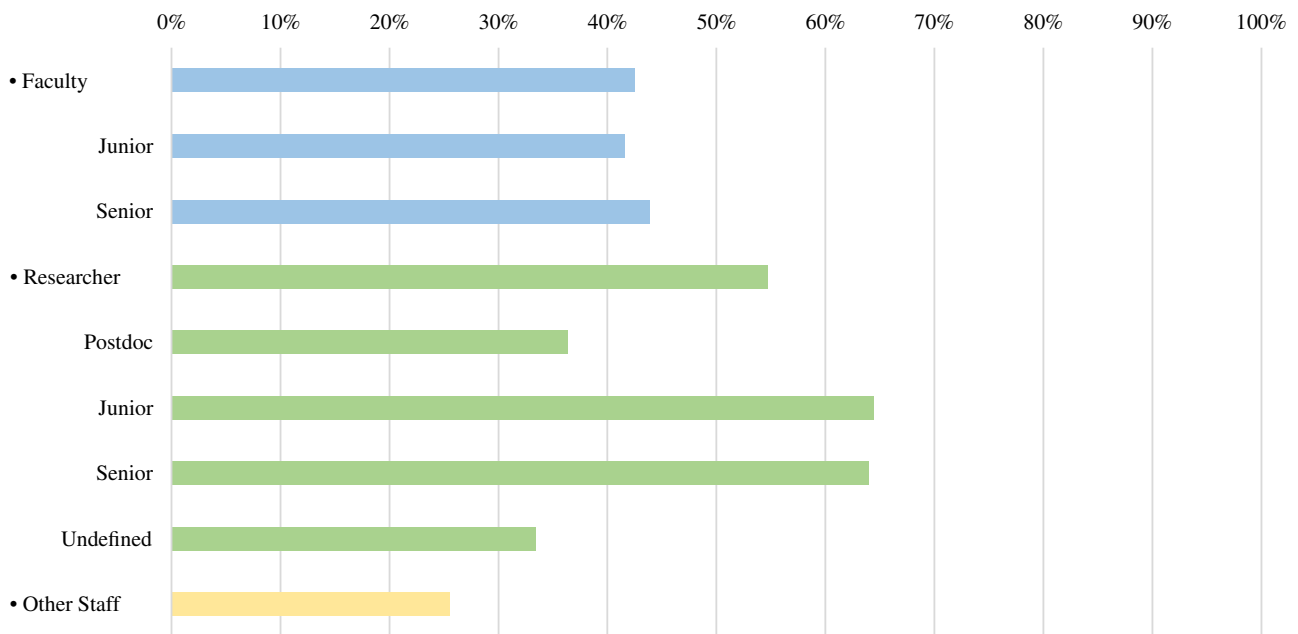


FIGURE 6 ORCID adoption by job category and rank.

recording the profile creation date, the date of the last update, and the person's identity, biography, affiliations and works. Part of these data appear on the online version of ORCID profiles (Fig. 2).

We tagged the profiles mentioning the named entities relevant to the Toulouse scientific area: 12 cities (the largest ones being Auch, Castres, Foix, Rodez, Tarbes and Toulouse), university names (*Champollion*, *Capitole*, *Mirail* and its new name *Jean Jaurés* and *Paul Sabatier*) and 13 school acronyms (e.g., ENSEEIHT, SUPAERO, TBS). This tag with the count of matching entities proved helpful to the visual inspection performed, that we detail in the next section.

Matching the staff registry to ORCID profiles: Semi-automated approach

With the ORCID data collected, we faced three cases for a given person's identity used as query:

1. No matching ORCID profile.
2. One matching ORCID profile.
3. Between 2 and 20 matching ORCID profiles.

The first author (MDH) performed a visual inspection of all matches. Candidate profiles were validated based on the

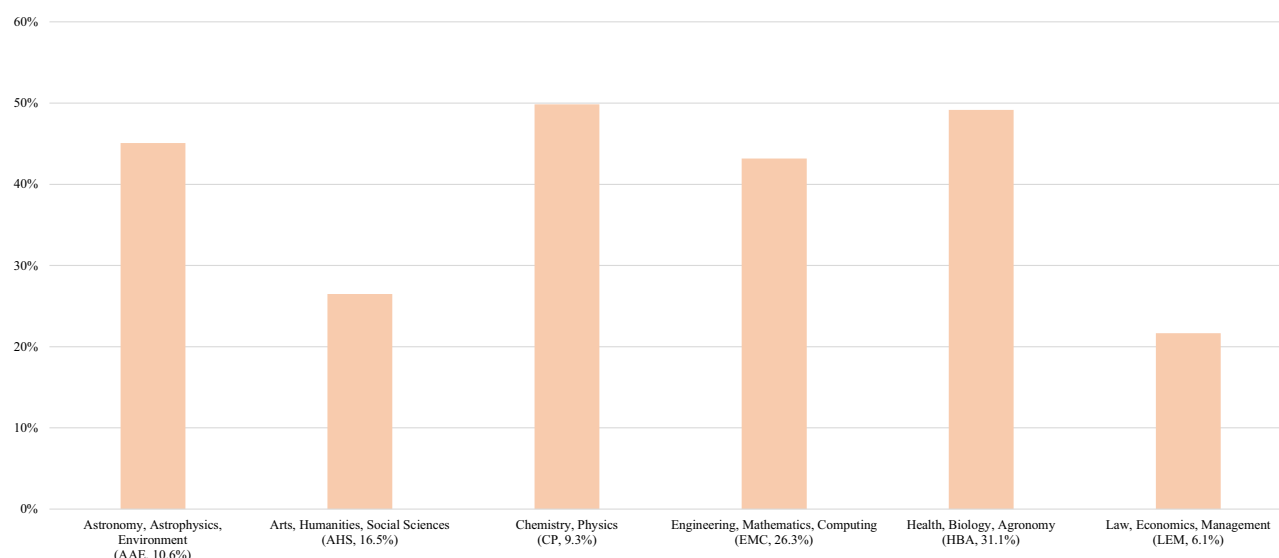


FIGURE 7 ORCID adoption by discipline.

biography and affiliations (listed in the profiles and on the publications appearing as works). The major pitfall concerned 'empty' profiles registered for a given identity but showing no works. Such profiles were validated with a quasi-certainty for discriminant identities (rare associations of first and last names), using Google Scholar to check the absence of homonyms.

Each pair of matching identity–profile pair was annotated with one of these codes:

- Code '0' when the looked-up identity and the ORCID profile did not match (e.g., different first names).
- Code '1' when the looked-up identity and the ORCID profile matched and the profile contents featured evidence of the Toulouse scientific area.
- For empty ORCID profiles:
 - Code '?' when no homonyms were detected via Google Scholar searches for the looked-up identity, suggesting that there is only one known author with this identity.
 - Code '??' when homonyms were detected via Google Scholar searches.

We considered as ORCID adopters the 2,789 identities with code '1' or '?'. This is a conservative estimate of the number of ORCID adopters in the Toulouse scientific area since some identities coded '??' were discarded because of lacking evidence.

QUANTITATIVE RESULTS

As of September 2020, 41.8% of the considered population had an ORCID profile. Setting aside the 'Other staff' job category, the ORCID adoption amounts to 46.9%. These two figures reflect

some disparities when broken down by job category or discipline, as discussed in this section.

ORCID adoption by job category

Researchers were the most likely (54.7%) to register an ORCID profile, with a varying adoption per rank (Fig. 6). Junior and senior tenured researchers registered more ORCID profiles (64.0% on average) than postdocs (36.4%). This imbalance is all the more surprising as postdocs usually have an intense research activity and they are expected to care for the visibility of their research (Nicholas et al. 2020).

Faculty members were less likely to register an ORCID (42.5%) and there is little difference with respect to seniority. This lesser adoption can result from two factors. On the one hand, faculty members are overrepresented in disciplines with a lesser use of ORCIDs: *Art, Humanities, Social Sciences* and *Law, Economics, Management*. On the second hand, part of the faculty members are more active in teaching and management roles than in research *per se*.

The other staff category has the lowest ORCID adoption rate (25.5%). They are engineers, technicians and assistants part of a research lab. Their work is acknowledged in the papers and they are sometimes co-authors of their lab's publications.

ORCID adoption by discipline and sex

The average ORCID adoption per discipline is 39.2%, ranging from 21.7% for *Law, Economics, Management* to 49.8% for *Physics, Chemistry* (Fig. 7). This result is in line with the publication habits of these disciplines and the use of identifiers that is well established in the hard sciences. The sex imbalance of adopters reflects the one we observed in the population (Fig. 5) except for

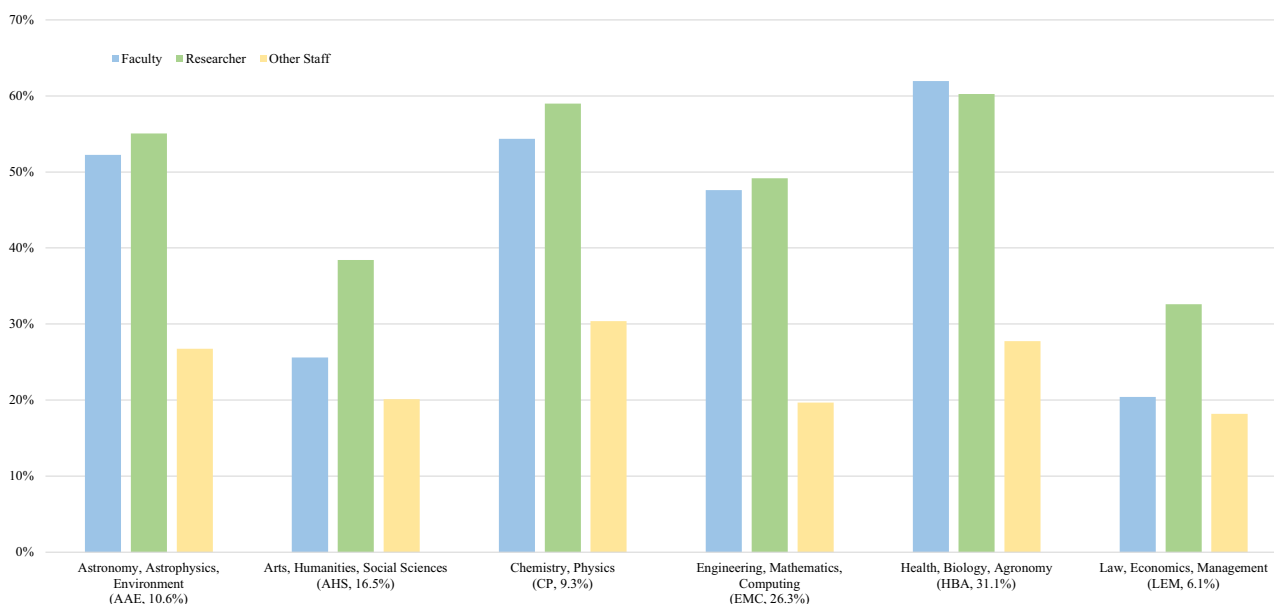


FIGURE 8 ORCID adoption by job category and discipline.

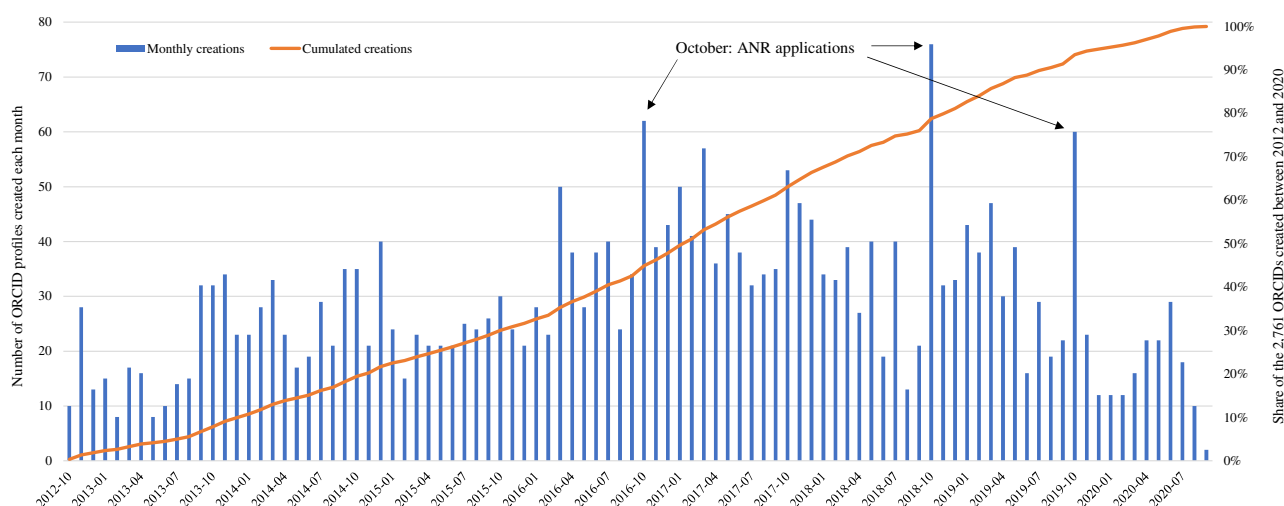


FIGURE 9 ORCID profile creation through time.

Health, Biology, Agronomy where women adopters are 5% less than women versus men in this discipline.

Figure 8 shows the adoption of each discipline per job category. Researchers are top adopters for all disciplines but *Health, Biology, Agronomy*. The promotion of ORCID by research institutes¹⁰ might be a favouring factor here. In *Law, Economics, Management* researchers and faculty members adopted ORCID at a lower rate than average. Other Staff have a high adoption rate in AAE, HBA and PC; these are disciplines whose labs were more likely to list support personnel as co-authors (see Section 6.2).

10. See for instance <https://doranum.fr/wp-content/uploads/DoRANum-fiche-ORCID-maj-2020-10-08.pdf>.

ORCID adoption through time

ORCID was launched on 16 October 2012. Each profile retrieved via the ORCID API includes its creation date. We used this information to plot the dynamics of ORCID adoption. The first ORCID profile created by a researcher in the Toulouse area dates back from 20 October 2012. They were 51 early adopters in 2012 from the Toulouse area: 10 in October, 28 in November and 13 in December.

The number of newly created ORCID profiles between October 2012 and September 2020 is shown in Fig. 9. Creations follow a steady rate with peaks in October, which coincides with grant applications to the French National Research Agency (*Agence Nationale de la Recherche*). This agency requires

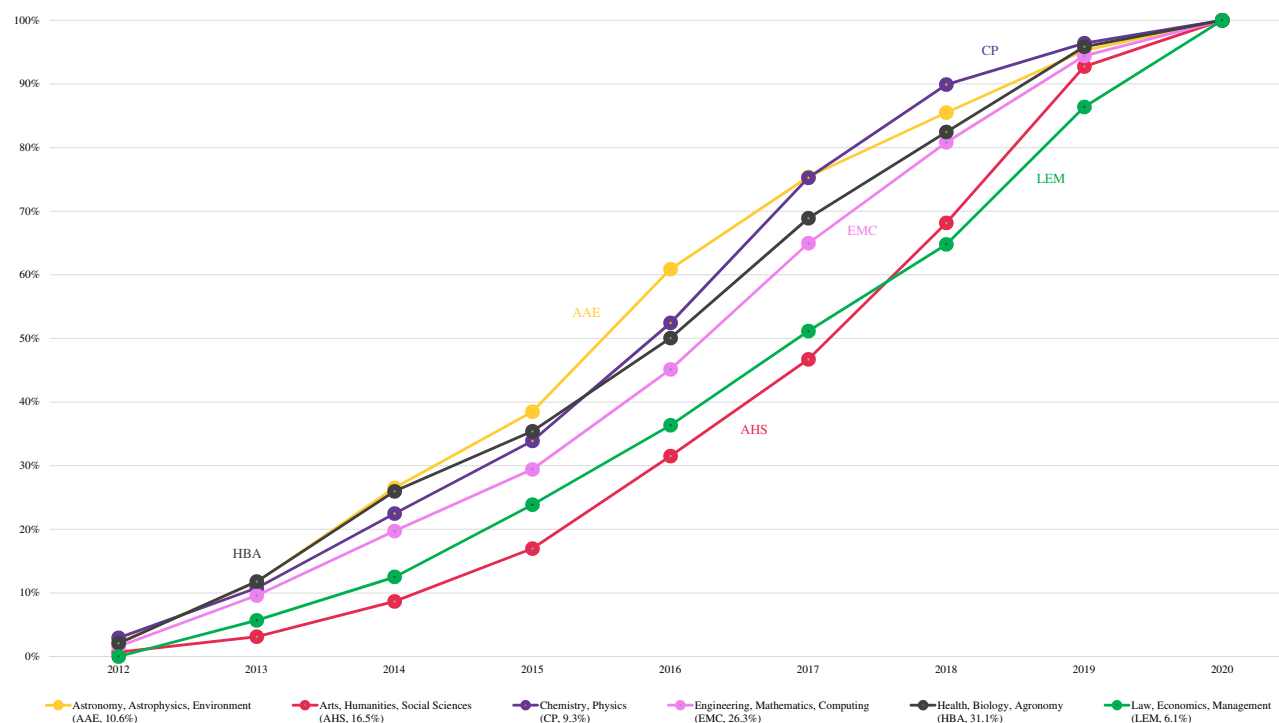


FIGURE 10 Cumulated share of ORCID profiles by discipline.

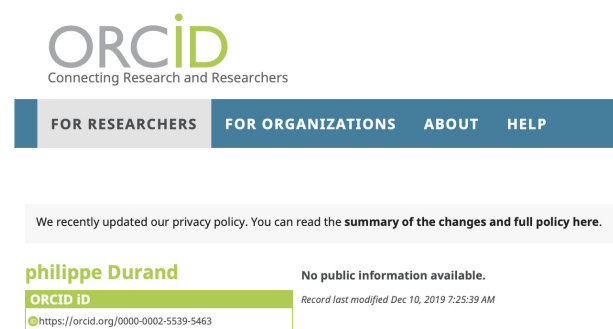


FIGURE 11 Example of an 'empty' ORCID profile.

consortium members to be identified with an ORCID: usually the PI and at least one partner per institution involved.¹¹

We split the cumulated red line of Fig. 9 to plot the cumulated number of profile creation for each discipline (Fig. 10). Disciplines gather into two groups characterized by different growth types. Throughout the 2012–2020 period, AAE–EMC–HBA–PC showed a greater adoption rate compared to AHS–LEM. For the first group, HBA adopted ORCID early, then AAE became a forefront adopter. Despite a slow start, PC is nowadays the leading adopting discipline. For the second group with lower adoption, LEM was quicker to create ORCID profiles compared to AHS until 2018 when AHS showed the largest

adoption rate among all six disciplines. The rationale behind the changing adoption speed is unknown. We can only speculate on the different publishing houses active in EMC–AAE–HBA–PC (i.e., STEM–Science, Technology, Engineering, and Mathematics) versus AHS–LEM (i.e., SSH–Social Sciences and Humanities) and their use of ORCID:

- STEM publishers are generally large-sized firms (e.g., Elsevier, Springer, Wiley). Most of them are members of the ORCID consortium and they integrated ORCID into their peer-review and production system. Authors and reviewers are encouraged to create and link their ORCID to their profile on the submission/review platform (Johnson et al., 2018, p. 162).
- SSH publishers are more diverse and smaller publishing houses (e.g., local University Presses). Many are small-sized companies publishing works on a specific topic, such as Dalloz in Law, Vrin in Philosophy, Presses Universitaires de Rennes in the SSH. Most of these publishers are not members of the ORCID consortium and their authors are not asked to provide an ORCID at submission stage.

QUALITATIVE RESULTS

Diversity of ORCID profile (mis)uses

Let us recall that researchers from the Toulouse area register an ORCID themselves (no institutional automated registration is

11. See page 20 in <https://anr.fr/fileadmin/aap/2019/aapg-anr-2019-Guide.pdf>

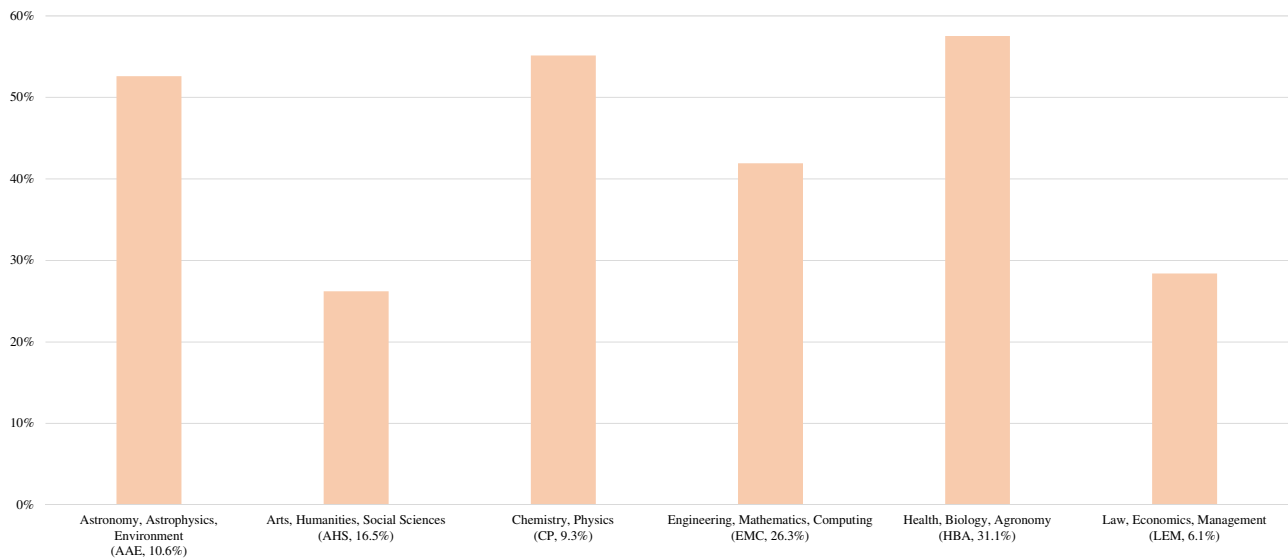


FIGURE 12 Share of ORCID profiles with at least one publication.

TABLE 1 Online presence of 15 *Highly Cited Researchers* (HCR) as per Clarivate Analytics with their ORCID and Google Scholar profiles, if any

ID	Discipline	ORCID				Google Scholar
		Profile	Biography	Affiliation	Publications	
1	Agronomy	x		x	x	x
2	Biology	x		x		
3	Ecology					x
4	Ecology	x		x		x
5	Economics					x
6	Engineering	x		x	x	x
7	Health	x				x
8	Health	x			x	x
9	Health	x	x		x	
10	Health					
11	Health	x				
12	Health	x				
13	Health	x		x	x	
14	Health	x			x	
15	Astronomy	x				x

Note: ORCID owners (with a Profile) might have a Biography, an Affiliation and Publications listed or not.

performed) and fill the associated sections, such as identity, employment biography and works (Section 2). The ideal ORCID profile as illustrated in Fig. 2 lists comprehensive data about the author and his/her publications. Looking at the ORCID profile in our corpus, we found that many are far from such thoroughly completed profile with up-to-date data. Most profiles appear to be incomplete and even empty: Fig. 11 only shows ‘philippe

Durand’ as the author’s identity (there were five such empty profiles for this name at the time of data extraction). This homonym issue is even worse with surnames such as Wang (Youtie et al., 2017).

For our corpus, Fig. 12 shows the percentage of ORCID profiles with at least one publication listed. All disciplines considered, only 48.3% of the profiles feature one or more works. This is

larger than the 26.8% expected from the ORCID statistics page.¹² Using a per discipline breakdown, *Health, Biology, Agronomy* has the largest rate of completed profiles (57.5%). *Arts, Humanities, Social Sciences* is the least adopting discipline with only 26.2% of completed profiles (but certainly not completed comprehensively). There is no direct way to check how comprehensive an ORCID profile is. One should consider the works listed with care and do not infer any selection process from the owner of the profile: the listed works might just have been pushed by a few publishers. Another point relates to the default visibility setting for new data added. When profile owners decide 'private by default' (Fig. 1) and never connect to ORCID and change the visibility to public, the works are attached to the profiles but invisible to the public.

Based on the visual inspection of 100+ profiles, we noted the following intriguing characteristics regarding two aspects:

- The Biography and Employment sections are supposed to list the successive affiliations of the profile owner. Yet, some users list foreign universities that never employed them. It appeared that these universities are affiliations of colleagues with whom the profile owner collaborated. There is also a propensity to list prestigious institutions. For researchers employed in French national research organisms (e.g., CNRS and INSERM) and affiliated to a lab in the Toulouse area, we noted that some of them mention the Paris headquarters (e.g., CNRS Paris) instead of the regional lab (e.g., CNRS IRIT Toulouse). And yet, they use their 'local' affiliation in the published papers.
- The works section does not seem comprehensive for most ORCID profiles that contain only a fraction of the publications of the profile owner. We noticed several profiles of highly productive researchers missing hundreds of bibliographic records when compared with the Web of Science. We also raised quality concerns: some bibliographic records did not show any publication date, some titles and authors were misspelled. These errors suggest that profile owners entered the bibliographic data manually themselves. They might have ignored that they could have automatically retrieved metadata by providing the DOIs or other identifiers they wanted to add to their profiles.

A researcher can create multiple ORCID profiles linked to a several of his/her email addresses. One can only speculate that it was easier and faster to create a new ORCID instead of searching in one's archives for the ID and associated password created a long time ago. We found 18 individuals from our corpus who created multiple ORCID profiles: eight cases of a completed versus empty profile, seven cases of two empty profiles, and three cases of completed versus completed profile (with a varying degree of completion). The profile created last was not always the one completed by the researchers. We do not

suspect any fraud related to fake profile creation as reported in (Teixeira da Silva, 2021).

One may wonder: why do not people create ORCID profiles and, when they own one, why do not they fill it properly? ORCID is not the only platform allowing the creation of a profile for an author to list his/her publications (Boudry & Durand-Barthez, 2020; French & Fagan, 2019; Tran & Lyon, 2017). Some authors might own such a profile and view ORCID as a platform providing identifiers only, disregarding its bibliography curation capability. These authors might invest more efforts into Google Scholar and the likes at the expense of their ORCID profiles. We tested this hypothesis by tabulating (Table 1) the use of ORCID and Google Scholar for 15 individuals from the staff registry who were listed as 'Highly Cited Researchers' by Clarivate Analytics.¹³ No clear conclusions emerge: a majority of the most productive researchers in the Toulouse area failed to complete their ORCID profiles although 80% of them created an ORCID. Surprisingly, one of these Highly Cited Researchers in health sciences has no online presence at all, neither on ORCID nor on Google Scholar.

Focus on the ORCID profiles by 'other staff'

Slightly more than 25% of the 'other staff' have created an ORCID profile (Section 5.1). Including these staff as co-authors or not varies according to the disciplines. Other staff were included as co-authors in 7 out of the 30 labs in *Health, Biology, Agronomy*. It also happened in *Astronomy, Astrophysics, Environment and Physics, Chemistry*. Some extra cases occurred in *Engineering, Mathematics, Computing*. This virtually did not happen in *Law, Economics, Management and Art, Humanities, Social Sciences*. Being listed as co-authors is an incentive for the creation of ORCID profiles among the 'other staff' job category.

CONCLUSION

Introduced in 2012, ORCID has become a key global infrastructure to disambiguate scholarly authors' identities. This study reported on the adoption of ORCID in one of the largest scientific areas in France. Contrary to Youtie et al. (2017) who focused on Wang as a single surname of interest, we considered the entire multidisciplinary research workforce of a French scientific metropole. Matching the 6,607 staff members affiliated to the 145 labs based in the Toulouse area in 2016 with the ORCID registry, we showed a steadily increase for all research disciplines. The overall adoption is 41.8% with varying percentages among job ranks and disciplines. Faculty members in *Health, Biology, Agronomy* are leading adopters (62.0%) as opposed to those in *Law, Economics, Management* with 20.4% ORCID profiles. Such differences might result from various contexts: discipline-related incentives enacted by research institutions, funding agencies and the publishing industry. For the 60% ORCID

12. <https://orcid.org/statistics>

13. <https://recognition.webofsciencegroup.com/awards/highly-cited/2019/>

profiles with contents provided, a qualitative analysis reveals they were not comprehensively filled. Profile sections documenting the biography, employment history, other identifiers, awards and grants received are frequently lacking. The same goes for publications with incomplete lists of works. We acknowledge that the people hired between 2017 and 2021 are not considered in this study. Likewise the people who left Toulouse since 2016 are included in this study performed in 2020. All in all, our results suggest that a majority of researchers in the Toulouse area have little use of ORCID's or understanding (or interest?) for identifiers.

There are only a few comparable studies of ORCID adoption worldwide. Dasler et al. (2017) analyzed the ORCID profiles created over 2012–2016, stressing breakdowns by discipline and location. This study at the global scale lacks a reference population, contrary to two studies on French scientific areas published later. The adoption in Toulouse of 41.8% (the present study) is higher than the 17.1% reported by (Boudry & Durand-Barthez, 2020) for the Caen University which is smaller in research workforce. We have no clue about this threefold observed difference and more research is needed to decipher this imbalance. The limited adoption of ORCID—and related misuses—calls for a better education to the issue of identity management for academics and the use of ORCID to tackle the issue of homonyms among authors worldwide, as stressed by the Surgisphere affair (Piller, 2020).

ACKNOWLEDGEMENTS

We are grateful to the two referees and the Editor-in-Chief for the invaluable remarks and comments that improved the manuscript.

AUTHOR CONTRIBUTIONS

GC and MDH: Conceptualization; Methodology; Formal analysis; Visualization; Writing – original draft; Writing – review & editing; GC: Software; MDH: Data curation.

REFERENCES

- Angermuller, J. (2017). Academic careers and the valuation of academics. A discursive perspective on status categories and academic salaries in France as compared to the U.S., Germany and Great Britain. *Higher Education*, 73(6), 963–980. <https://doi.org/10.1007/s10734-017-0117-1>
- Bello, M., & Galindo-Rueda, F. (2020). *The 2018 OECD international survey of scientific authors* (OECD Science, Technology and Industry Working Papers). OECD. <https://doi.org/10.1787/18d3bf19-en>
- 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*, 15(9), e0238583. <https://doi.org/10.1371/journal.pone.0238583>
- Brown, J., Demeranville, T., & Meadows, A. (2016). Open access in context: Connecting authors, publications and workflows using ORCID identifiers. *Publications*, 4(4), 30. <https://doi.org/10.3390/publications4040030>
- Chevallier, T. (2001). French academics: Between the professions and the civil service. *Higher Education*, 41(1–2), 49–75. <https://doi.org/10.1023/a:1026758611096>
- Dasler, R., Deane-Pratt, A., Lavasa, A., Rueda, L., & Dallmeier-Tiessen, S. (2017). *Study of ORCID adoption across disciplines and locations* (Tech. Rep.). British Library. Horizon 2020 THOR project report. doi: <https://doi.org/10.5281/zenodo.841777>
- Fabre, R., Egret, D., Schöpfel, J., & Azeroual, O. (2021). Evaluating the scientific impact of research infrastructures: The role of current research information systems. *Quantitative Science Studies*, 2(1), 42–64. https://doi.org/10.1162/qss_a_00111
- French, R. B., & Fagan, J. C. (2019). The visibility of authority records, researcher identifiers, academic social networking profiles, and related faculty publications in search engine results. *Journal of Web Librarianship*, 13(2), 156–197. <https://doi.org/10.1080/19322909.2019.1591324>
- Grossetti, M., Maisonobe, M., Jégou, L., Milard, B., & Cabanac, G. (2020). wSpatial organisation of french research from the scholarly publication standpoint (1999–2017): Long-standing dynamics and policy-induced disorder. In C. S. Lévy (Ed.), *EPJ web of conferences* (Vol. 244, 01005). EDP Sciences. Complexity and Disorder Meetings 2018–2019–2020. <https://doi.org/10.1051/epjconf/202024401005>
- Haak, L. L., Fenner, M., Paglione, L., Pentz, E., & Ratner, H. (2012). ORCID: A system to uniquely identify researchers. *Learned Publishing*, 25(4), 259–264. <https://doi.org/10.1087/20120404>
- Haak, L. L., Meadows, A., & Brown, J. (2018). Using ORCID, DOI, and other open identifiers in research evaluation. *Frontiers in Research Metrics and Analytics*, 3. <https://doi.org/10.3389/frma.2018.00028>
- Haak, L. L., Petro, J. A., Simpson, W., Demeranville, T., Wijnbergen, I., Hershberger, S., Wilmers, C., Kiermer, V., Hook, D., & Buster, M. (2019). ORCID 2019 annual report (Tech. Rep.). ORCID. Retrieved from. <https://doi.org/10.23640/07243.12009153.v1>
- Hanson, B., Lawrence, R., Meadows, A., & Paglione, L. (2016). Early adopters of ORCID functionality enabling recognition of peer review: Two brief case studies. *Learned Publishing*, 29(1), 60–63. <https://doi.org/10.1002/leap.1004>
- Harzing, A.-W. (2015). Health warning: Might contain multiple personalities—the problem of homonyms in Thomson Reuters essential science indicators. *Scientometrics*, 105(3), 2259–2270. <https://doi.org/10.1007/s11192-015-1699-y>
- Heusse, M.-D. (2016). Faire parler les données sur la recherche grâce au web sémantique: le projet VIVO. In *Actes de l'atelier Valorisation et Analyse des Données de la Recherche (VADOR) organisé dans le cadre du congrès* (pp. 19–25). INFORSID.
- Heusse, M.-D., & Cabanac, G. (2020). Adoption de l'identifiant chercheur ORCID: le cas des universités toulousaines. In S. Dupuy-Chessa (Ed.), *INFORSID'20: 38^e congrès de l'Informatique des Organisations et Systèmes d'Information et de Décision* (pp. 19–34). Éditions Inforsid. Retrieved from. http://inforsid.fr/actes/2020/INFORSID_2020_p19-34.pdf
- Johnson, R., Watkinson, A., & Mabe, M. (2018). *The STM report: An overview of scientific and scholarly publishing* (5th ed.). International Association of Scientific, Technical and Medical Publishers. Retrieved from. https://www.stm-assoc.org/2018_10_04_STM_Report_2018.pdf
- Meuric, L. (Ed.). (2020). *L'état de l'Emploi scientifique en France*. Ministère de l'enseignement supérieur, de la recherche et de l'innovation. Retrieved from. <http://www.enseignementsup-recherche.gouv.fr/pid35205/etat-de-l-emploi-scientifique.html>

- Nicholas, D., Watkinson, A., Abrizah, A., Rodríguez-Bravo, B., Boukacem-Zeghmouri, C., Xu, J., Świgoń, M., & Herman, E. (2020). Does the scholarly communication system satisfy the beliefs and aspirations of new researchers? Summarizing the Harbingers research. *Learned Publishing*, 33(2), 132–141. <https://doi.org/10.1002/leap.1284>
- Piller, C. (2020). Who's to blame? These three scientists are at the heart of the Surgisphere COVID-19 scandal. *Science*. <https://doi.org/10.1126/science.abd2252>
- Sabine, S. (2014). You know who you are, but do others? *Share: the newsletter of the Australian National Data Service* 18. Retrieved from https://www.ands.org.au/__data/assets/pdf_file/0004/388822/share-issue-18.pdf
- Sivertsen, G. (2019). Developing current research information systems (CRIS) as data sources for studies of research. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 667–683). Springer. https://doi.org/10.1007/978-3-030-02511-3_25
- Soete, L., Schneegans, S., Eröcal, D., Angathevar, B., & Rasiah, R. (2015). A world in search of an effective growth strategy. In S. Schneegans (Ed.), *UNESCO science report: Towards 2030* (pp. 20–55). Paris: UNESCO Publishing. Retrieved from <http://unesdoc.unesco.org/images/0023/002354/235406e.pdf>
- Teixeira da Silva, J. A. (2021). Abuse of ORCID's weaknesses by authors who use paper mills. *Scientometrics*, 126(7), 6119–6125. <https://doi.org/10.1007/s11192-021-03996-x>
- Tran, C. Y., & Lyon, J. A. (2017). Faculty use of author identifiers and researcher networking tools. *College & Research Libraries*, 78(2), 171–182. <https://doi.org/10.5860/crl.78.2.171>
- Youtie, J., Carley, S., Porter, A. L., & Shapira, P. (2017). Tracking researchers and their outputs: New insights from ORCIDs. *Scientometrics*, 113(1), 437–453. <https://doi.org/10.1007/s11192-017-2473-0>