



HAL
open science

A study on data curation perceptions and practices in France

Alina Danciu, Nicolas Sauger

► **To cite this version:**

Alina Danciu, Nicolas Sauger. A study on data curation perceptions and practices in France. European DDI Users Conference (EDDI19), Dec 2019, Tampere University (FSD), Tampere, Finland. hal-02470555

HAL Id: hal-02470555

<https://sciencespo.hal.science/hal-02470555>

Submitted on 15 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

What do you hear when I say data curators? A study on data curation perceptions and practices in France.

Alina Danciu, Nicolas Sauger

EDDI19 – 11th Annual European DDI Users Conference

Tampere, Finland

December 3, 2019 – December 4, 2019

The rise of the data curator?

Curation, traditionally: collections of artifacts in a museum and works of art in a gallery

*Data curators prepare and enrich research data to make them findable, accessible, interoperable and reusable (FAIR)**

Data curation

- ***Data curation** is the active and on-going management of data through its lifecycle of interest and usefulness to scholarship, science, and education*
- ***Curation activities** enable data discovery and retrieval, maintain quality, add value, and provide for re-use over time**
- The Inter-university Consortium for Political and Social Research (ICPSR) website marks **1962 as the date of their first Survey Data Archive.**

*(Cragin, Melissa; Heidorn, P. Bryan; Palmer, Carole L.; Smith, Linda C. (2007). "An Educational Program on Data Curation". ALA Science & Technology Section Conference. Retrieved 7 October 2013.)

Data curation in France

- **Legislative context** evolves in France (Vidal National Plan for Open Science, GDPR)
- Incitation to **publish research data**
- Different professionals: archivists, data curators, librarians, DPOs, researchers....
- Some have a *traditional* role, others are newer to this field
- Different statuses in France



Let's take a closer look!

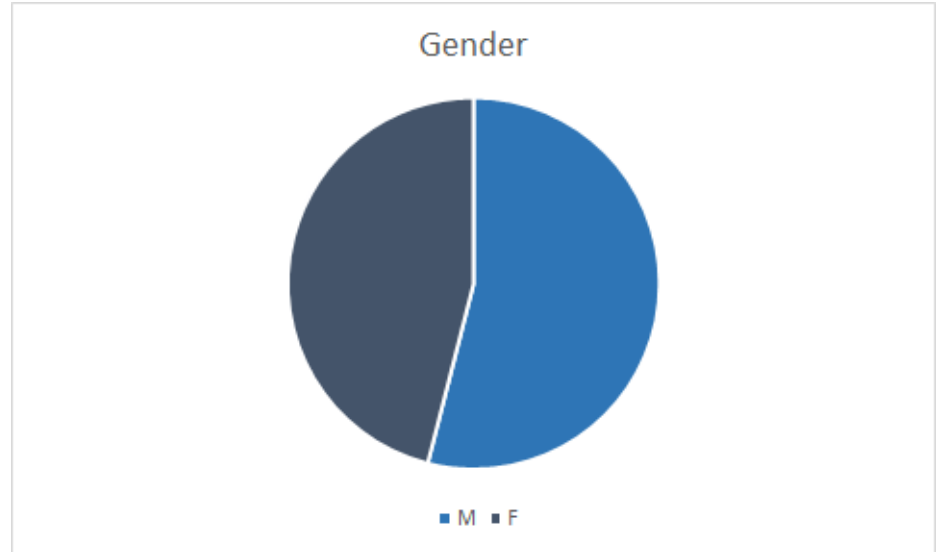
- **Online survey** sent to all the mailing lists of the engineers/research officers of the French National Centre for Scientific Research + RDA France
- Also published on **Twitter** (Impressions: 2849 | Engagements: 102)
- **Fieldwork:** 2019 November, 12 - 2019 November, 27
- 215 responses

W
H
O
A
N
S
W
E
R
E
D

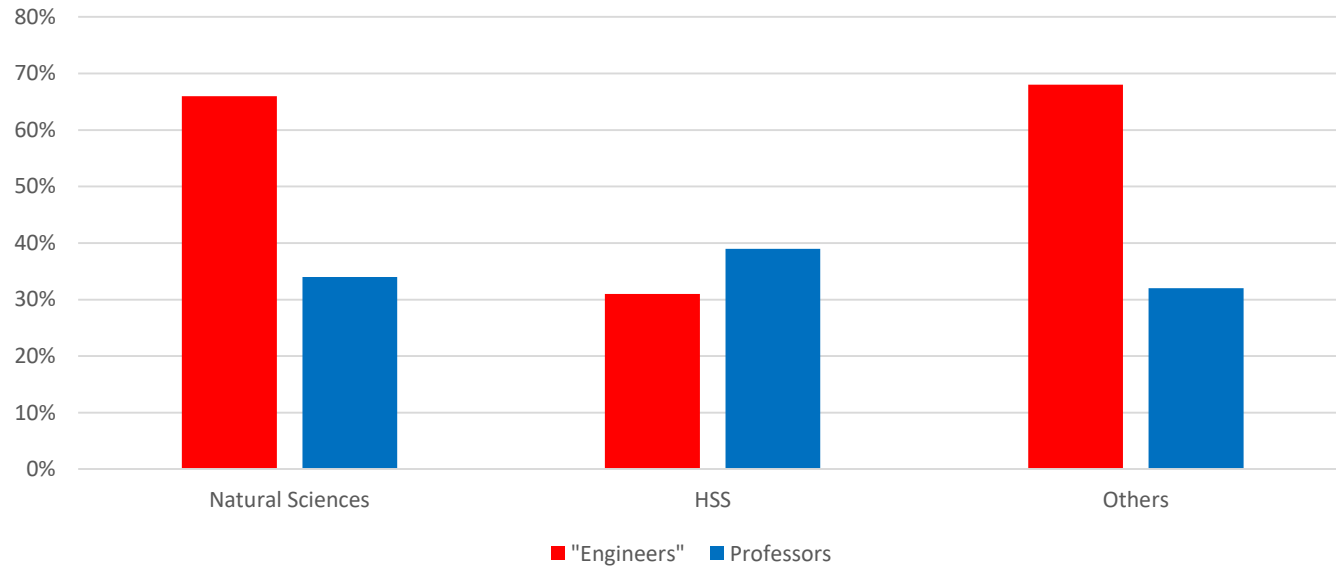
| PROFESSION | % | Count |
|---|--------|-------|
| Archivist | 1.86% | 4 |
| Librarian | 0.93% | 2 |
| Chercheur-se ou enseignant-chercheur-se | 32.09% | 69 |
| Documentalist | 3.72% | 8 |
| Engineers | 46.98% | 101 |
| Post-doctoral researcher | 4.19% | 9 |
| Technician or Assistant Engineer | 4.19% | 9 |
| Others | 6.05% | 13 |
| Total | 100% | 215 |

Who answered: age and gender

| AGE | % |
|-------------------|--------|
| 18 - 29 years old | 8.45% |
| 30 - 49 years old | 68.08% |
| 50 - 64 years old | 22.07% |
| 65 years and more | 1.41% |



Professional status vs areas of academic studies



Nota bene: differences not statistically significant


Amount of time dedicated to "data activities" in comparison to other activities by status

| | Engineers | Professors |
|-------------------------------|--------------|---------------|
| Data collection | 17.5% | 14% |
| Data management | 26.6% | 17.2% |
| Data analysis for publication | 12.5% | 16.3% |
| Other activities | 43.3 % | 52.4 % |

Amount of time dedicated to "data activities" in comparison to other activities by data domain

| | Natural sciences | HSS | Others |
|-------------------------------|------------------|--------------|--------|
| Data collection | 16.8% | 16.9% | 12.8% |
| Data management | 25.7% | 22% | 20.2% |
| Data analysis for publication | 11.1% | 16.8% | 11.8% |
| Other activities | 46.3% | 44.3% | 55.2% |

Standards

- 105 people declared using a standard when documenting their data
- Research officers more likely to use standards to document their data: 52% use at least one
- In natural sciences, people are more likely to use a standard
- From 81 people who cited at least one standard: 8 cited  DDI[®]
- Other standards used by people using DDI: DC, EAD, Isa-TAB, METS, SEDA

| | Natural sciences | HSS | Others |
|-----|------------------|-------|--------|
| Yes | 59.8% | 32.0% | 63.3% |
| No | 40.2% | 68.0% | 36.7% |

Have you already...

Published data on a repository?

| | Natural sciences | HSS | Others |
|-----|------------------|------|--------|
| Yes | 56.1 | 44.1 | 53.3 |
| No | 43.9 | 55.9 | 46.7 |

How often do you perform these activities?

Data documentation : 89,6% of the engineers declare to have already done it (professors: 76,6%)

Data integration in information systems: 74,8% of the engineers (professors: 61,1%)

| Data integration in information systems | Natural sciences | HSS |
|---|------------------|-------|
| Regularly | 49.4% | 31.4% |
| Sometimes | 13.6% | 14.7% |
| Rarely | 12.3% | 17.6% |
| Never | 24.7% | 36.3% |

| Long-term archiving | Natural sciences | HSS |
|---------------------|------------------|------|
| Regularly | 42.0 | 19.6 |
| Sometimes | 19.7 | 20.6 |
| Rarely | 17.3 | 25.5 |
| Never | 21.0 | 34.3 |

Have you already payed to:

- **Have access to an article: 37.9%**
- Publish an article: 26.7%
- Have access to data: 28.4%
- **Publish data: 5.7%**

What's next?

- **Data will be published** after an embargo period
- We intend to take a closer look at the data
- **Get in touch** with us in case you would like to replicate this study in your country!
- We can help with the survey tools

THANK YOU!

Get in touch!

alina.danciu@sciencespo.fr
nicolas.sauger@sciencespo.fr



Follow us on Twitter!

@cdsp_ScPo
@alina_danciu_
@saugerscpo