Implementing a full green open access policy at Inria – the reasons of success

Laurent Romary

Inria, director for scientific information and culture

Overview

- The general context: Scientific information policies and the French open science plans
- The French national repository HAL – covering all higher education and research institutions
- A focus on Inria and its open science policy
- Lessons learnt

Scientific information and open science — why do we need national and institutional policies?

- Supporting researchers in managing and communicating scientific research productions
  - Publications, data, software
  - At the service of science and mankind
- Impact
  - Visibility, citability, measure
- Scientific transparency
  - Proving, comparing, reproducing
- Quality
  - Formats, metadata, authorities
  - Researchers/library specialists
- Budget control
  - Subscriptions, APCs, investments
- Ensuring our (digital) sovereignty
  - Setting up a corpus of our research productions
  - Sustainable public infrastructures, text and data mining, indicators

MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'INNOVATION
Liberté
Égalité
Fraternité

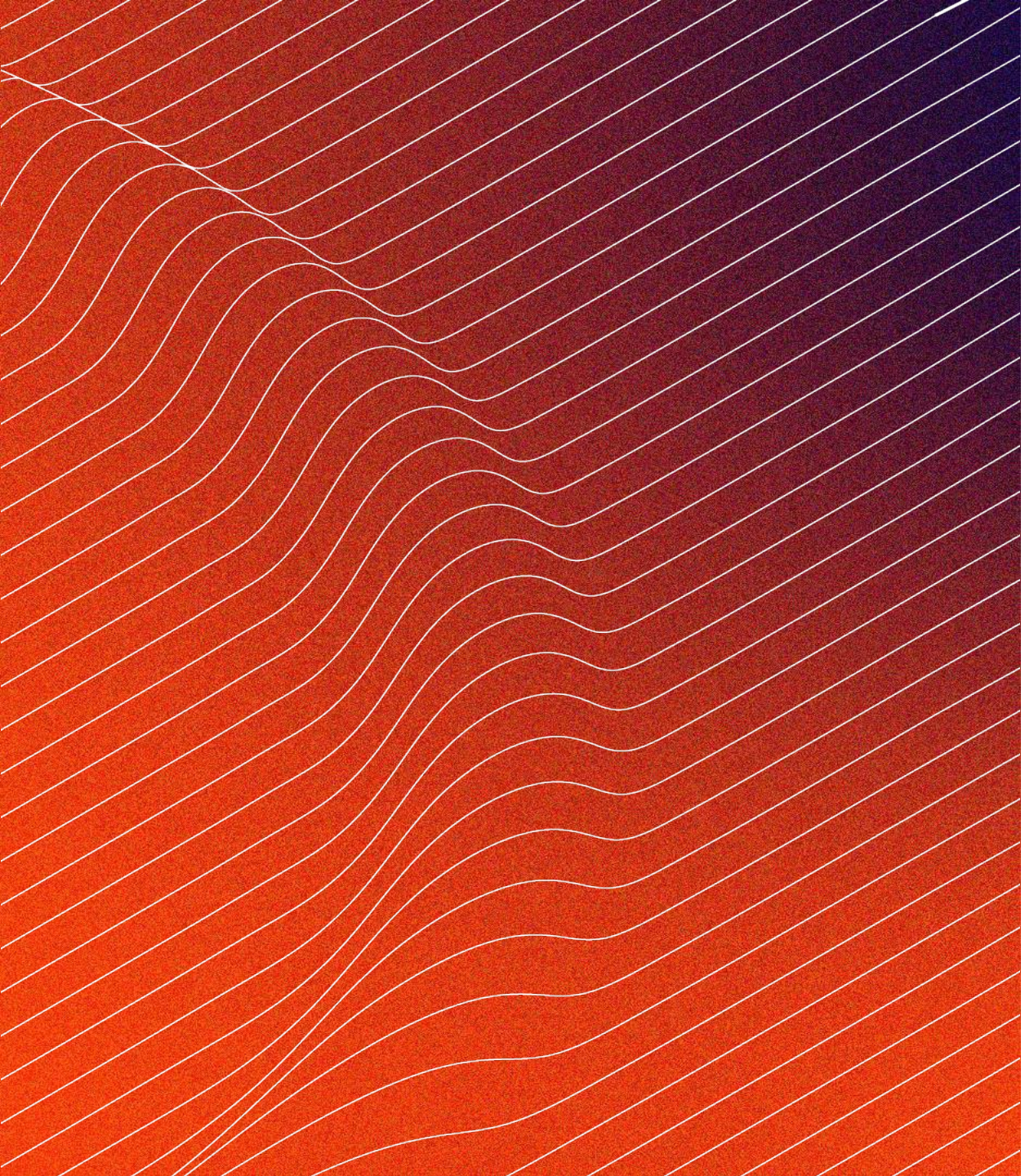
Open science: (French) national context

- **2016:** Loi pour une république numérique (Lemaire)
  - Allows self archiving with maximal embargo periods for accepted authors' manuscripts
  - Open data became mandatory for all publicly funded data, including research data
- **2018**: Premier plan national pour la science ouverte
  - Creation of French Open Science Committee
  - Setting up a National Open Science Fund
- **2021**: Deuxième plan national pour la science ouverte
  - New action plan with 70+ items
  - The pillars of open science: publications, research data, research source codes and software, transforming scholarly practices

https://www.ouvrirlascience.fr/second-national-plan-for-open-science/

Main infrastructures for open science

- **HAL** – the multidisciplinary national publication repository
  - Publications, theses, reports etc. (1 million+ documents, 3 million+ references)
  - Additional services: conference management, overlay journal platform (Episciences)
  - Inria is one of the hosting institutions of HAL (with CNRS and INRAe)
- **Software Heritage** – the international repository that collects, preserves, and shares all software that is publicly available in source code form
  - Harvests all (GitHub, Gitlab etc.) open software forges - 2 billion+ source files
  - Collaboration with Unesco
  - Inria is a co-founder to SH – e.g. linking HAL and Software Heritage
- **Recherche Data Gouv** – the national open research data repository
  - Initiated by the French national committee for open science
  - Dataverse solution - set up by INRAe
  - Officially on air since 8 July 2022



HAL
science ouverte

The multidisciplinary open archive for French research

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
Liberté
Égalité
Fraternité

cnrs

Inria

INRAE

HAL in 5 five key points

1. A public, sustainable and accountable infrastructure
2. A community-driven governance
3. A pooling platform that ensures cost control and innovation
4. An ever-growing multidisciplinary archive with more than 1M full-text documents
5. A range of services for researchers and research performing organisations

A HAL Odyssey

2001-2021

1 331 434 full text documents as of 2024-04-24

HAL's growth during its 20 years of existence was marked by important, major, even decisive events, some of which are pinpointed on the timeline.

More on HAL's history on www.ccsd.cnrs.fr/en/highlights/

- Structuring events
- Major developments
- Political events
- Interoperability
- Milestones
- Portals
- Creation of a new portal

TEL
Creation of an archive dedicated to theses

CREATION OF HAL
by the CCSD

HAL V2
New functionality: portals and collections

Partnership with the international open archive
TRANSFERT TO ARXIV

HAL-SHS
Creation of an archive dedicated to social sciences and humanities

Partnership with the open archive dedicated to economics
TRANSFERT TO REPEC

Partnership with Inserm (French National Institute of Health and Medical Research)
TRANSFER TO PUBMED CENTRAL

100,000
documents in HAL

Coordinated approach at a national scale for the open archiving of scientific publications
INTER-INSTITUTION PROTOCOL

MEDIHAL
Creation of an archive dedicated to image, sound and video

For long-term preservation of files
AGREEMENT WITH CINES

100th
portal

HAL V3
Major evolution with new services: idHAL, AuréHAL, APIs, search engine...

25 partners join to use HAL as a shared platform
INTER-INSTITUTION PARTNERSHIP

EASY DEPOSIT
Evolution of the interface, automatic collect of metadata

The French Ministry of Higher Education, Research and Innovation (MESRI) adds HAL to its roadmap of research infrastructures
HAL: A NATIONAL RESEARCH INFRASTRUCTURE

500,000
documents in HAL

Researchers can import their work from HAL to their ORCID profile
INTERCONNECTION WITH ORCID

The PNSO supports HAL, in the implementation of its 1st axis: increasing open access to publications
1ST NATIONAL PLAN FOR OPEN SCIENCE

Partnership with the universal software archive
TRANSFER TO SOFTWARE HERITAGE

MIGRATION FROM PRODINRA TO HAL-INRAE
More than 62,000 new documents in HAL

GOVERNANCE
The Committee for the Open Science adopts a new funding model for HAL and a new shared governance for the CCSD

EQUIPEX +
HALiance project is a laureate for PIA3

970,000
documents in HAL

2ND NATIONAL PLAN FOR OPEN SCIENCE

The MESRI roadmap of research infrastructure includes an apparatus dedicated to open science and scientific publications, combining HAL, Episciences and Sciencesconf
HAL+ : A NATIONAL RESEARCH INFRASTRUCTURE

2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021

HAL core principles

Sustainability

- Soveraign data hosting, backup and long-term preservation (CC in2p3, CINES)
- Open Source code
- Community driven governance
- Business model (institutional support and projects financing)

Quality

- Documentary, technical and legal check before online publication
- Respect for the authorship of the document
- Alignement with reference data and use of identifiers (ORCID, idHAL, ROR, project codes, etc.)

Interoperability

- Transfer of deposits to ArXiv and PubMedCentral
- Provision of files in the REPEC database.
- Completion of the ORCID account
- Display of publishers' policies regarding self-archiving (Sherpa-Roméo)

Visibility

- OAI-PMH repository and API
- Exposure of data in RDF (triplestore DataHAL)
- Search engines (Google Scholar, Google, PubMed via Linkout, Isidore)
- International directories (OpenDoar)

Benefits of the HAL (centralized) model

Cost savings

Using HAL free up universities resources so that they can focus on their core functions

---

CCSD

- *Daily management of the front and back office*
- *Design and IT developments*
- *SEO*
- *Authentication and access control*
- *Modernizing of legacy*
- *Input quality control*
- *Provision of user help*

Regulatory compliance

Simplification of compliance with numerous and evolving standards and protocols

---

CCSD

- *Alignment with authority files and adding of PIDs (ORCID, ISNI, ROR, etc.)*
- *Compliance with funders (ANR, OpenAire)*
- *Certification (CTS in process)*
- *Open APIs, RDF, OAI-PMH, Sword, Notify*

Innovation

Network and collaborative ecosystem accelerate the adoption of next-generation technologies and the adding of new services

---

CCSD

- *Overlay services (Episciences)*
- *Interoperability with international OR (PMC, arXiv, RePeC, SWH)*
- *Research CV*
- *Statistics and Altmetrics*
- *Collections of publications*
- *Automated completion of the bibliographic metadata*

Resiliency

Use and implementation within HAL of standardized processes and good practices

---

CCSD

- *Reference data*
- *Long term preservation of the publications (OAIS)*
- *Redundant systems*
- *Stability of the URLs*
- *Citability*

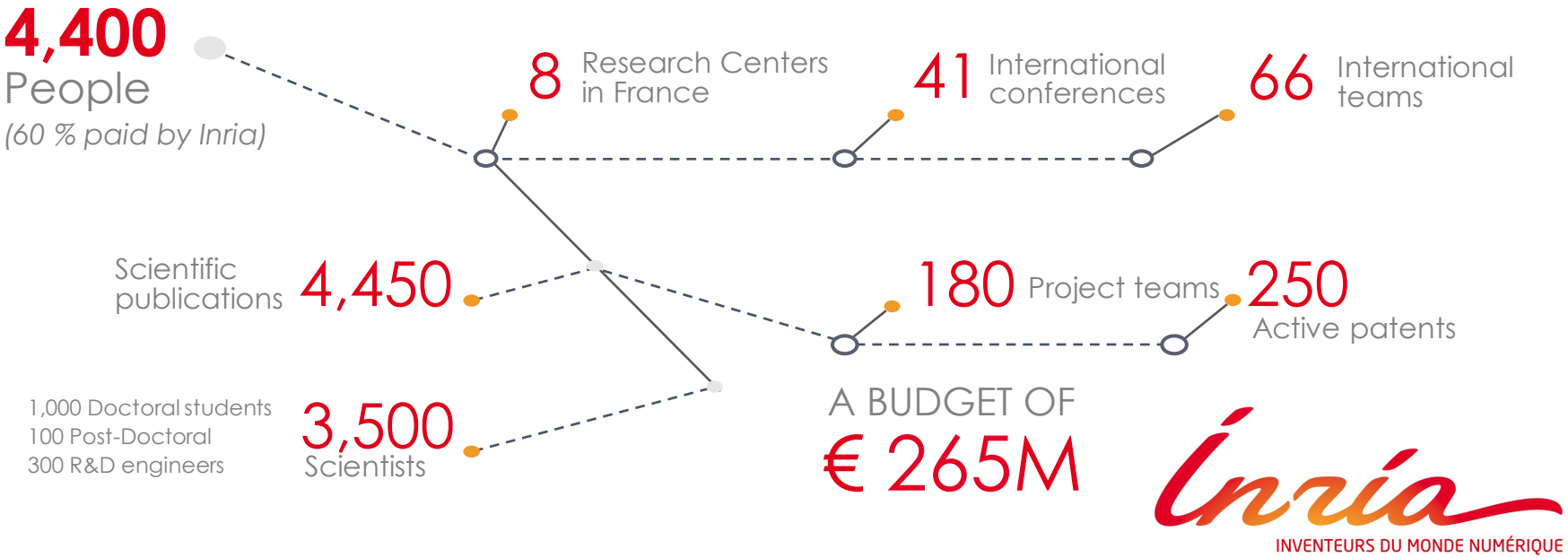
1 national open archive, 141 institutional portals

92%

of French universities use HAL

Open Science at Inria – a comprehensive policy

The underlying vision



Objectives of Inria's scientific information policy

- Maximising the **dissemination of our scientific assets** (visibility and swift communication of knowledge), for a reasonable price
- Constitution of a **reliable** and **sovereign institutional corpus** (documentation, preservation, access), with clear public governance principles
- Contribution to **shaping the scientific communication landscape** in terms of editorial processes and usage made of scientific productions

Inria scientific information policy in concrete terms

- Deposit mandate on all scientific publications (in HAL, CC-BY)
  - **Condition to appear in annual research reports**
  - Encouragement towards preprints
- Central budget for APCs (aka SIS - Supervised injection sites)
  - Forbidding hybrid open access
  - **Publisher's PDF to be deposited in HAL (with equal licence)**
  - Management of a national dashboard of costs and journals (OpenAPC)
  - No refunding requests to funders (FR, EU)
- Engaging in developing new publication models and infrastructure
  - **Editorial support to *Episciences* based journals**
  - Investment and support to *Software Heritage*
  - **Editorialization (IFIP DL, conferences)**
- Research data support group
  - Data management plan, support to the use of the new Recherche Data Gouv infrastructure
  - National network of scientific contacts/champions for research data
- Printed material as disposable goods
  - Creation of a central collection of reference works
  - **On going digitization of legacy publications (on HAL)**
- Engaging in text and data mining
  - Contribution to the national open science monitor
    - Software and data set citations

See an overview under: https://www.inria.fr/en/open-science-inria-role

Welcome to HAL-Inria

Hal-Inria allows to access all publications in HAL and, for Inria team members, to deposit the full text of their scientific contributions

Zoom on...

Episciences survey: users highlight advantages

épisciences.org

A survey was conducted in the form of a questionnaire and distributed to the editors of the 11 journals active on Episciences in April 2018.

The survey highlighted the strengths of Episciences:

- an economic model considered virtuous by the journals on Episciences;
- publication of references of articles, volumes and sections on a personalised website;
- independent evaluation by members of editorial boards outside traditional private editorial circuits;
- technical support and day-to-day assistance in using the platform.

Read the results of the survey

Reminder: The Inria Scientific Edition team accompanies you in your epijournal project in computer science and applied mathematics.

LATEST DEPOSITS

Nour El Houda Ghanmi. Etude DFT de l'influence des ions fluorure sur les mécanismes de croissance des oxydes nanoporeux par oxydation anodique. Chimie-Physique [physics.chem-ph]. Université Pierre et Marie Curie - Paris VI, 2017. Français. ⟨NNT : 2017PA066354⟩. ⟨tel-01879538⟩

Mahmoud Teimouri. D'une pensée moderniste à une approche paysagère : étude du rôle de l'identité dans les approches de rénovation urbaine à Téhéran contemporain (depuis la fondation de l' organisation de la rénovation urbaine de Téhéran en 1975). Art et histoire de l'art. Université Panthéon-Sorbonne - Paris I, 2015. Français. ⟨NNT : 2015PA010550⟩. ⟨tel-01879537⟩

Enrica Montalban. Long-lasting effects of operant conditioning and cocaine on D1 pyramidal neurons in prefrontal cortex and on the D1 and D2 striatal neurons mRNAs. Neurons and Cognition [q-bio.NC]. Université Pierre et Marie Curie - Paris VI, 2016. English. ⟨NNT : 2016PA066723⟩. ⟨tel-01879536⟩

SEARCH

Termes de recherche

LINKS

- Documentation
- Helpdesk Inria
- contact

NEWS

SAVE THE DATE "INTEROPÉRABILITÉ ET PÉRENNISATION DES DONNÉES DE LA RECHERCHE: COMMENT FAIR EN PRATIQUE ? RETOURS D'EXPÉRIENCES" *(10/9/18)*
Paris, November 27th, 2018

NEW UPDATE OF HAL RELEASE NOTES (IN FRENCH) *(6/9/18)*

AN EXTREMUM STRONG NEWS FOR OPEN SCIENCE (IN FRENCH) *(7/13/18)*

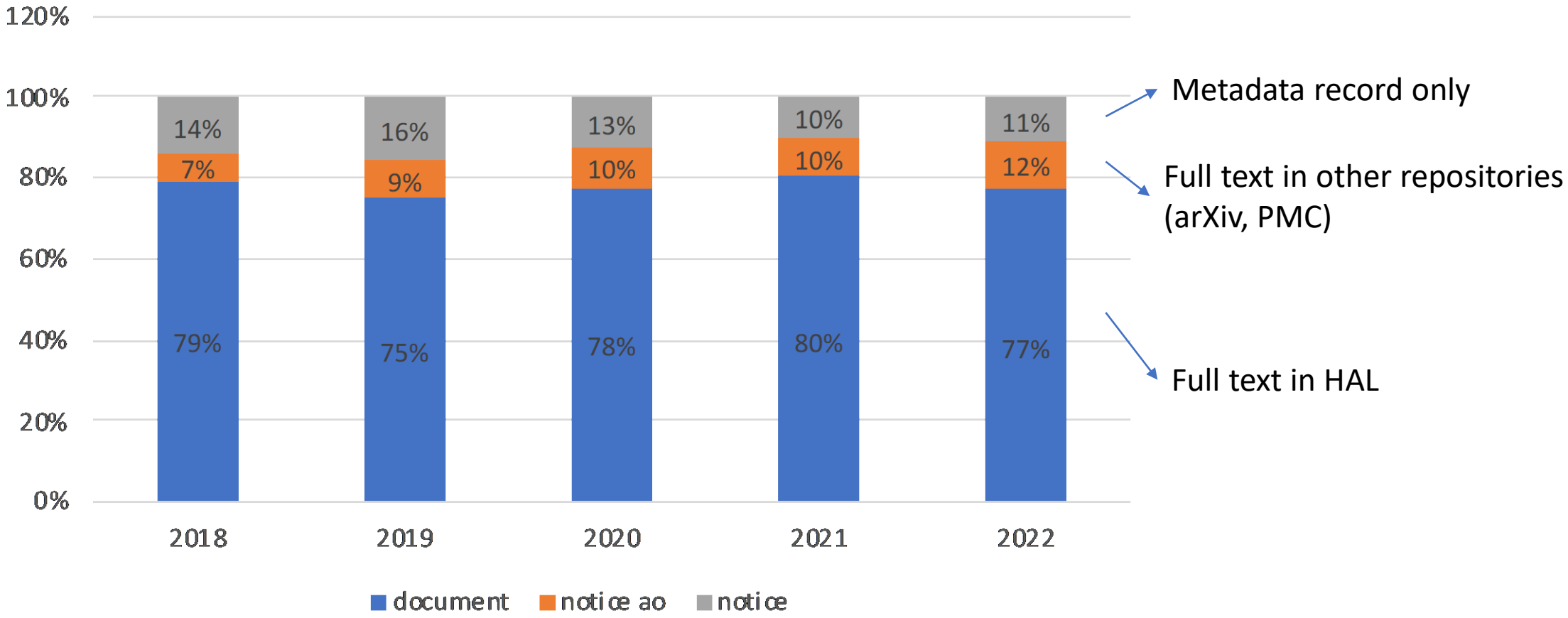
SEMINAR PDS/EHESS INTERVENTIONS ARE NOW AVAILABLE IN AUDIO *(6/22/18)*

WHY COLLECT, PRESERVE AND SHARE TEH SOURCE CODE OF ALL THE SOFTWARE THAT IS AVAILABLE? *(6/19/18)*

NUMBER OF FULL TEXT DOCUMENTS

581 298

A successful policy ☺



Full text availability rate for Inria publications in HAL for all publication types

Lessons learnt - 1

- Patience…
  - From encouraging deposit (from 2005 onwards) to deposit mandate (since 2013)
  - Setting up library teams
  - Working hand in hand with researchers
  - Providing services
    - E.g. one deposit – multiple usages (reporting, web page, bibliographies/exports)
- Main difficulty : having a coherent policy
  - « What goes in the annual report has to be in HAL »
  - Being flexible with borderline cases (book chapters, books, multiple non-Inria authors)
- Importance of an underlying legal framework
  - French right of secondary publication (*Loi pour une république numérique,* 2016)

Lessons learnt - 2

- Importance of a coherent scientific information policy endorsed at all levels within an institution
  - Management: understanding the added value of open science (reliability, visibility, efficiency)
  - Researchers: must be directly involved in the implementation of the policy and benefit from services
  - Library staff: dedicated to providing support, ensuring quality and coverage
- Mastering all aspects of scientific information
  - Publications, data and software, and related authorities (authors, affiliations)
  - Peer review, reporting and assessment
  - Research cartography, text and data mining, machine learning
  - And costs!

Next steps for France

- Getting more institutions in the deposit mandate loop
  - The relation to reporting is key
- Systematizing deposit for all open access content
  - Open works detector: e.g. articles in open access journals
- Providing a dashboard of the progress made
  - Open science monitor – publications, data sets and software
- Working on diamond open access platforms
  - E.g. Episciences: overlay journals on publication repositories

Merci de votre attention !