

DECENTRALIZED RESEARCH DATA MANAGEMENT INTRODUCING SOVISU+

David REYMOND

Université de Toulon - Open Science Officer
Accredited Research Director and Assistant Professor
Library and Information Science

Without CRIS, a struggle

- Recurrent need to identify the list of publications
- Gathering publication databases (and/or publications)
 - Identify laboratory affiliations on publications (difficult)
- A real painstaking task, based on cleaning a list of publications that the institution purchased and whose content was not the institution's researchers!

The worse point:

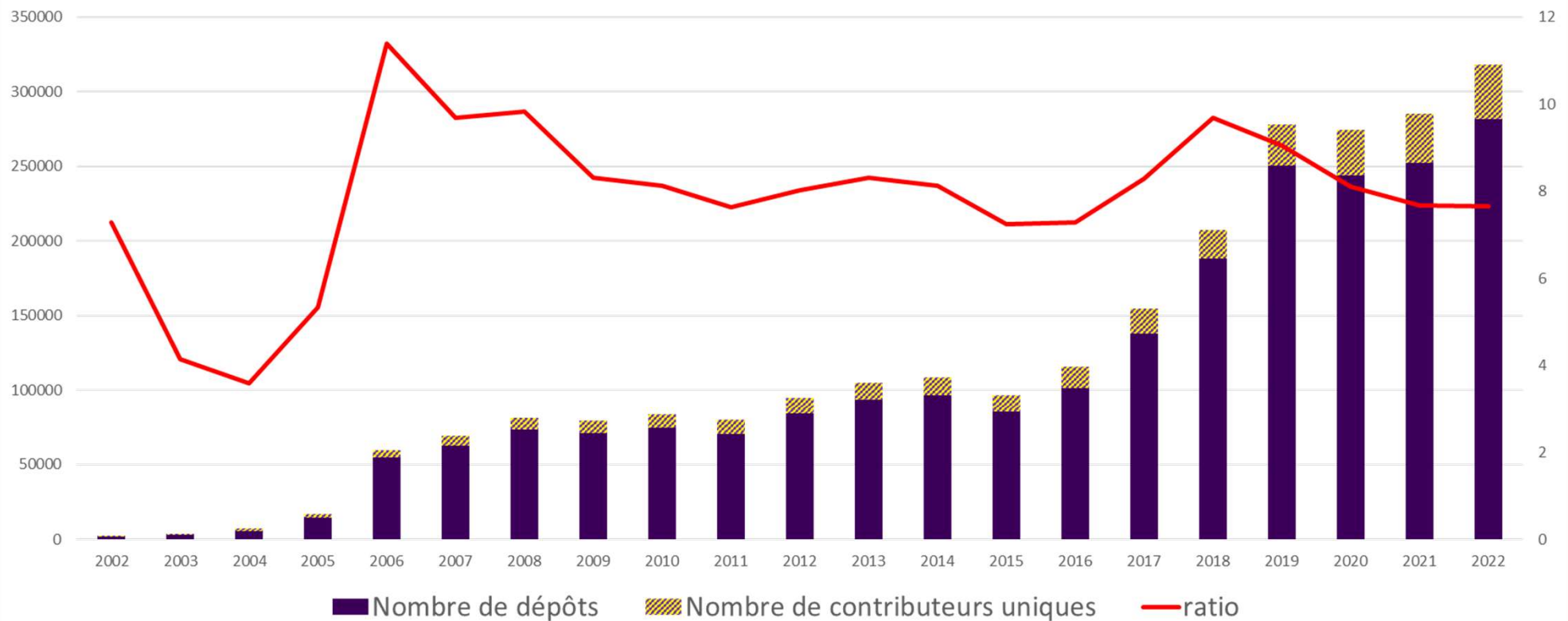
- Publications are attributed to laboratories and individuals
- No one can attest the reliability
- Who, other than the authors, could ?

CNRS (CENTRE NATIONAL DE LA RECH SCIENTIFIQUE)
CENTRE NAT DE LA RECH SCIENTIFIQUE(CNRS)
CT NAT DE LA RECH SCIENTIFIQUE
CNRS CENTRE NATIONAL DE LA RECH SCIENTIFIQUE
CENTRE NAT DE LA RECH SCIENTIFIQUE CNRS
CNRS DIRE
CT NAT DE LE RECH SCIENT
SCIENT C N R S CT NAT DE LA RECH
CT NAT DE LA RECH CNRS
LE CENTRE NAT DE LA RECH SCIENT
CT NAT DE LA RECH SCIENTIFIQUE—CNRS—
SANTR NASONAL DE LJA RESHERSH SJANTIFIK
CT NAT DE LA RECH SCIENTIFIQUE—CNRS
CENTER NAT DE LA RECH SCIENT C N R S
CT NAT DE LA RECH SCIENT PUBLI
PHARMAMENS E CT NAT DE LA RECH SCIENT CNRS
ENTRE NAT RECH SCIENT
CT NAT DE LA RESERCHE SCIENT CNRS
CENTRE NATIONAL DE LA RECH SCIENTIFIQUE—CNRS
CENTRE NAT DE LA RECHERCH SCIENT
EHNSTITJU NAS ONAL DE LJA RESH
CENTRE NATIONAL DE LA RECH SCIENTIFIQUE—CNRS

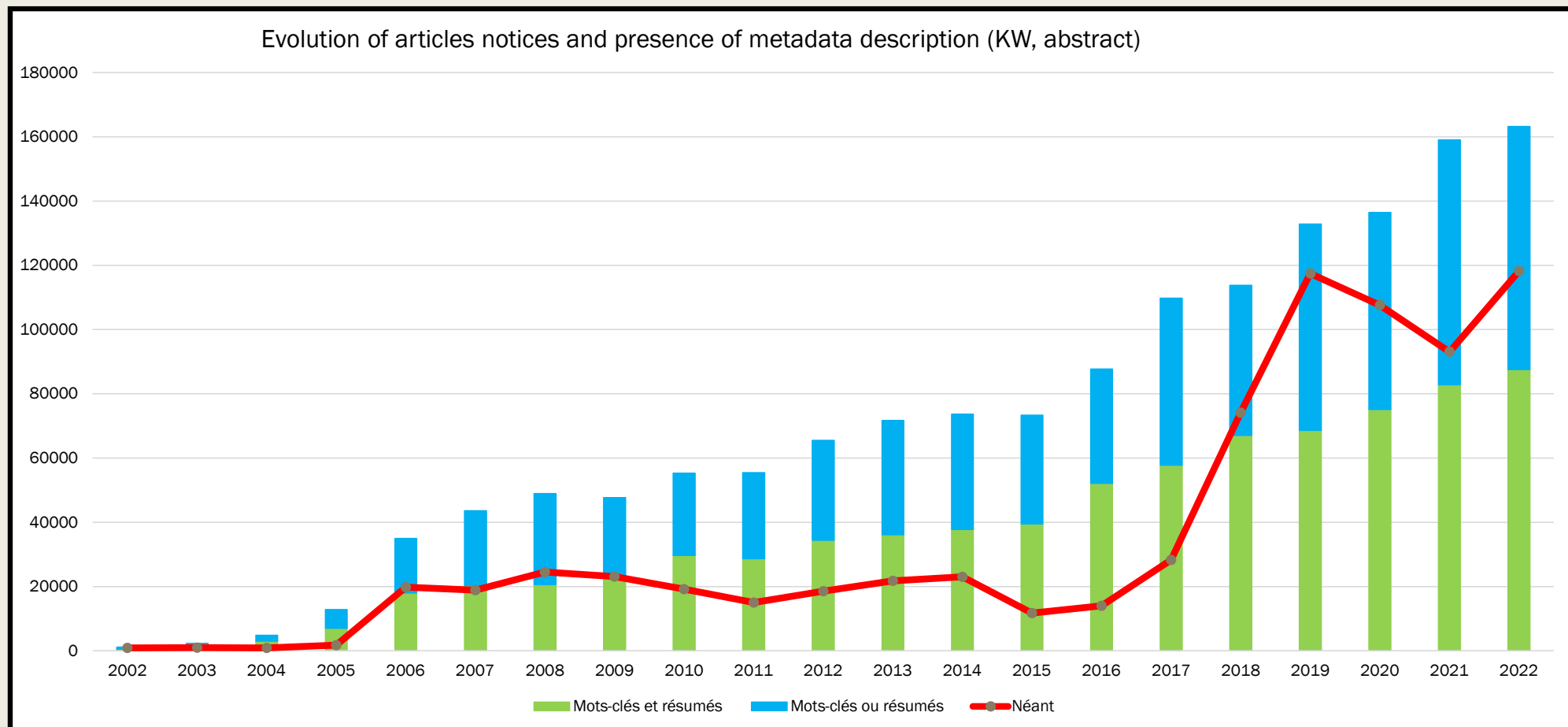
Perverse leverage effect



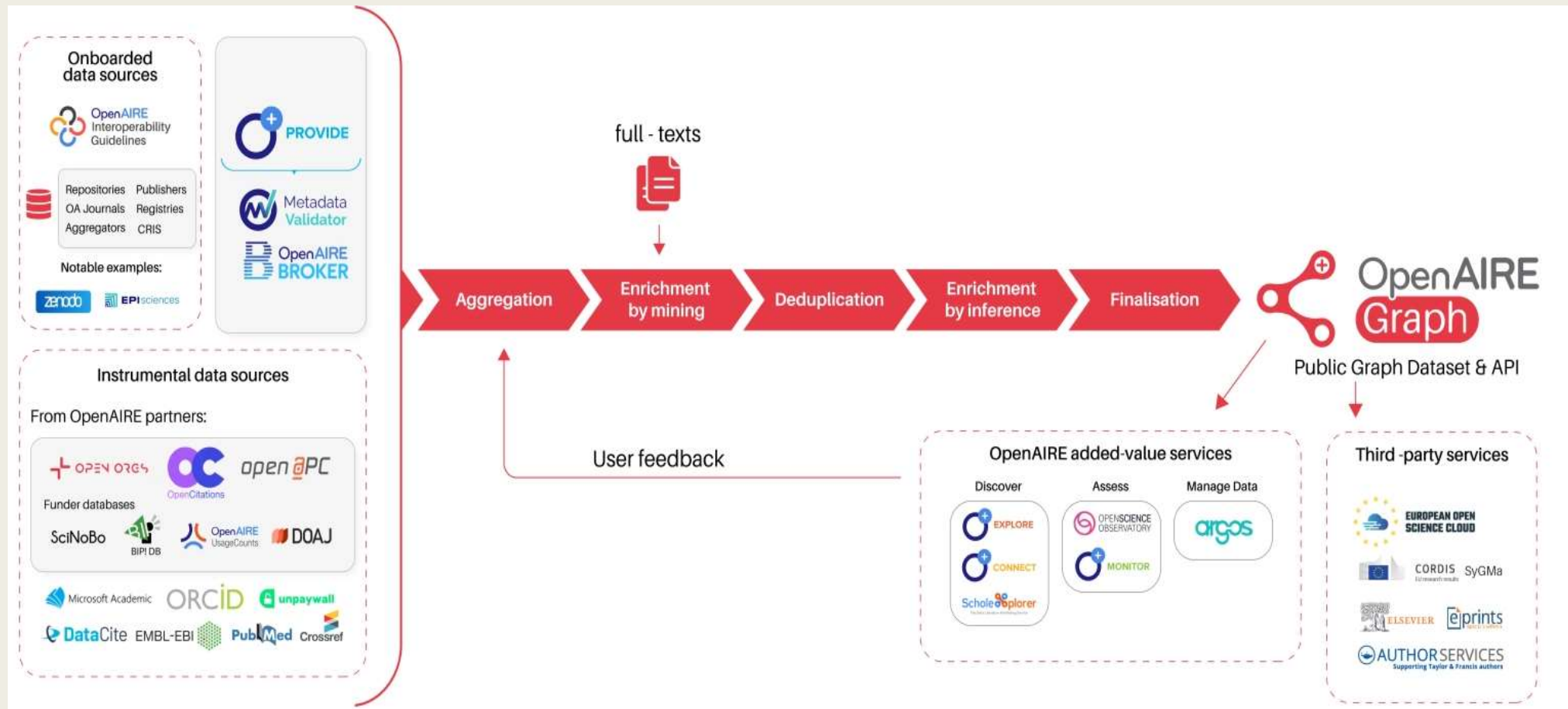
Evolution du nombre de dépôts, de contributeurs et le ratio (dépôt/contributeur)



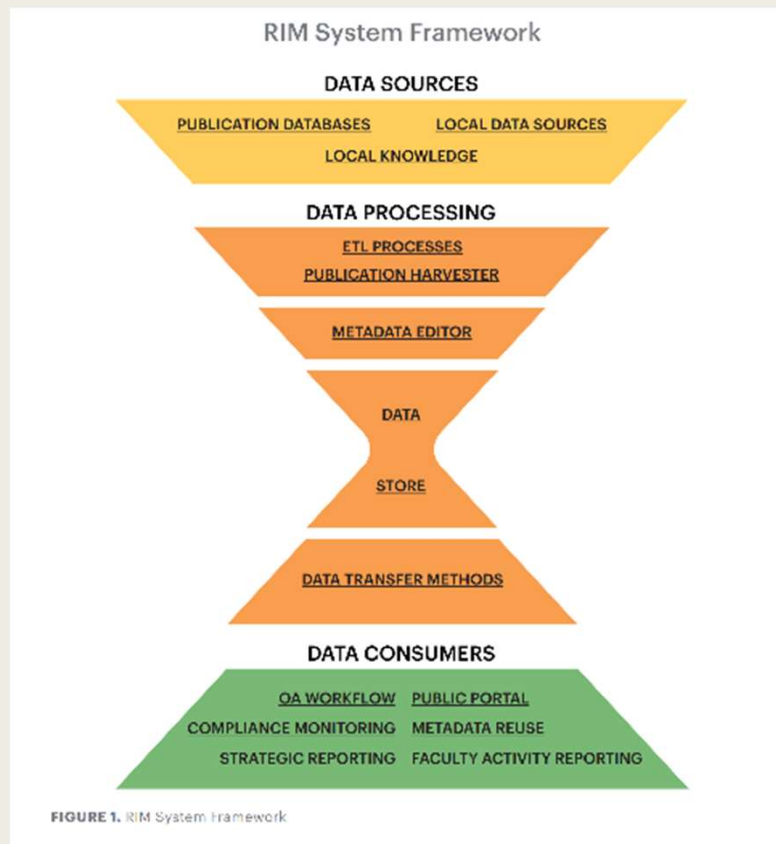
OS policy side bad effect, not really FAIR



TDM → Garbage In, Garbage Out



A « French » challenge?

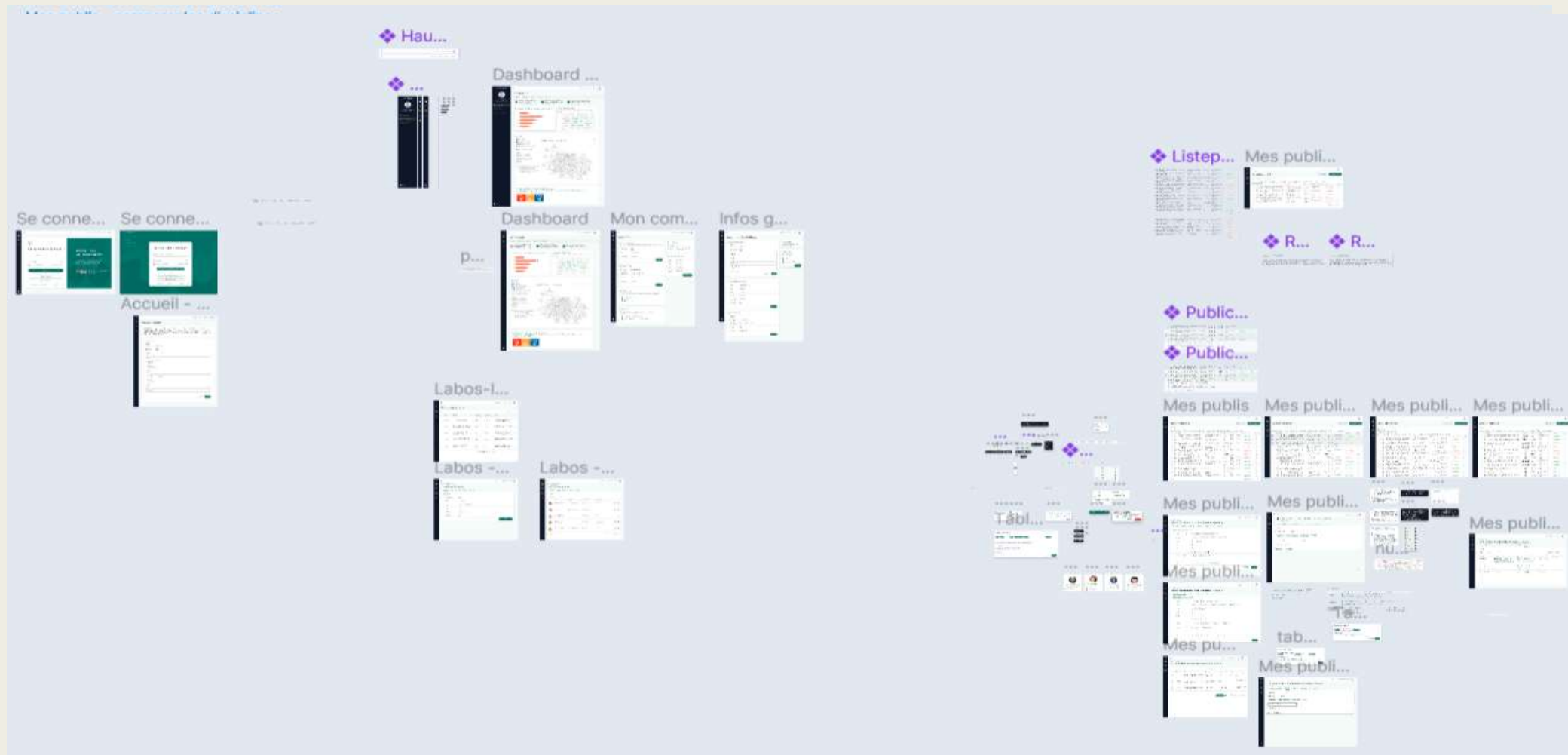


- Respect the layered structure of higher education and research
 - Keep the *members* list up to date
 - Keep the *production* list up to date
 - Index expertise's, experts and production
 - Data quality requirement ([six key principles for Scientific Data](#))
 - *Involve the members in their own data verification/curation*
 - *Verify data authenticity and integrity of certain information*
 - Simplify: Say it once
 - One step beyond: use Linked Data transparently
- Map out (cartography, KG)

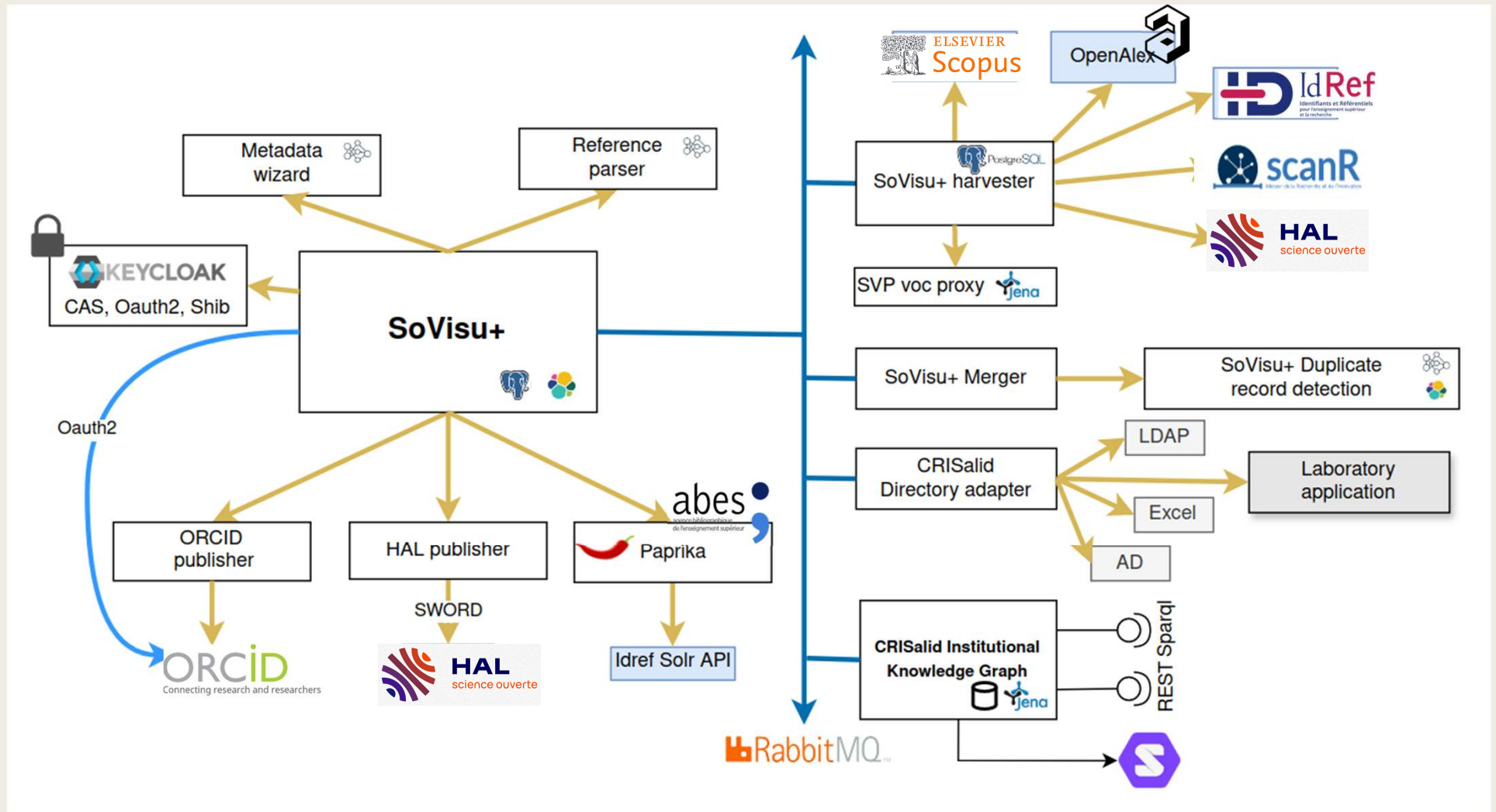
Philosophy

- Shortcomings of current centralized systems
 - *Built in silo's architecture, RIMs may handle **errors**, **unverified information** and **reliability***
 - *By their side, for **researchers** it's **mandatory to manage** their own production. Dissemination in indexes is dependent of metadata quality*
 - *So, the job is **done at least twice but separatively***
- Introduction to SoVisu+ and its decentralized model
 - *Provide researcher a tool for managing their production (as a first step)*
 - *Help them in checkiing their (meta)information (Pids, affiliation, research outputs)*
- We challenge this way centralized sytem and avoid two main issues
 - *Fragmentation and data quality issues*
 - *Problems with silo's interoperability and data obsolescence*

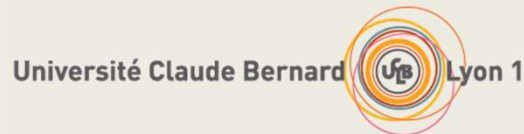
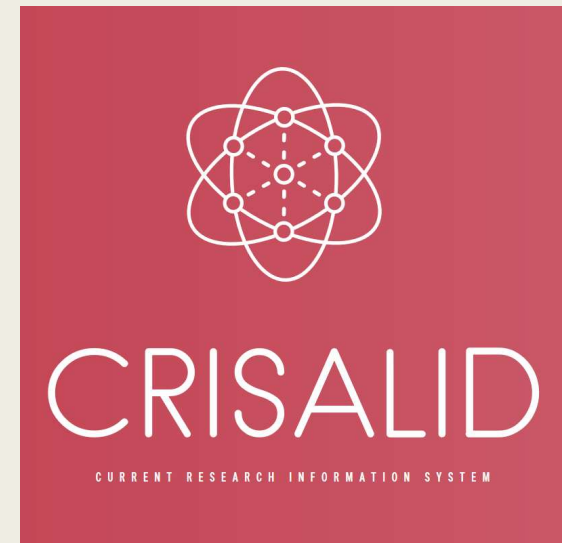
SoVisu+ overview



Micro services architecture (SoVisu+ 05/2024)



A great support community : CRISalid



Others ?

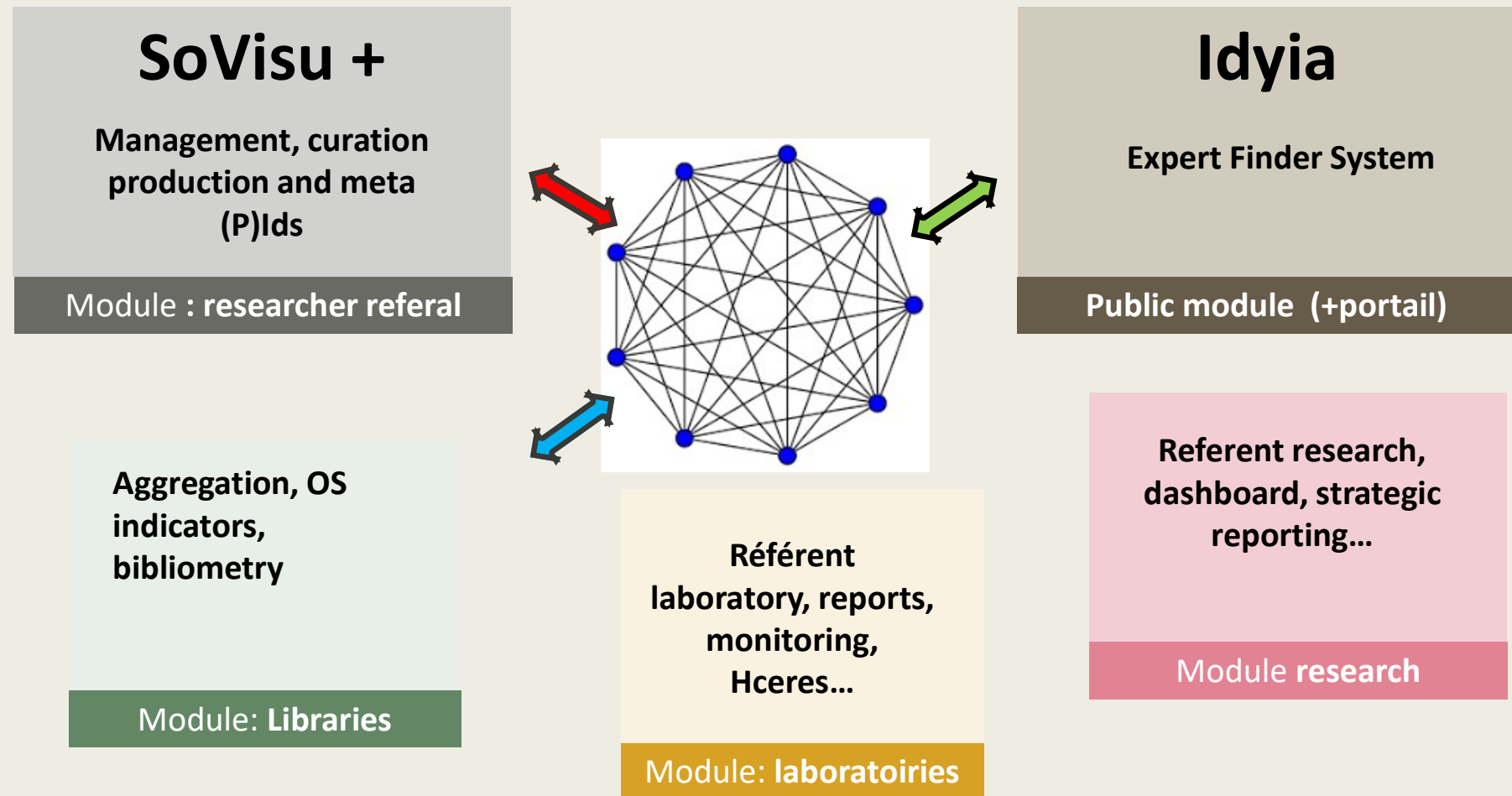




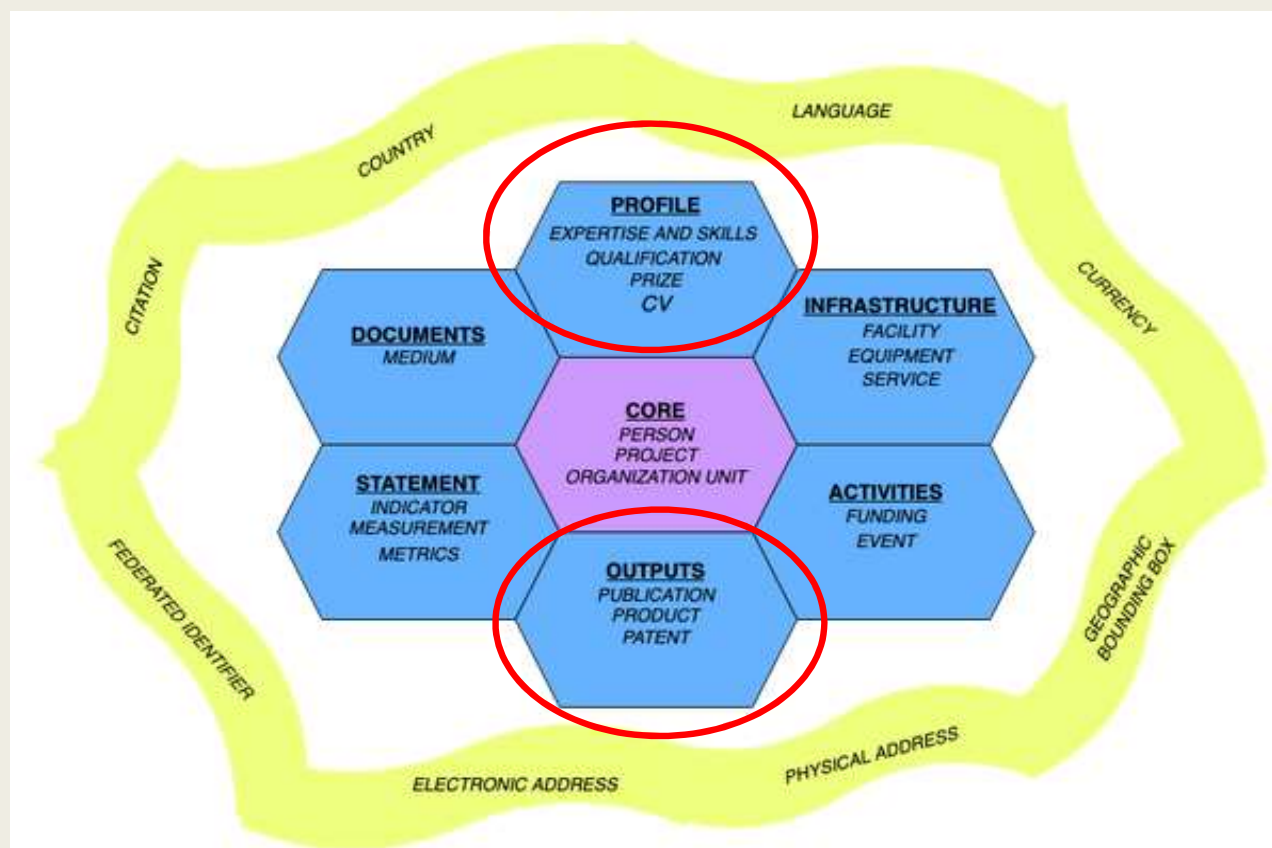
CRISolid's projects

- CRISolid is a community of researchers, documentalists, and computer engineers.
- Built Institutional Knowledge Graph
- Develops Tools and services for Open Science,
- Maintenance, support...
- In five years long cycle

Modularity and functions



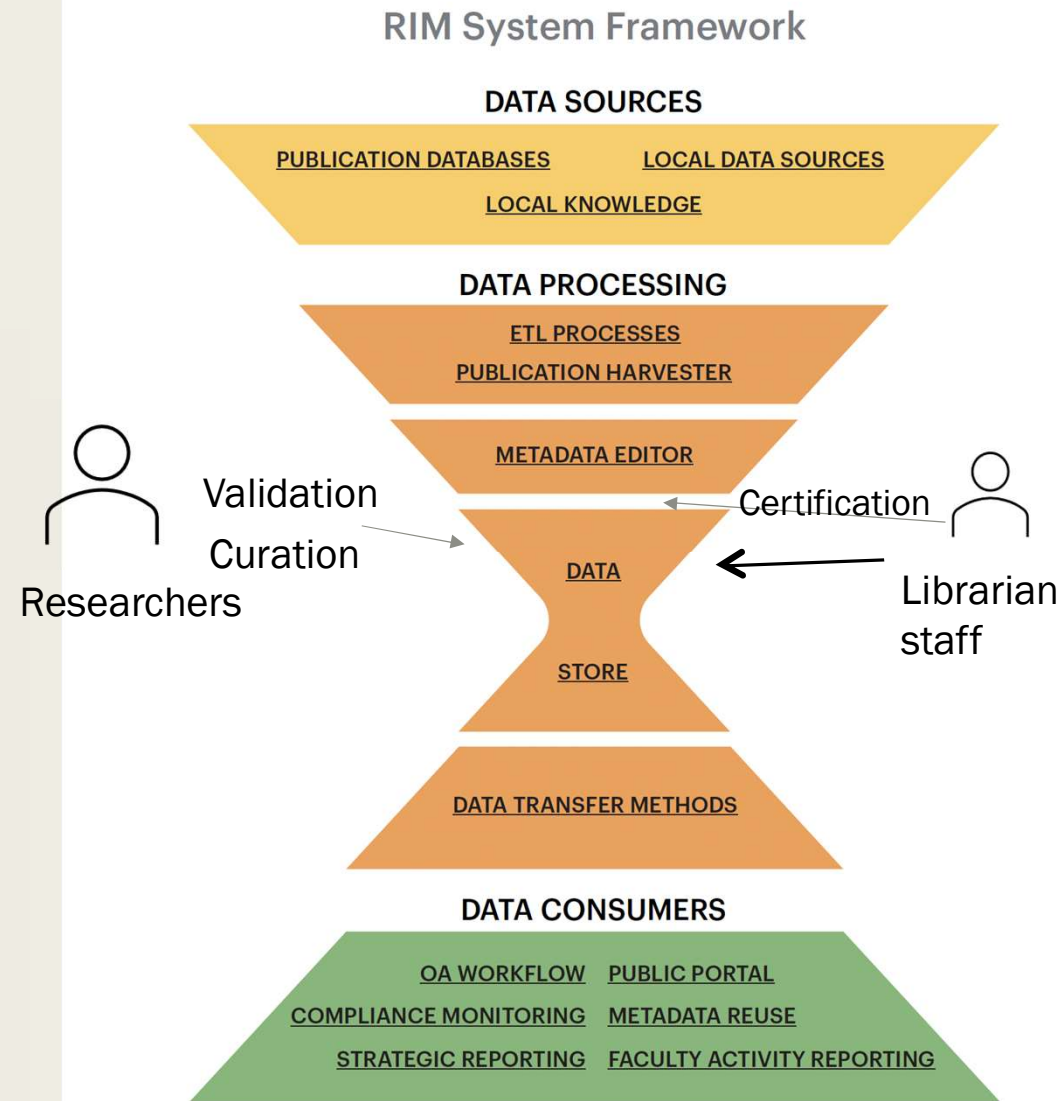
Strategy in the scope of the Common European Research Information Format (CERIF)



Take a broader view

- « Research information management (RIM) systems support the transparent aggregation, curation, and utilization of data about institutional research activities. »

Bryant, Rebecca, Jan Fransen, Pablo de Castro, Brenna Helmstutler, and David Scherer. 2021. *Research Information Management in the United States: Part 1—Findings and Recommendations*. Dublin, OH: OCLC Research. <https://doi.org/10.25333/8hgy-s428>.



Conclusion

- Strategy for building and maintaining a **collaborative robust** knowledge graph
- Importance of researcher-driven updates and certifications
- IA and decentralized Web technologies to improve acceptability and empower users
- SoVisu+'s approach to a more comprehensive and transparent evaluation support as everyone is involved in the process
- Challenges: ensuring international interoperability

Q&A

- Thanks

David REYMOND – david.reymond@univ-tln.fr

References

- **IAP input into the UNESCO Open Science Recommendation** July 2020, The InterAcademy Partnership (IAP), Trieste. <https://www.interacademies.org/publication/iap-input-unesco-open-science-recommendation>
 - Tabariès, A., & Reymond, D. (2021, January 27). Comparaison de méthodes d'extraction de mots-clés non supervisées. *Atelier TextMine - 21ème édition conférence Extraction et Gestion des Connaissances (EGC)*, Montpellier, France.
 - Reymond, D. 2021b. « SoVisu, terrains de recherches SIC en cours ». présenté à Réunion IMSIC AXE 1 Enjeux et usages des dispositifs sociotechniques numériques et des mutations informationnelles Captas, documentation et numérique, juin 18.
 - —. 2022a. « SoViSu : visibilité et lisibilité en SO ». présenté à Séminaire IMSIC AXE 1 *Brevets, captas et Science Ouverte*, février 24.
 - —. 2022b. « Involving researchers in Open Science: the SoVisu innovative solution ». Dans *EOSC Symposium, session MVE beyond EOSC Future and the INFRAEOSC07*. Prague, Tchéquie.
 - —. 2023. « Les chercheurs dans la boucle de la Science Ouverte. Et l'IA ? » présenté à Séminaire ObTIC, Sorbonne Université, Paris, France, mai 11.
 - —.(2023, Juin) Le dispositif “SoVisu”. Séance publique de l'atelier “cartographie de la recherche”, Université Paris 1 Panthéon-Sorbonne, Paris, France. ([hal-04141467](https://hal.archives-ouvertes.fr/hal-04141467))
 - —., Lapôtre, R. « SoVisu+: starting point and foundations of a national CRIS » présenté à la conférence **Semantic Web in Libraries (SWIB)** 23, Berlin, Allemagne, sept 12.
 - —. , Buccheri, A., Dornbush, J., Dubernard, H., Lapôtre, R., (2023, sep) SoVisu+ (SVP). 4^e séance publique de l'atelier “cartographie de la recherche”, Université Paris 1 Panthéon-Sorbonne, Paris, France
1. Bryant, Rebecca, Janet Fransen, Pablo de Castro, Brenna Helmstutler, et David Scherer. « Research Information Management in the United States: Part 2—Case Studies Research Information Management in the United States: Part 1—Findings and Recommendations », 2021. <https://doi.org/10.25333/8HGY-S428>.
 2. Langer, André, Dang Vu Nguyen Hai, et Martin Gaedke. « SolidRDP: Applying Solid Data Containers for Research Data Publishing ». In *Web Engineering*, édité par Maria Bielikova, Tommi Mikkonen, et Cesare Pautasso, 12128:399-415. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-50578-3_27.
 3. Samba, Andrei Vlad, Essam Mansour, Sandro Hawke, Maged Zereba, Nicola Greco, Abdurrahman Ghanem, Dmitri Zagidulin, Ashraf AboulNaga, et Tim Berners-Lee. « Solid: a platform for decentralized social applications based on linked data ». *MIT CSAIL & Qatar Computing Research Institute, Tech. Rep.*, 2016.