International FAIR Convergence Symposium
27 November 2020 - 04 December 2020
Remotely

Issues and obstacle to the sharing of sensitive data in Europe
Who are the stakeholders and what are their responsibilities?

Authors: **Romain David**, Audrey Richard, Diana Stepanyan, Krista Versteeg, Jonathan Mineau-Cesari, Michel-Yves Mistou, Laurence Mabile, Hervé Raoul.

Reducing the risk of misuse of data is particularly crucial in the context of research on highly pathogenic agents. But as demonstrated during the COVID19 crisis, **sharing data of high quality is a sine qua non condition to compare research results at a large scale**. Our actual challenge is to **ensure compliance with the FAIR principles while taking into account the dangers and pitfalls concerning dual use type data** / samples (because there are often some good reasons to not share).

Identifying stakeholders and organising their interaction is strategic to ensure that data will be **accessible easily** while observing in a highly secure environment the essential rules with regards to the rights of access and reuse of sensitive data (*as open as possible, as closed as necessary*). To that aim, preparing all stakeholders and improving their FAIRness literacy must be prioritized by all institutions (David *et al.*, 2020).

Dual Use Research Concern (DURC)

Dual Use Research of Concern (DURC) is defined by the United States Government Policy for Oversight of Life Sciences as "*Life sciences research that, based on current understanding, can be reasonably anticipated to provide knowledge, information, products, or technologies that could be directly misapplied to pose a significant threat, with broad potential consequences, to public health and safety, agricultural crops and other plants, animals, the environment, materiel, or national security.*"

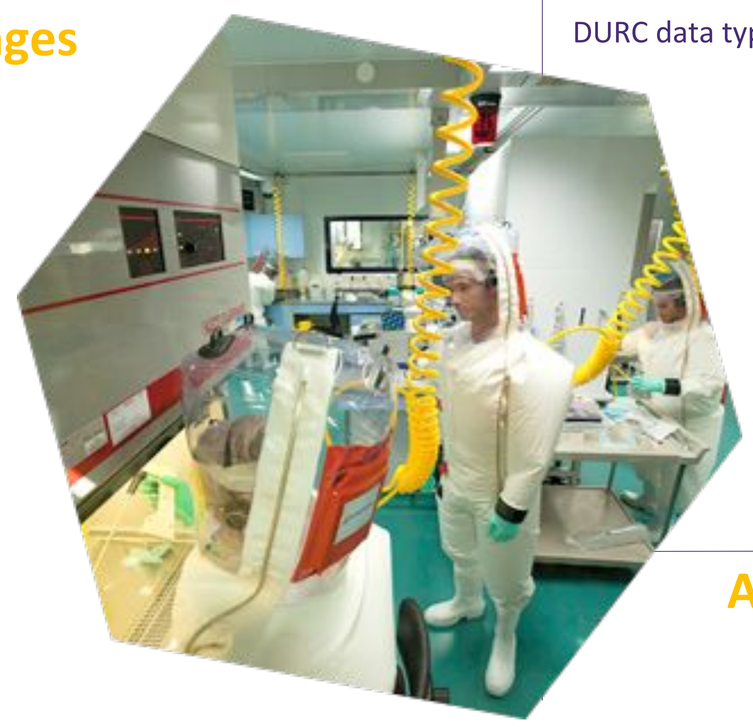
Many countries or regions have adopted laws to regulate exchanges concerning DURC. As an example, in the European Union, dual use biological materials are listed in "Council Regulation (EC) No 428/2009".

DURC data typically falls into one or more of these categories:

- Demonstrates how to render a vaccine ineffective
- Enhances the harmful consequences of a pathogen or toxin or renders a non-pathogen virulent
- Increases the transmissibility of a pathogen
- Alters the host range of a pathogen or toxin
- Enables evasion of diagnostic or detection modalities
- Enhances the susceptibility of a host population to a pathogen or toxin
- Generates or reconstitutes certain eradicated or extinct pathogens or toxins
- Enables weaponization of a biological agent or toxin.

Dual Use Research Concern (DURC) management challenges

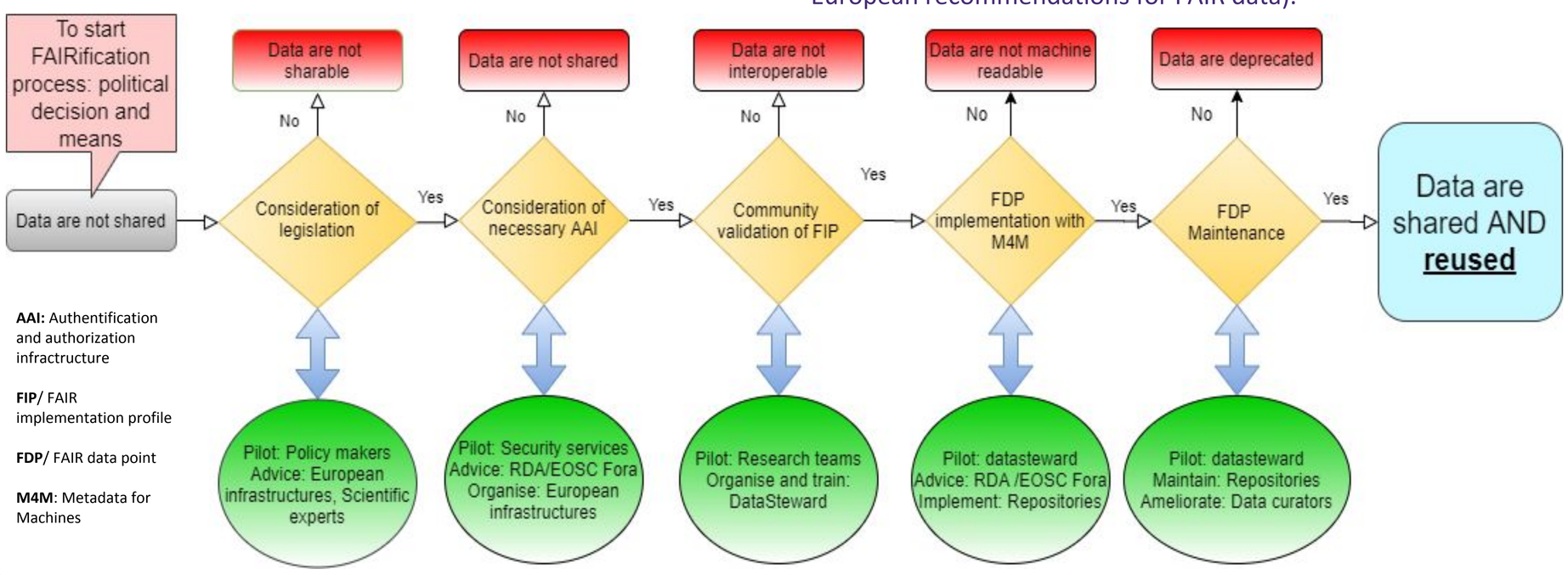
- How can we take these aspects into account so that in the meantime **FAIR principles and legislations are applied** ?
- How to **increase the possibilities of sharing dual use data** to reduce overlap of services? and to encourage collaboration on DURC while securing their uses?
- How to **reward and credit back good practices**?
- *What data should be prohibited to share*?
- *Who should be prohibited* from sharing or accessing data which may be classified as DURC?



AAI services, FAIR Data Points, and Fair Implementation Profile for DURC

Dual use metadata must be machine readable (like Metadata for machine - M4M implemented in Go-FAIR) to be reused at large scale. They can contain sensitive content, therefore an authentication and authorization mechanism must be implemented by an authorized infrastructure (AAIs are part of the 6 European recommendations for FAIR data).

Stakeholders and their roles for each critical step in data FAIRification.



AAI: Authentification and authorization infractructure

FIP/ FAIR implementation profile

FDP/ FAIR data point

M4M: Metadata for Machines

Access to the various FAIR data points -FDP (landing page containing the data while respecting all the FAIR criteria) to access the data is done through the AAI services. To describe FAIR ressources usable to build FDP for sensitive data, we have replicated a Fair Implementation Profile (FIP) developed by the VoDAN team (as we ve dealing with the same scientific objects) Pergl Sustkova et al., 2019.

References :

- Six Recommendations for implementation of FAIR practice by the FAIR in practice task force of the European open science cloud FAIR working group. https://op.europa.eu/en/publication-detail/-/publication/4630fa57-1348-11eb-9a54-01aa75ed71a1
- David, R., Mabile, L., Specht, A., Stryeck, S., Thomsen, M., Yahia, M., Jonquet, C., Dollé, L., Jacob, D., Bailo, D., Bravo, E., Gachet, S., Gunderman, H., Hollebecq, J.-E., Ioannidis, V., Le Bras, Y., Lerigoleur, E., Cambon-Thomsen, A. and Research Data Alliance – SHAring Reward and Credit (SHARC) Interest Group, T.R.D., 2020. FAIRness Literacy: The Achilles' Heel of Applying FAIR Principles. Data Science Journal, 19(1), p.32. DOI: http://doi.org/10.5334/dsj-2020-032
- http://www.data-intelligence-journal.org/p/47/ Schultes FIP https://www.go-fair.org/how-to-go-fair/fair-implementation-profile/
- FDP. https://www.go-fair.org/how-to-go-fair/fair-data-point/
- M4M. https://www.go-fair.org/how-to-go-fair/metadata-for-machines/

Take home messages

"Chefs d'orchestre" are needed. European research infrastructure networks are the best structured to do that.

We recommend the following:

- **setting up pre-approved data sharing agreements** for highly secured repositories.
- considering DURC in actual Data stewardship should be developed.
- to be cautious while **being too careful and classifying all data in dual uses; it would be another pitfall**, and would prevent re-use despite public health issues.
- **to identify actors** at least inside a country or inside a community of countries (e.g. Europe) **and organize their interactions** as reducing duplication of effort and improving trial design concerning Dual Use research is a challenge,
- **to enable dual use data sharing information, an authorized metadata should be available centrally on an intergovernmental web page** with explicit authority contact and / or be applied considering officials sharing conventions.

ERINHA Advance has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824061

Author affiliations : 1) ERINHA (European Research Infrastructure on Highly Pathogenic Agents) AISBL, FR, 2) MIRRI (Microbial Resource Research Infrastructure), INRAE, FR, 3) MIRRI, Université Paris-Saclay, INRAE, CIRM (Centre International de Ressources Microbiologiques), Unité MaIAGE, FR, 4) UMR 1027 Inserm - Université Paul sabatier Toulouse III, FR

contact: romain.david@erinha.eu https://orcid.org/0000-0003-4073-7456, **Twitter: @ERINHA_RI**