

Transnational Access to Confidential Microdata

Progress and Impact for Research

Joint UNECE and Eurostat expert meeting on statistical Data Confidentiality

1-3 December 2021

Roxane Silberman

Maria Alkhoury

Centre d'Accès Sécurisé aux Données

Secure Data HuB



C/SD/C

Introduction

- ❖ Highly detailed (confidential) administrative microdata increasingly used by researchers and important for public policies evaluation
- ❖ Most of the time, not gathered at European and international levels
- ❖ Although access for research increasingly possible at national levels, it remains uneven across borders
- ❖ Feedbacks from existing experience important to move forward and remove obstacles
 - CASD (Centre d'accès sécurisé aux données) in France: 10 years' experience
 - Is there a significant demand? Who? Which data? What kind of projects?



The current landscape

Two main issues

❖ Legal issues

- How to sue people across borders in case of breaches ?

❖ Practical issues depending on modes of access

○ On-site access

- Requires travelling for non-resident researchers
- Financial and time burdens on researchers

○ Remote execution

- No direct access to the data
- Involves batch processing

○ Remote access

- Researchers can “see the data”
- Does not automatically entail access across borders



Current developments

❖ Legal issues

- Equivalence of sanctions and contracts (OECD discussions, Eurostat)
- GDPR for EU, adequacy decisions and contracts for others

❖ Modes of access

- On-site access (also decentralized) still important
- Remote execution: not much welcome by researchers, automatic outputs checking limits,
- Synthetic data difficult to implement widely
- Remote access in development (Statistics Denmark as a pioneer, CBS, CASD, Finland ..., Eurostat soon expected), international projects as NordMAN, IDAN
- However still hesitations re. access across borders: UK, Germany, US...





CASD
Centre d'Accès
Sécurisé aux
Données

The French Secure Data Hub

CASD access modalities

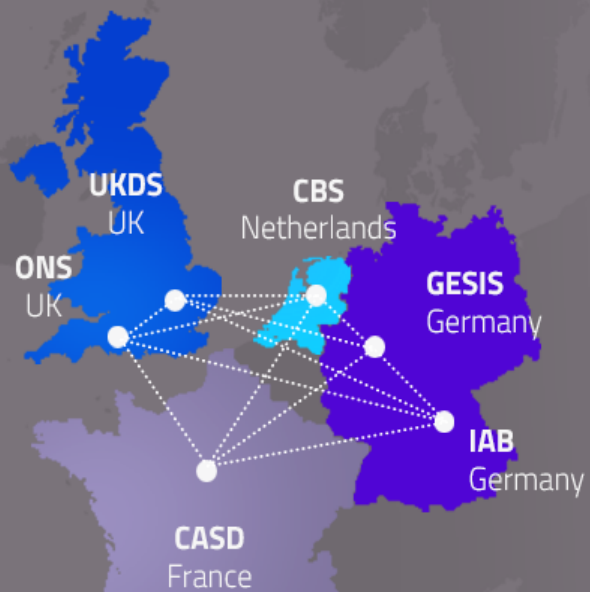
- ❖ **Set up in 2010**, CASD, the Secure Data hub, is a **public institution** with a purpose of organizing and implementing **secure access services to confidential data** for non-profit research, study, evaluation or innovation activities
- ❖ Opting for **a remote access**: a system designed and fully controlled by CASD.
- ❖ Access to the data through a dedicated secured access terminal **the SD-Box** with a **biometric authentication**
- ❖ **Assessment of data access requests** via an independent authority
- ❖ CASD in charge of **contracts** and **enrolment**

The SD-BOX



Opening access across borders

- ❖ Access from EU and AELE countries to all CASD data **was opened in 2011**
- ❖ **In 2018** access to some CASD data was opened to researchers located in North America (Canada and the United-States) under certain conditions
- ❖ CASD partner of IDAN (International Data Access Network)



IDAN

INTERNATIONAL DATA ACCESS NETWORK

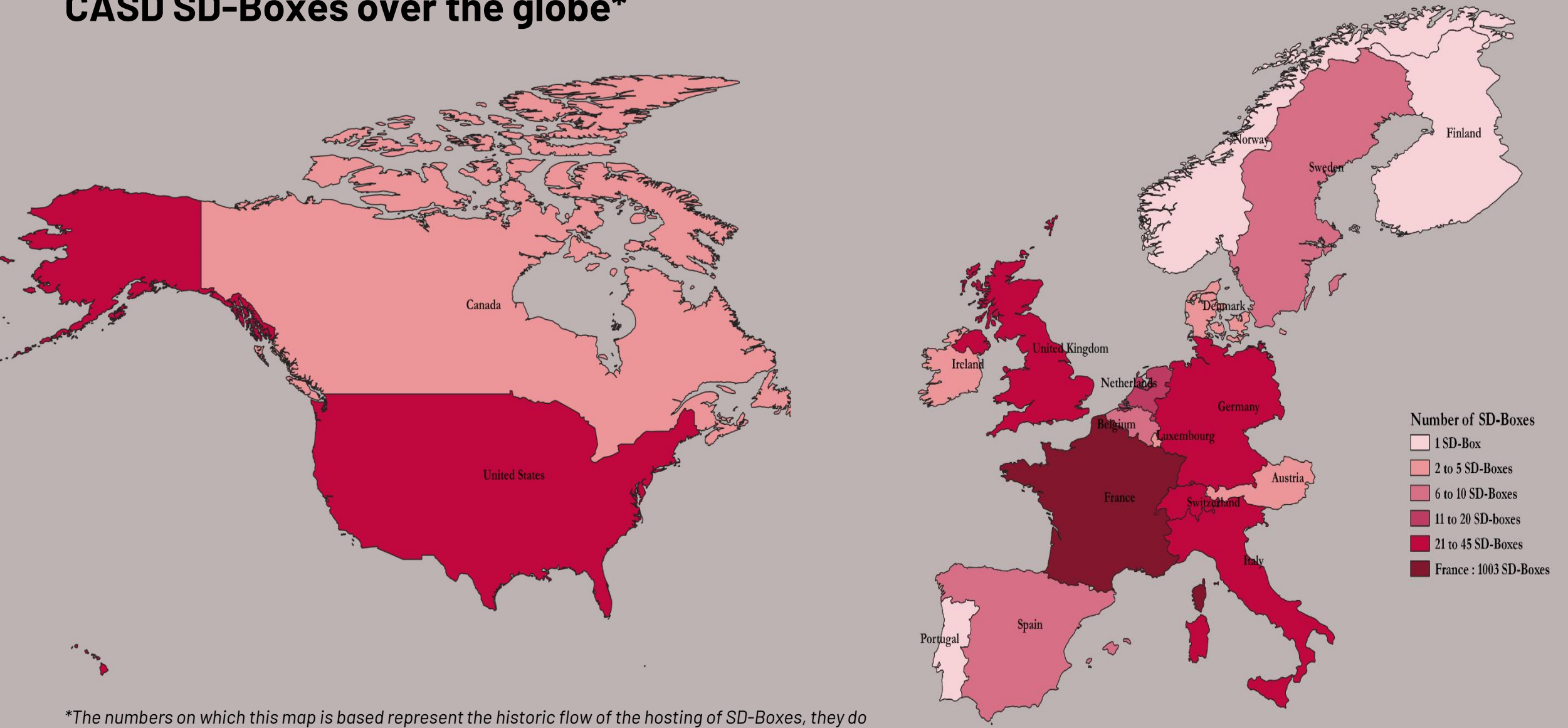


10 years of access
across borders

CASD access points deployed

- ❖ **217 access points (18%)** have been requested to be deployed abroad since 2017
- ❖ **120 SD-Boxes (16%)** over a total of 730 are currently hosted outside of France
- ❖ 5 countries host **about 70% of SD-Boxes** outside of France: **United Kingdom, United States, Italy, Germany and Switzerland**
- ❖ New European countries each year
- ❖ **12 SD-Boxes** were sent to the US in 2019 when access was authorised from this territory

CASD SD-Boxes over the globe*

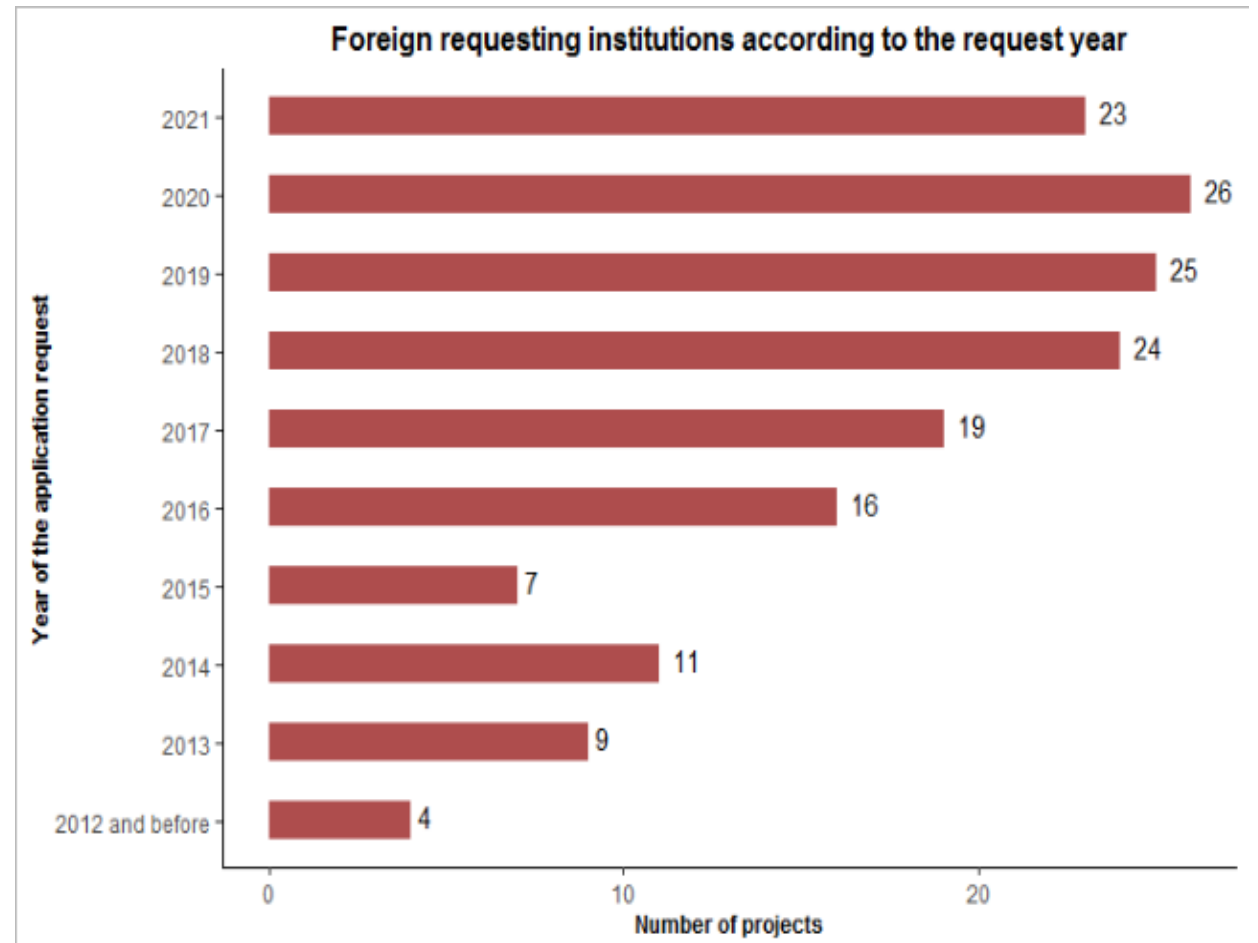


*The numbers on which this map is based represent the historic flow of the hosting of SD-Boxes, they do not represent the number of SD-Boxes currently hosted by these countries

With more projects led by foreign institutions

- ❖ In **15.5%** of all CASD projects (frequently involving several institutions), the requesting institution* for the project is located outside of France
- ❖ A steady increase: from **4 projects** between 2010 and 2012, to **26 projects requested** in 2020

*The requesting institution is the institution of the application procedure. This information is collected by the CSS. It is the institution which is liable when it comes to processing of personal data.



And collaboration between researchers from different countries

❖ Projects involving researchers from different countries (including or not France)

Projects according to the number of countries of the institutions of their members

Number of countries	Number of projects	Percentage
One country: France	811	76.9%
One foreign country	100	9.5%
Two countries	119	11.3%
Three countries	21	2.0%
Four countries	4	0.4%
Total	1055	100.0%

Three types of usage

❖ Supporting researchers' mobility

- Short visits for meetings, workshops or conferences
- Mobility (mainly of French students) for PhD or post-doc position
- Mobility for a teaching position

❖ Use of original and rich data

- Frequently: administrative and linked data

❖ Comparative work

The background features a dark blue gradient with intricate white patterns. On the left side, there is a dense, semi-circular arrangement of thin white lines and small white dots, resembling a stylized globe or a complex network. The right side of the image is more open, with scattered white dots and faint, larger-scale white lines that create a sense of depth and movement.

Conclusion

Supported by continuous investments

- ❖ Translation to English
- ❖ Distant enrolment in the case of CASD
- ❖ Harmonizing data in order to promote comparative work (CASD & IAB work)
- ❖ Use of data from different countries (IDAN project)



We thank you for your attention

CASD 