



seriss

SYNERGIES FOR EUROPE'S
RESEARCH INFRASTRUCTURES
IN THE SOCIAL SCIENCES

www.seriss.eu
@SERISS_EU

Administrative Data Linking: Enriching Administrative Data with Surveys

Tom Emery (NIDI/GGP/ODISSEI)



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 654221.



Background

- What is SERISS
- Administrative Data Linkage Types
- What is the GDPR





SERISS

- SHARE, ESS, EVS, GGP, CESSDA, WageIndicator
 - Infrastructures, not just surveys!
- Working together to find joint solutions to joint problems
- WP6 – Ethical, Legal and Technical Challenges in New forms of Data
 - Social Media data
 - Administrative Data Linkage
 - Biomarkers





Administrative Data Linkage

- How do the social sciences use administrative data?
- What potential future developments exist?
- How will these be affected by new legislation? (i.e. GDPR)

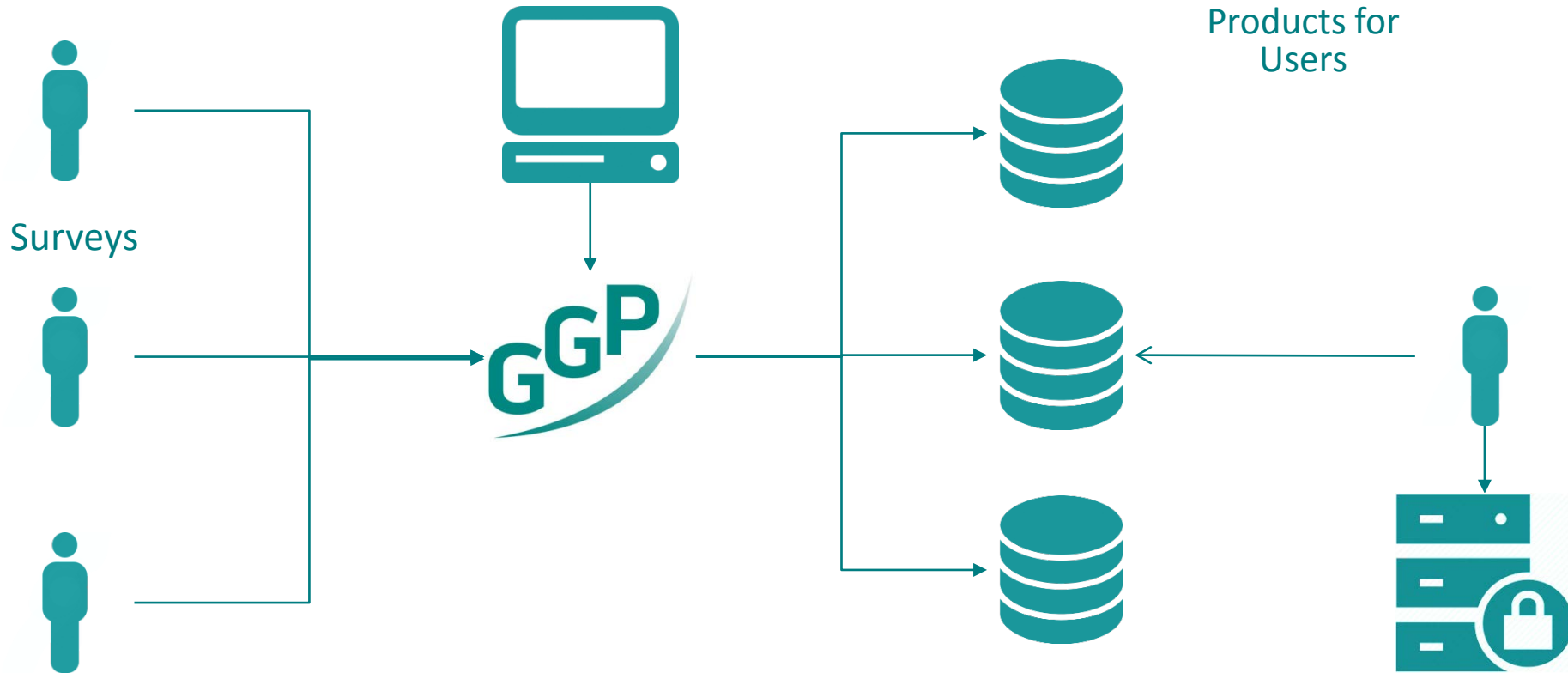




GGP Background

- GGP is a longitudinal, cross-national survey currently conducted in 19 countries
- Survey focuses on families and demographic behavior
- In several countries respondents are linked to population registers
- Less common is integration of administrative data in PUF





Post Hoc – Single Access Linkage

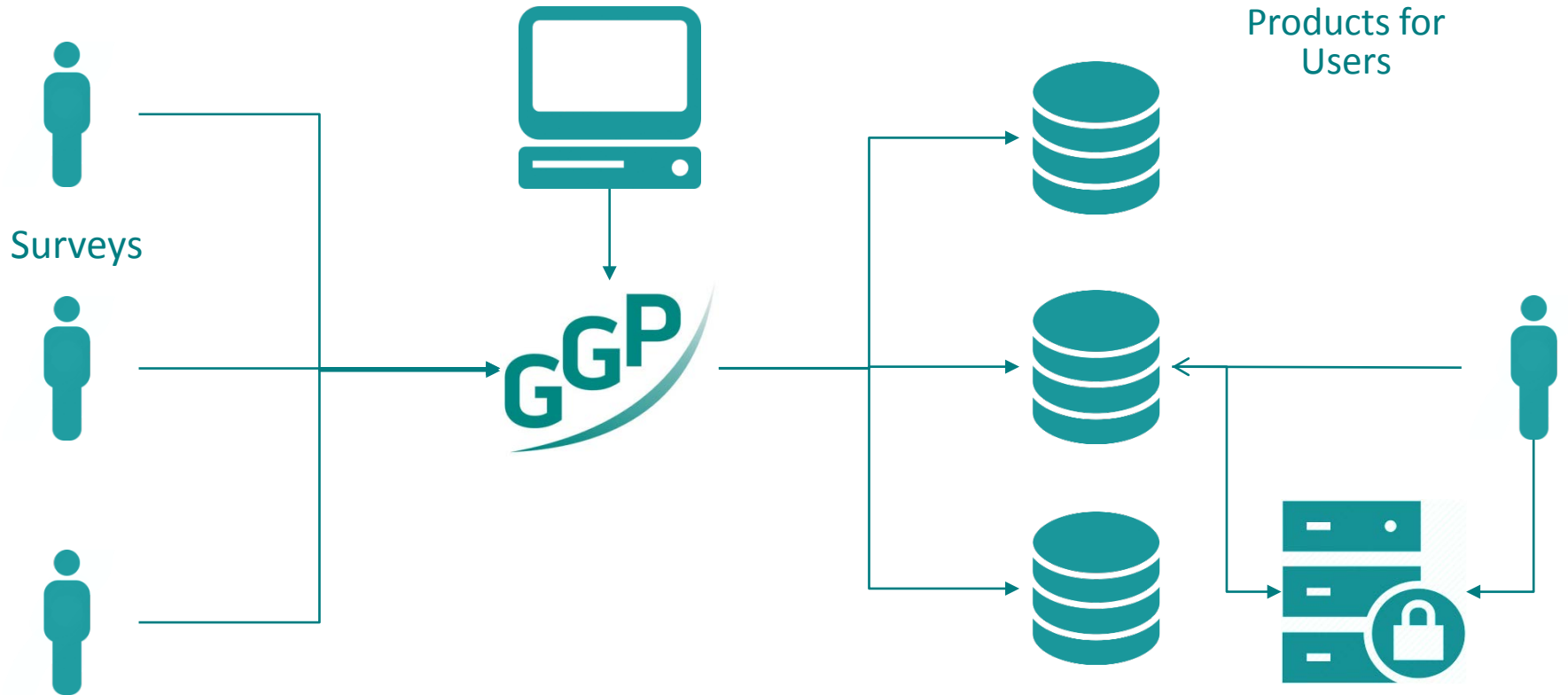




Post Hoc – Single Access Linkage

- Examples: Overview in Emery, Cabaco & Groenewold, 2017
- Advantages:
 - Specific data for specific questions
 - Filling in the gaps of survey questions
- Limitations:
 - Time-consuming
 - Network dependent
 - Access is highly restrictive
 - High Duplication, Low Replication
 - Only Post-Hoc: Can't undo non-response, can't integrate in data collection





Post Hoc – Infrastructural Linkage

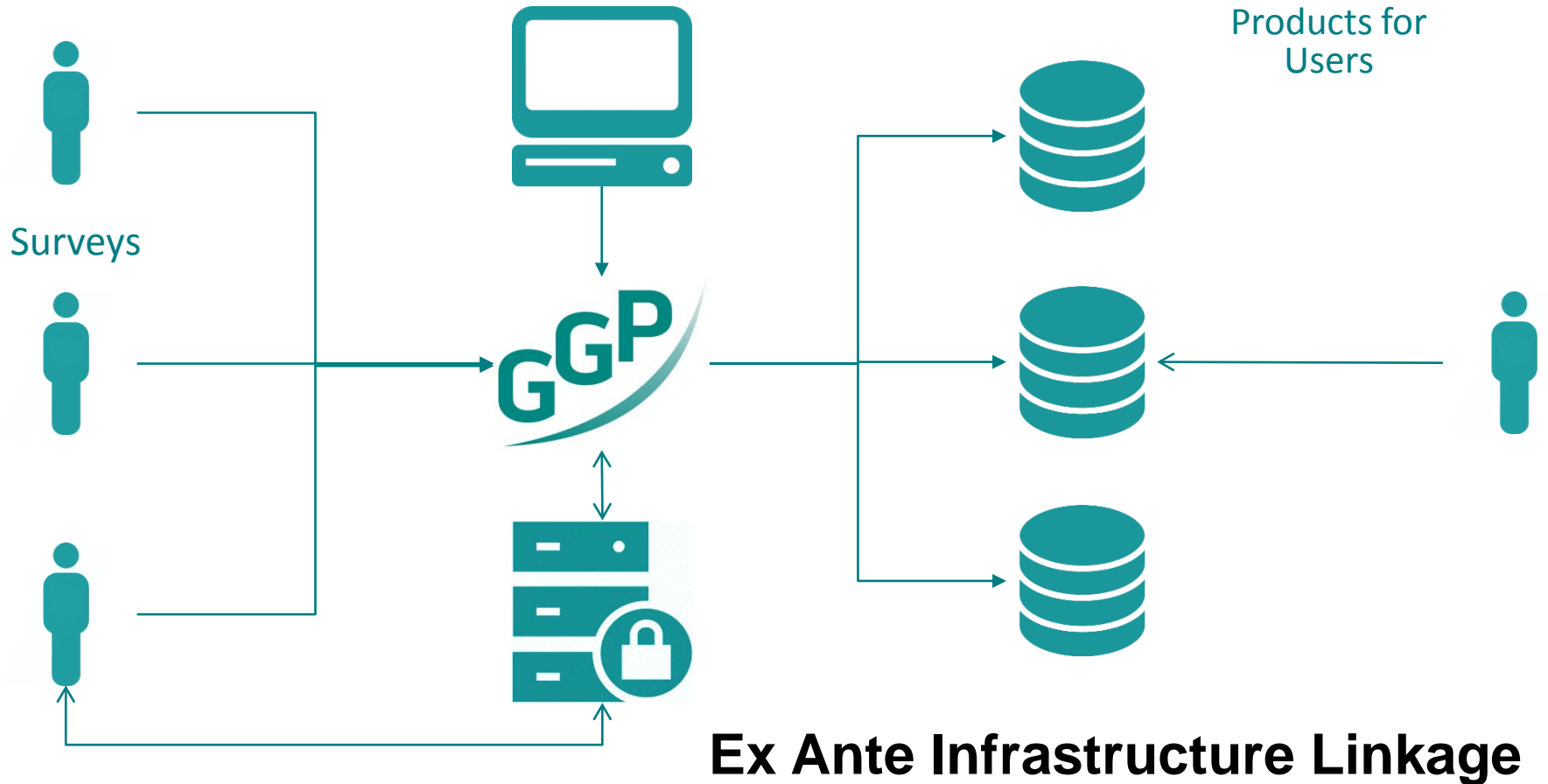




Post Hoc – Infrastructural Linkage

- Examples: ADRN, DWB, NSO's
- Advantages
 - Cheaper than going alone
 - Potential to do comparative work
 - Easier to create a feedback loop between Admin and Survey data processes
- Limitations:
 - Expensive (Disclosure Checks)
 - Access is highly restrictive
 - Inefficient: High Cost, Low Output
- SHARE-RV
 - Cheaper and Easier Access
 - Separate applications
 - Not integrated within Data Processing







Post Hoc – Infrastructural Linkage

- Examples: GGP, ESS
- Advantages:
 - The added value is baked into the survey: better response rates, lower attrition, better quality data
 - Integration is seamless in the eyes of the researcher
 - Potential for experimental usage
- Limitations:
 - NSO's need to transfer data Ownership to infrastructures
 - Few NSO's want to do this due to data protection





Example of Sweden

- Carried out by Stockholm University in collaboration with Statistics Sweden
- 9,688 Respondents aged 18-79 interviewed in 2012-13 (Response Rate of 54.7%)
- Phone survey with postal follow up and registry linkages
- Registry follow up provided every 3 years (2nd Wave is being prepared for release now)





GDPR Background

- General Data Protection Regulation
- Comes into force across the EU in March 2018
- Harmonizes approach to data protection across MS





The Status of Research

- Research should be considered within principle of purpose limitation
 - *may avoid restrictions on secondary processing and on processing sensitive categories of data (Recital 50)*
 - *As long as they implement appropriate safeguards, these organizations also may override a data subject's right to object to processing and to seek the erasure of personal data (Article 89)*
 - *...but NSO is still gatekeeper. Not researchers!*





Right to Data Portability

- *“The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided”(Article 20)*
- *In exercising his or her right to data portability pursuant to paragraph 1, the data subject shall have the right to have the personal data transmitted directly from one controller to another, where technically feasible.*
- *...But NSO's are exempt!*





Squaring the Circle

- *Are NSO's able to transfer data to RI's under Recital 50 and Article 20 combined?*
 - *Does this require a respondents consent?*
 - *Can NSO's wave their right to Article 20 exemption?*
 - *What restrictions should be placed on RI's if they are granted access?*
 - *Can RI's redistribute this data? Under what conditions?*
- *RI's are in a weak position. They are clearly doing research, they can collect informed consent, and they have the technical means to do it.....*
- *But the law places integration at the NSO's discretion*





Next Steps

- Engaging with NSO's
- Highly dependent on National Interpretations
- Workshop in Brussels in January
- Consultation with Article 29 Working Group





seriss

SYNERGIES FOR EUROPE'S
RESEARCH INFRASTRUCTURES
IN THE SOCIAL SCIENCES

www.seriss.eu
 @SERISS_EU

Thank you



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 654221.