An Phríomh-Oifig Staidrimh
Central Statistics Office

Privacy from the Outset

REQUIRED

(Research Ethics in QUestionnaIRE Design)

10th Jan, 2020

John Dunne

CSO, Ireland

An Phríomh-Oifig Staidrimh
Central Statistics Office

Todays talk

Walk through the Research Project journey

using the **GSBPM** as a roadmap (Generic Statistical Business Process Model)

with a **Privacy by Design** approach (Privacy from the Outset, Privacy by Default)

incorporating the **PIA/DPIA** toolkit (Privacy <Data Protection> Impact Assessment)

using the **5 safes** framework as a high level check

We will also talk a bit about **Data Governance**, **Metadata** and **PETs (Privacy Enhancing Techniques/Technologies)** including **pseudonymisation/anonymization**

Finish with a reference list and questions

An Phríomh-Oifig Staidrimh
Central Statistics Office

Project Roadmap
GSBPM

Specify Needs

Design — **Privacy By Design/Default**

Build

Collect

Process

Analyse

Disseminate — **Operational**

Evaluate

An Phríomh-Oifig Staidrimh Central Statistics Office

Privacy by Design
7 Principles

1. **Proactive** not Reactive; **Preventive** not Remedial
2. Privacy as the **Default Setting**
3. Privacy **Embedded** into Design
4. Full Functionality – **Positive Sum**, not Zero Sum
5. End to End Security – **Full Lifecycle Protection**
6. **Visibility** and **Transparency** – Keep it **Open**
7. **Respect** for User Privacy – Keep it **User Centric**

Incorporated into GDPR

An Phríomh-Oifig Staidrimh
Central Statistics Office

Five Safes Framework

Safe Projects

Safe Data

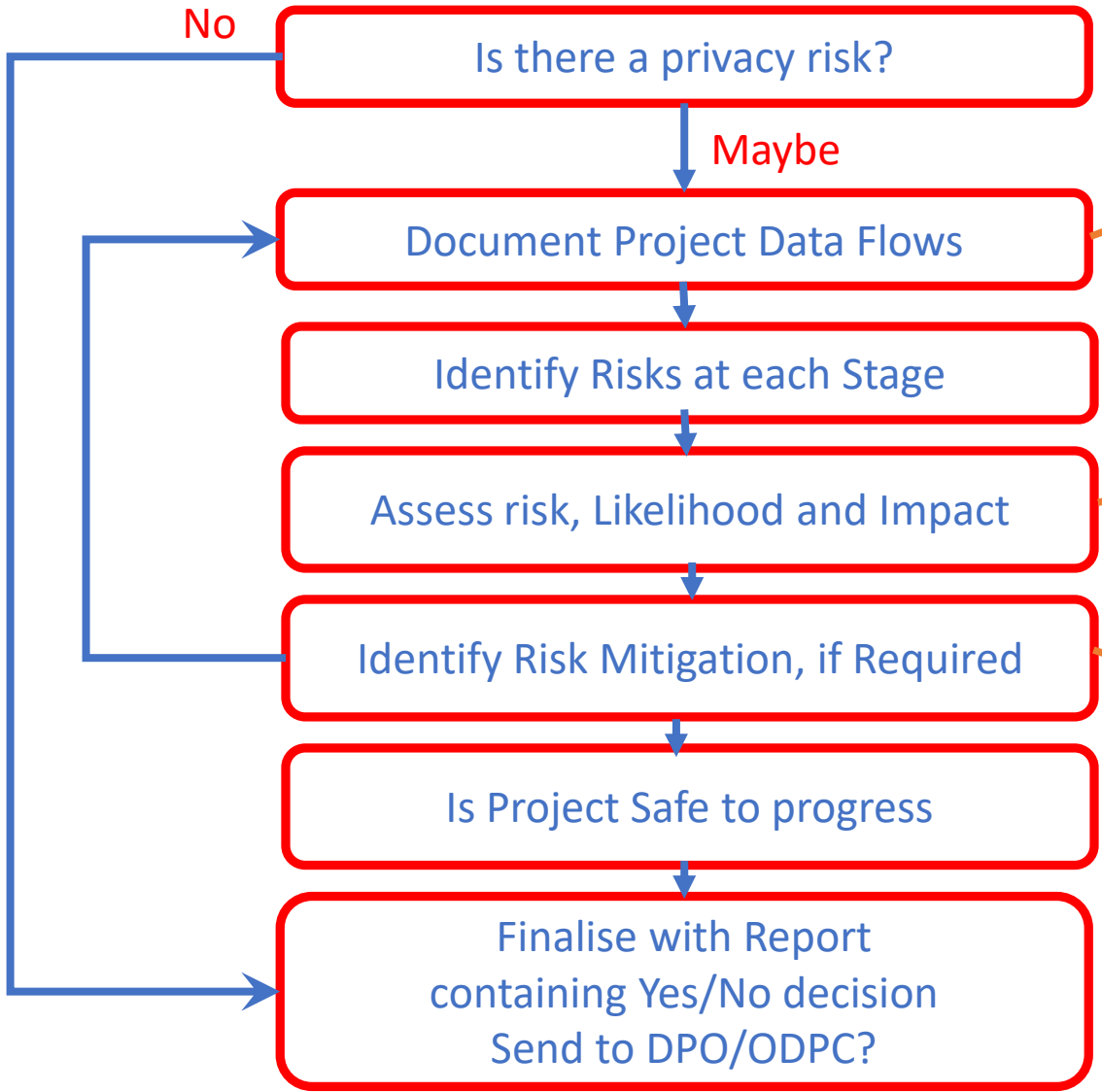
Safe People

Safe Settings

Safe Outputs

An Phríomh-Oifig Staidrimh
Central Statistics Office

Summary PIA/DPIA Process



Social Licence
Data Governance
Metadata
Special categories of data

Identify where to reduce risk

Identify Privacy Enhancing Techniques
to implement
Also data breach protocols

Big benefits to PIA process even if not mandatory
- including addressing openness and transparency (6)

An Phríomh-Oifig Staidrimh
Central Statistics Office

Data Governance

For each Stage of data processing

- Know who the data custodian is (Accountability)
- Know what datasets there are(Metadata)
- Know where the data is, including copies, backups (Data Management)
- Know who should have access/ has access (Access Control)
- Know what data to delete and when (Data Retention)

An Phríomh-Oifig Staidrimh
Central Statistics Office

Role of Metadata

Know it **(List of datasets and locations)**

Understand it **(Data Dictionaries)**

Evaluate it **(Background documentation, questionnaires etc)**

An Phríomh-Oifig Staidrimh Central Statistics Office

Some Collect Stage Considerations

Primary data, Secondary data

> Is there a requirement for primary data collection

Passive data collection, Active data collection

Single data source or multiple data sources

Sample frames

Content – Personal data, Special categories of data

Single data source or multiple data sources

Considerations - informed consent, legal basis, data controller/processor relationship, access control

An Phríomh-Oifig Staidrimh Central Statistics Office

Some Process Stage Considerations

Separation principle

- separate identity from attribute data, use pseudonymisation

Record linkage

- Deterministic, Probabilistic, Statistical

Use pseudonymisation where relevant

Considerations – data minimisation, access control, privacy risk

An Phríomh-Oifig Staidrimh Central Statistics Office

Some Disseminate Stage Considerations

Ensure Aggregate outputs (tables) are not disclosive

> Concept of primary and secondary confidentiality

Be familiar with Statistical Disclosure (SDC) techniques (see references)

Archive anonymous dataset for other researchers (along with documentation)

Delete relevant datasets on project completion

Considerations – Statistical Disclosure Control, Anonymisation, Data Retention

An Phríomh-Oifig Staidrimh
Central Statistics Office

Anonymisation V Pseudonymisation

Personal data

Pseudonymisation is an anonymization technique but does not in itself fully anonymise data

Before classifying somethings as anonymous, consider

i) is it still possible to single out an individual?
ii) is it still possible to link records relating to an individual?
iii) can information be inferred concerning an individual?

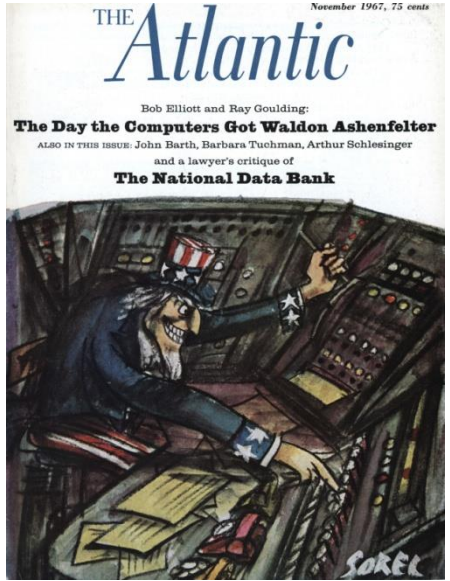
Pseudonymisation is primarily used as part of toolkit to minimise privacy risk

An Phríomh-Oifig Staidrimh
Central Statistics Office

Something from the past

Statistical Deja Vu: The National Data Center Proposal of 1965 and Its Descendants
Rebecca S Krauss - Journal of Privacy and Confidentiality (2013)

"Tyranny of the Statistic," *Christian Science Monitor*, July 29, 1966
"Computer Abuse Threatens Privacy," *Systems*, September 1966
"Computer as Big Brother," *Pittsburgh Post-Gazette*, August 1966
"Big Brother Never Rests," *Indianapolis Star*, August 15, 1966
"A Giant Peeping Tom," *Paterson (NJ) Evening News*, August 8, 1966



The National Data Center and Personal Privacy
By
Arthur R Miller

An Phríomh-Oifig Staidrimh
Central Statistics Office

Useful References (1)

GSBPM https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1

Privacy by Design

- https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf

PIA Google Search **site:hiqa.ie PIA**

DPIA Google Search **site:dataprotection.ie DPIA**

5 safes

- https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control/five-safes

Pseudonymisation and Anonymisation

- Google Search **site:dataprotection.ie Guidance on pseudonymisation and anonymisation**

An Phríomh-Oifig Staidrimh Central Statistics Office

Useful References (2)

Data Quality and Record Linkage Techniques by Herzog, Scheuren and Winkler (Textbook)

Handbook on Statistical Disclosure Control (SDC)

https://ec.europa.eu/eurostat/cros/content/handbook-sdc_en

Guidelines for Output checking (SDC Tabular output)

https://ec.europa.eu/eurostat/cros/system/files/dwb_standalone-document_output-checking-guidelines.pdf

Privacy Enhancing Technologies

Google Search **site:ENISA.EUROPA.EU Privacy Enhancing Technologies**

Irish Social Science Data Archive (ISSDA) http://**www.ucd.ie/issda**

An Phríomh-Oifig Staidrimh
Central Statistics Office

Thank you

Questions