Privacy Impact Assessment

Concetta Tania Di Iorio
Serectrix
tania_diiorio@virgilio.it

Privacy Impact Assessment of the B.I.R.O. Information System

Introduction:

Privacy impact assessment is a systematic and flexible process for evaluating a proposal/project in terms of its impact upon privacy, which has been specifically adapted to the BIRO context

Objectives:

To provide a definitive description of privacy risks, applicable privacy legislation and mitigation strategies adopted in the implementation and management of the BIRO Information System

Materials & Methods

The procedure involved 4 consecutive steps:

Step 1: Preliminary PIA

Step 2: Data Flow Analysis

Step 3: Privacy Analysis

Step 4: Final Report

Preliminary PIA

- Discussion on data flow: physical/logical separation of personal information/data
- Systematic review of the privacy literature:
  - Ovid Medline: 64 biomedical and 11 law articles *were identified*
  - Second search on Law Journals using the same criteria
  - A core set of fourteen papers was selected by comparing abstracts against main project objectives
- Papers were reviewed by the PT to complete a comprehensive report of the first step and identify a short list of possible candidate architectures.

Data Flow Analysis

- Delphi Consensus Procedure to define the best alternative, using the following materials:
  - **data flow tables** (DFT), including the possible scenarios for the collection, use and disclosure of personal information/data, with a number of possible options
  - **information flow questionnaire** (IFQ), to assign marks to each scenario/option
  - **overall consensus table** (OCT), ranking scenarios/options
- Materials were assembled using the procedure presented in the following figure

Procedure

Data Flow Table

CANDIDATE ARCHITECTURE 2: AGGREGATION BY GROUP OF PATIENTS

Scenario 1: Grouping condition directly set by statistical object (e.g. ordered frequency distribution of LOS by CENTRE to compute variability of medians)

Description of personal information / Data clusters	Collected by	Type of format	Used by	Purpose of collection	Transmission to BIRO: de-identification	Security mechanisms for data transmission	Format of BIRO Database	Disclosed to	Storage or retention site
NO aggregation size limit OR min aggregation N=5 patients per cell OR min aggregation N=5, only applicable for high critical privacy variables e.g. service centre, geographical site etc	BIRO partner	One Record for each aggregation level	BIRO partner (local engine), BIRO Consortium (central engine)	Computation of single BIRO statistical object for local and SEDIS reporting	OPTION 1: All DATE fields transmitted as in original OPTION 2: DATE fields approximated to time interval (e.g. months)	OPTION 1: Password access for local administrator prompting client program to send encrypted bundles to BIRO OPTION 2: Client program automatically sending encrypted data (agent)	Separate sets of aggregated tables linkable by predefined statistical criteria	OPTION 1: BIRO database administrator OPTION 2: All local database administrators	OPTION 1: BIRO Coordinating Centre OPTION 2: EU (DG-SANCO)
Aggregation across service centres									

Data Flow Questionnaire

SCENARIO 1:
Question 1. PERSONAL INFORMATION/DATA CLUSTER: DECISION 1

Option	Privacy				Information Content	Technical Complexity
	Identifiability	Linkability	Observability	Overall	Overall	Overall
No Aggregation size limit						
Min aggregation N=5 patients per cell						
Min aggregation N=5 patients per cell, only applicable for high critical privacy variables e.g. service centre, geographical site etc						

Overall Consensus Table

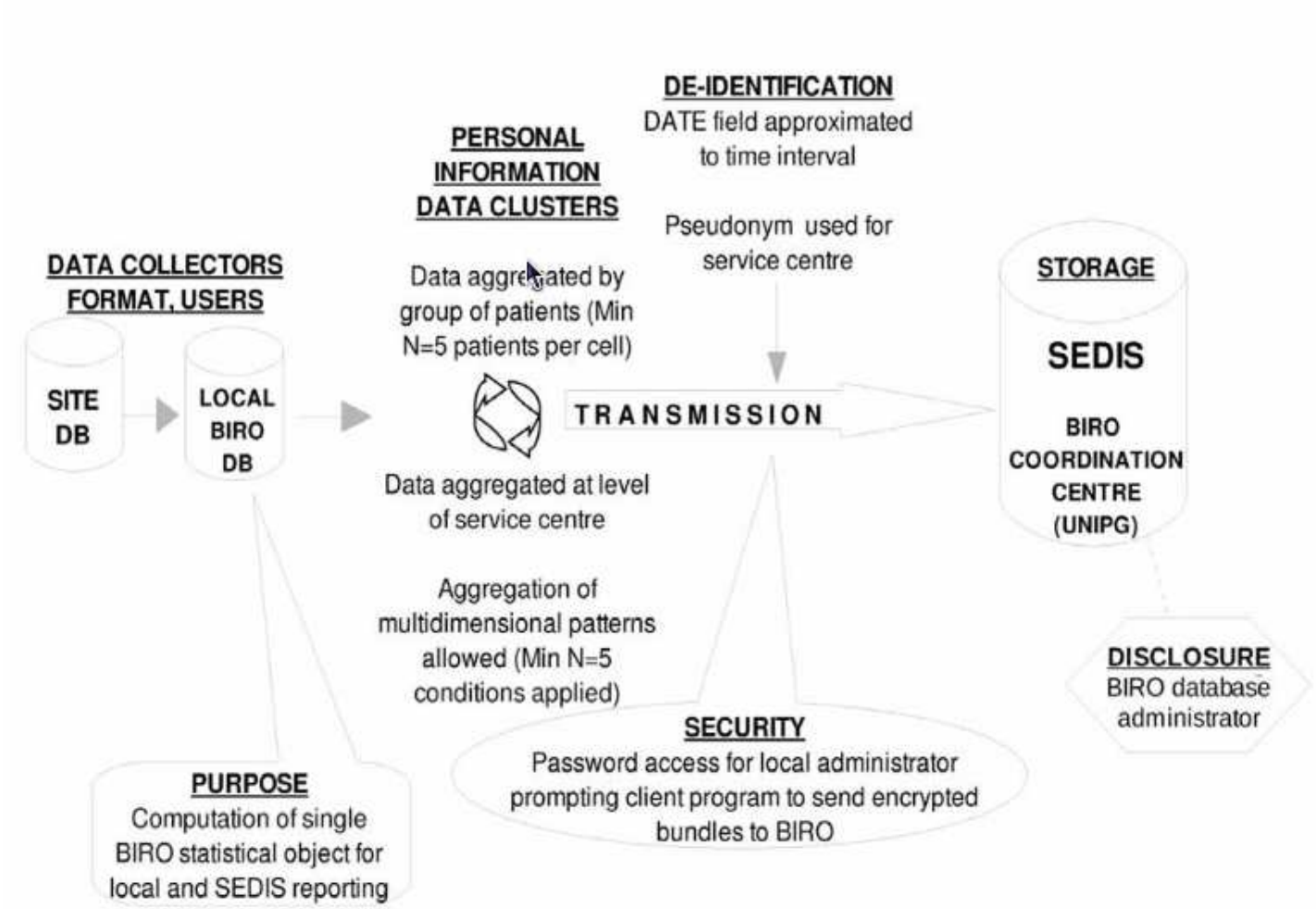
Privacy Analysis & Final Report

- **Privacy Analysis**
  - Cover issues arising in data transfer from local centres to the central database
  - Potential privacy risks identified and analysed through a summary table indicating mitigation strategies to be implemented
  - The level of risk was classified according to an ordinal scale of intensity
- **Final Report**
  - In depth analysis of the selected architecture
  - Compilation of all materials/results into an overall report

Results

- Three main candidate architectures were identified:
  - “individual patient data, de-identified through a pseudonym”
  - “aggregation by group of patients, with Centre’s IDs available in de-identified form, securely encrypted”
  - “Aggregation by Region”
- The Delphi panel selected the second one as the best alternative by ranking the three alternative scenarios, including options for their implementation

Best architecture



Discussion: Privacy Analysis

- The BIRO Information System involves the use of sensitive-medical data collected through diabetes registries within national boundaries and further processed for public health studies at international level

- At a general level, the kind of processing that takes place in the BIRO centres is legitimate ex Article 8 (3) of the Data Protection Directive

Discussion: Privacy Analysis (2)

- BIRO centres anonymise data before any transfer to the BIRO central database
- Recital 26 of the EU Directive, anonymisation allows personal data processing without consent: BIRO processing falls outside the scope of the data protection principles
- The BIRO system processes only statistical objects stored as aggregate comma delimited files
- Pseudonyms for Centres' IDs

Discussion: Privacy Analysis (3)

- The further processing by the global statistical engine cannot pose any privacy risk, either directly or indirectly
- Trans-border data flow envisaged in BIRO is legally viable according to the EU legislation.
- Publication of project results is performed to avoid any direct/indirect identification of data subjects and/or local centres

Privacy contingency risks

Element	Nature of risks	Level of risks			Comments	Mitigating Mechanisms
		Low	Medium	High		
Individual data: Pseudonym used for patients' IDs + Data is Aggregated (N=5 patient per cell)	Individual privacy	X			Pose an indirect risk to individual's privacy	Non-Reversible De-identification
Pseudonym used for Centres IDs	Non-Individual Privacy		X		Pose an indirect risk to Centres' privacy	Reversible De-identification + Reporting System: percentage
Data Transmission	Security Measures	X			Pose an indirect risk to individual's privacy	Encryption
Access to the BIRO network	Security Measures		X		Pose an indirect risk to individual's privacy	Secure applications Hacking tests
Global Statistical Analysis	Individual privacy + Non-Individual Privacy + Security Measures	X			Pose an indirect risk to individual's privacy and centres privacy	Non-reversible de-identification + Encryption

Conclusions

- Privacy impact assessment shows that the selected BIRO architecture fulfils privacy protection requirements by addressing and resolving broad privacy concerns from different angles.
- The architecture of the system flexibly affords the best privacy protection in the construction of an efficient model for the continuous production of European diabetes reports.
- The privacy impact assessment method developed and applied in B.I.R.O. may represent a general tool that can be used to design trans-border health information systems.