B.I.R.O.

PRIVACY IMPACT ASSESSMENT
STEP 2:
DATA FLOW ANALYSIS

DUNDEE MEETING
(26-27 March 2007)

Dr. Concetta Tania Di Iorio
Serectrix s.n.c
Tania_diiorio@virgilio.it

The PIA Process: A Reminder

The PIA provides a balanced approach that allows:

- to realize the best, most privacy protective solution for the B.I.R.O. Information System and
- to easily demonstrate that the very best possible solution has been delivered

The PIA process includes 4 steps:

- Step 1: Preliminary PIA
- Step 2: Data Flow Analysis
- Step 3: Privacy Analysis
- Step 4: PIA Report

What Has Been Done So Far?

At completion of step 1 of the Privacy Impact Assessment, the following objectives have been reached:

- The PIA Team has been set up
- A summary evaluation of potential privacy risks of the BIRO Information System have been carried out
- A Checklist of key privacy requirements/criteria has been produced
- The main alternatives for the BIRO architecture have been selected
- The Preliminary PIA Report has been successfully delivered to the Commission

STEP 2: DATA FLOW ANALYSIS

Objectives

1) To develop a detailed description and analysis of BIRO data flow

2) To identify the best privacy enhancing system architecture for BIRO

   (derived from a detailed description and in-depth analysis of the selected alternatives)

Step 2 - Objective 1

Developing a Detailed Description and Analysis of BIRO Data Flow

In order to document the BIRO data flow the PIA Team should undertake the following activities:

A. **Describe and analyse the BIRO Health Information System architecture through a <u>diagram</u>**

A. **Describe the information flow involved in project through**
   - Identifying **<u>clusters of personal information/data</u>** involved in BIRO System
   - Developing a detailed **<u>data flow table</u>**

Step 2 - Objective 1

Developing a Detailed Description and Analysis of BIRO Data Flow

TASK A: B.I.R.O. Diagram

The BIRO Health Information System Architecture Diagram should document:

- The general BIRO infrastructure architecture
- The flow of information through the system
- Any physical or logical separation of personal information/data and/or
- Security mechanisms that prevent improper access to personal information/data and/or
- Means to maintain any required separation

Step 2 - Objective 1

Developing a Detailed Description and Analysis of BIRO Data Flow:

TASK B: B.I.R.O. Information Flow

In order to describe the information flow involved in project, the PIA Team should:

- Identify **clusters of personal information/data** involved in BIRO System
  - Describe all personal data elements associated with the proposed system. As an example, a data cluster could be elements of patient identification (name, country of birth, ethnicity, etc.)
- Develop a detailed **data flow table**
  - describe the collection, use and disclosure of personal information/data in the BIRO project

Step 2 - Objective 1

Developing a Detailed Description and Analysis of BIRO Data Flow:

TASK B: B.I.R.O. Information Flow

INSTRUMENT: Data Flow Table

A detailed data flow table of personal information/data follow each data element or cluster from collection, use, disclosure and to disposition, in particular it should include:

- Information on data sharing, data retention and data disposal
- Information on:
  - the source of data
  - how information is acquired (directly, indirectly)
  - authority to collect
  - the use and purpose of collecting information (authority for use)
  - disclosure and retention (security levels for information)
  - how long information will be retained and
  - where it will be retained

Data Flow Table

Dimension / Candidate Architecture	Description of personal information/ data cluster	Collected by	Type of format (e.g. paper, electronic)	Used by	Purpose of collection	Disclosed to	Storage or retention site
A							
B							
C							

Link to Step 2 – Objective 2

Data Flow Table → **Questionnaire**

Step 2 - Objective 2
Identifying the Best Privacy Enhancing System Architecture for BIRO

- The activity consists in ranking the three BIRO Information System alternative architectures, identified in Step 1, through a Consensus Panel (modified Delphi Panel)

- The best scoring alternative will be implemented in the BIRO project

Step 2 - Objective 2
Identifying the Best Privacy Enhancing System Architecture for BIRO

Procedure

- Set up Consensus Panel (modified RAND Delphi Panel) to evaluate BIRO candidate architectures
- Define Panel Ranking Form through general consultation (Dundee Meeting+Electronic Communication+BIRO Forum, **April 2007**)
- Use Panel Ranking Form to assign marks to each criterion for all alternatives – **REMOTE** (Electronic Communication, **May 2007**)
- Consensus Panel PIA Meeting (Cyprus Investigator Meeting end **May 2007**)
- Analyse results and rank alternatives (**June 2007**)
- Select best scoring privacy enhancing system
- Finalise PIA Update Report **by July 2007**

INSTRUMENT
Questionnaire (Panel Ranking Form)

Candidate Architecture	Dimension	Scoring * Dimensions: Privacy Protection	Information Content	Technical Complexity
A	Description of personal information/ data cluster			
	Group patients by min N=5 per pattern	5	3	2
	Group patients by classes of Gender, Age	5	2	1
	Collected by			
	Used by			
	Type of format (e.g. paper, electronic)			
	Purpose of collection			
	Disclosed to			
	Storage or retention site			
	TOTAL	5	2	1

* Min=1, Max=5

Scoring Problems

- Definitions
- Identify major dimensions (scoring columns)
- Agree metrics
- Identify Scoring Dimensions
- Identify Weights for a Total Score
- Identify Composite Score

Fundamental scoring dimension: Privacy

A score on privacy can be based on three separate criteria:

1) Identifiability

2) Linkability

3) Observability

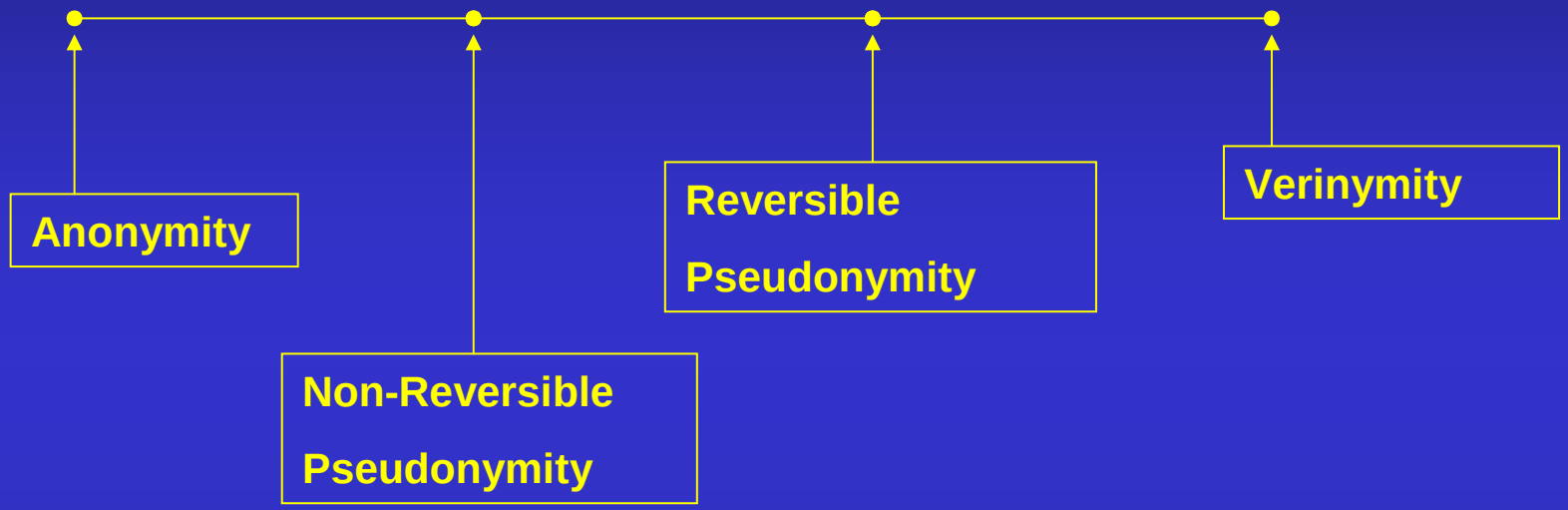
Step 2: Privacy Metrics

Criterion 1: Identifiability

- Measures the degree to which information is personally identifiable
- The Identity measurement takes place on a continuum, from full anonymity (the state of being without name) to full verinymity (being truly named)
- The goal of the PIA Team is always to decrease the amount of identity in a given system.
- A minimalist design approach should be employed and if identity data is not required, it should be intentionally removed from the architectural equation
- Many tools employing reversible and non-reversible pseudonymity are available for this purpose

Step 2: Privacy Metrics

Criterion 1: Identifiability



Potential Marks

Anonymity_____= 4

Non-Reversible Pseudonymity_____= 3

Reversible Pseudonymity_____= 2

Verinymity _____= 1

Step 2: Privacy Metrics
Criterion 2: Linkability

- Measures the degree to which data elements are linkable to the true name of the data subject
- Unlinkability means that different records cannot be linked together and related to a specific personal identity.
- Complex interrelations need to be taken into account: record linkage can be subtle, as it may be organized and/or made possible in different ways

Step 2: Privacy Metrics
Criterion 3: Observability

- Measures the impact of identity or linkability on the use of a system
- It considers any other factor relative to data processing (time, location, data contents) that can potentially affect the degree of identity and/or linkability (effect modifiers)

Step 2: Privacy Metrics Conclusions

- Although the proposed metrics do not produce objective measurements (need to identify/develop standards)…
- they can represent the building blocks of a scoring system underpinning a fair comparison of different solutions
- Goal of the PIA Team is to minimize the degrees of identifiability, linkability and observability
- A single privacy score for each questionnaire item can be obtained from a weighted average of the proposed criteria

Privacy in the context of other fundamental dimensions

- A privacy score must take into account other fundamental dimensions of the BIRO information system
- Goal of the system is to compute quality of care and outcomes indicators
- The impact of BIRO on privacy should be a trade-off between:
  - higher levels of privacy protection
  - relevance of information content in relation to target diabetes indicators
  - minimal technical complexity

- The scoring system must produce a composite indicator incorporating the above dimensions to support a final decision on the candidate best architecture

Step 2: Deliverable

The privacy facilitator shall provide the

Data Flow Report (D5.2)

by July 2007