B.I.R.O.

WP 5: PRIVACY IMPACT ASSESSMENT
STEP 2: DATA FLOW ANALYSIS

Ranking BIRO Architectures through the Data Flow Table and Information Flow Questionnaire

CYPRUS MEETING
(23-25 March 2007)

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

Data Flow Analysis (Step 2) Objectives

Objectives:

- ❑ to describe the information flow occurring through the BIRO system
- ❑ to identify the target BIRO architecture.

By means of the *data flow analysis* the PIA Team primarily aims:

- ❑ to develop a detailed description and analysis of BIRO data flow
- ❑ to identify the best privacy enhancing system architecture for BIRO (derived from a detailed description and in-depth analysis of the selected alternatives)

PIA Team Tasks

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

- ❑ to describe and to analyse the BIRO Health Information System architecture through a *diagram*
- ❑ to describe the information flow involved in the project through
    - identifying clusters of personal information/data involved in BIRO System
    - developing a detailed *data flow table*
- ❑ to develop an *information flow questionnaire* from the data flow table
- ❑ to rank candidate architectures based on marks given to each option on the basis of standard criteria involving privacy, information content and technical complexity.

Materials and Methods

- ❑ BIRO HEALTH INFORMATION SYSTEM DIAGRAM
- ❑ DATA FLOW TABLE
- ❑ INFORMATION FLOW QUESTIONNAIRE
- ❑ ARCHITECTURES RANKING

Materials and Methods
1) BIRO Health Information System 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

Materials and Methods
2) Data Flow Table

- ❑ The *data flow table* is a specific tool developed in order to in depth describe the dynamics involved in both data collection and information exchange procedures
- ❑ Data flow tables shall be used for each of the candidate architectures identified in PIA previous step
- ❑ It includes details of personal information/data and how they are handled along the entire process: from collection, use, disclosure and to disposition.

Materials and Methods
2) Data Flow Table: How to describe the BIRO Data Flow

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

- ❑ identify clusters of personal information/data involved in BIRO System
- ❑ describe all personal data elements associated with the proposed system (example: a data cluster could be elements of patient identification e.g. 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
- ❑ list the different options available for data collection and exchange in each BIRO candidate architecture

Materials and Methods:
2) Data Flow Table: Information to be Included in Data Flow Tables

The data flow table includes information on:

- ❑ data sharing, data retention and data disposal
- ❑ source of data
- ❑ acquisition (direct, indirect)
- ❑ authority to collect
- ❑ use and purpose of collecting information (authority for use)
- ❑ disclosure and retention (security levels for information)
- ❑ how long information is retained for
- ❑ where it is retained

The data flow table should highlight all major components to be taken into account in order to rank the different BIRO alternative architectures (described in Step 1 of the PIA process).

Materials and Methods
3) Information Flow Questionnaire

- ❑ The questionnaire has been distributed on the 13th of May 2007
- ❑ Each member of the PIA Team has been asked to fill in the questionnaire *independently* and return it to the BIRO Coordinating Centre by the 18th of May 2007
- ❑ The *information flow questionnaire* has been defined using the various individual components listed in the data flow table
- ❑ The various options have been grouped to specify the different solutions available for the definition of the final structure of the BIRO information system
- ❑ Each item has been evaluated on the basis of three different criteria:
  - privacy protection
  - information content
  - technical complexity

Scoring Dimensions

- ❑ 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.

Scoring Dimension 1. Privacy

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

- ❑ Identifiability
- ❑ Linkability
- ❑ Observability

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 Privacy Architect and the PIA author 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

Linkability & Observability

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

Criterion 3: Observability

- Measures the degree to which identity or linkability may be impacted from 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)

Materials and Methods:
4) Architectures Ranking

The candidate architectures will be evaluated taking into account the results of the questionnaire, according to the following procedure:

- ❑ average marks will be produced for each dimension of any BIRO alternative architecture
- ❑ Those average marks will be communicated to PIA Team Members at the beginning of the Delphi session
- ❑ A discussion will be opened over eventual disagreements on average marks
- ❑ The Delphi Consensus Panel will take any decision by majority (50% + 1), if an agreement is not reached through discussion
- ❑ The best scoring BIRO candidate alternative will be selected