

The Capish Information Model – Simplify Access to Your Data

Anna Berg and Catharina Dahlbo
Capish Nordic AB, Malmö, Sweden

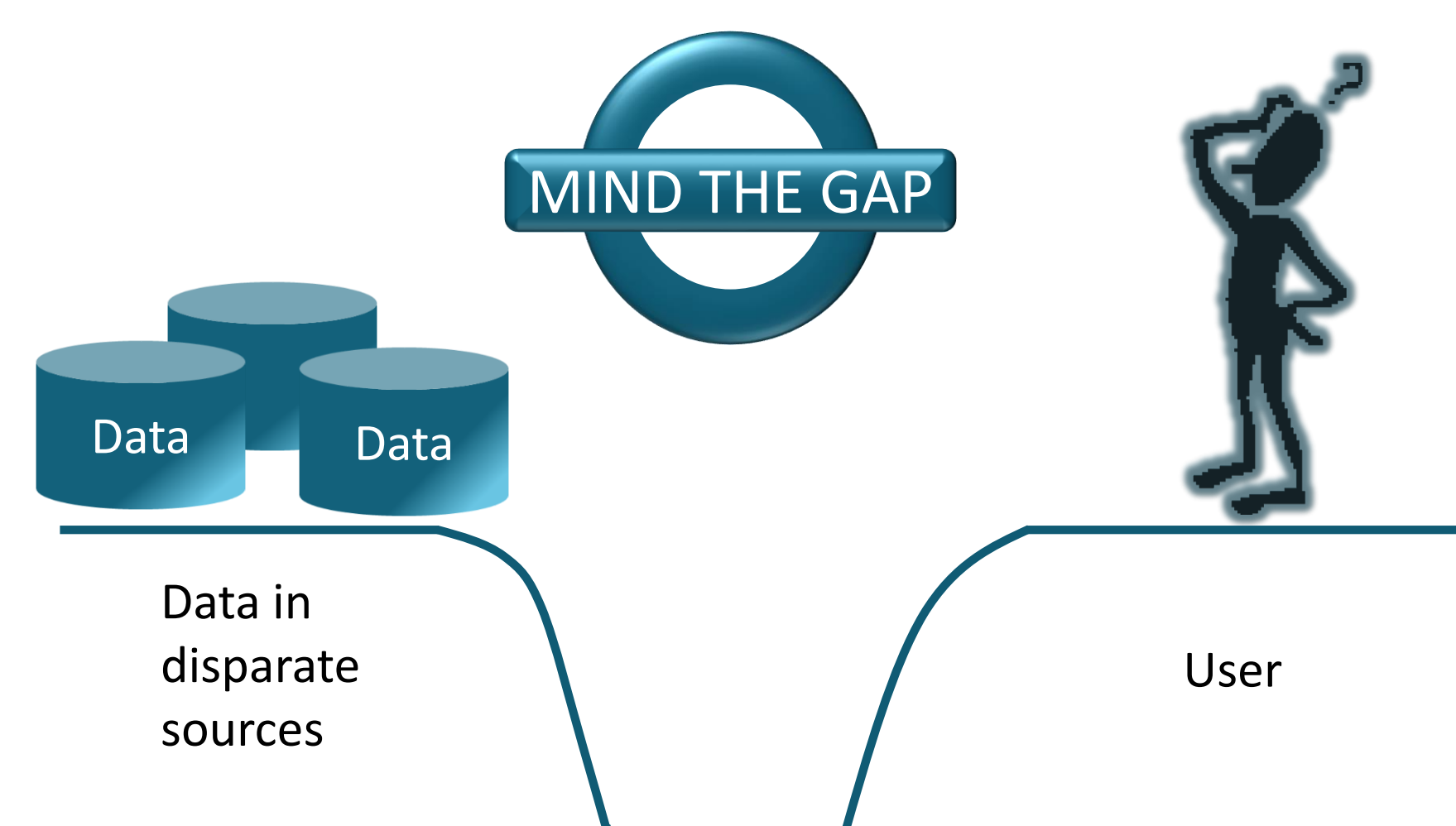


Figure 1. Bridging the gap between user and data is one of the challenges in data transparency.

Clinical Data Transparency Challenges

Opening up clinical data for a wider group of users, including the public, requires the ability to bridge the gap between user and data (Figure 1). Users need to be provided with the possibility to evaluate and analyze the available information, irrespective of its location or the application that created it.

The main challenge is to make the data available in a unified, well-defined and easy to understand format that minimizes the need of database knowledge (Figure 2).

Other major challenges are maintenance of data quality, preservation of patient privacy and the safeguarding of intellectual property rights.

Another important aspect for making data accessible to a wider user group, is to provide them with a flexible and easy to understand tool, which allows them to navigate and search the data with ease.

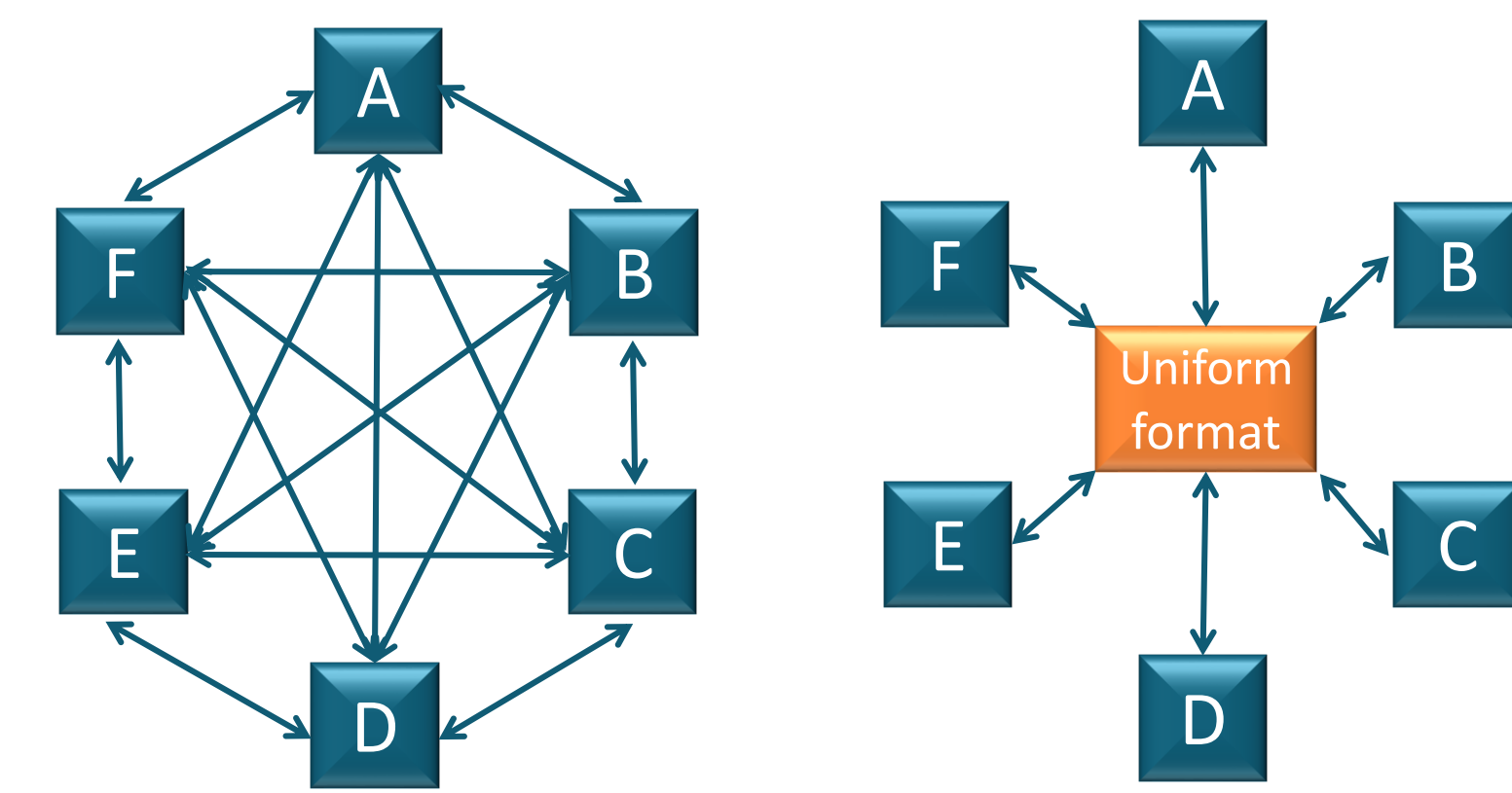


Figure 2. A uniform data format reduces the need for mapping "everything to everything".

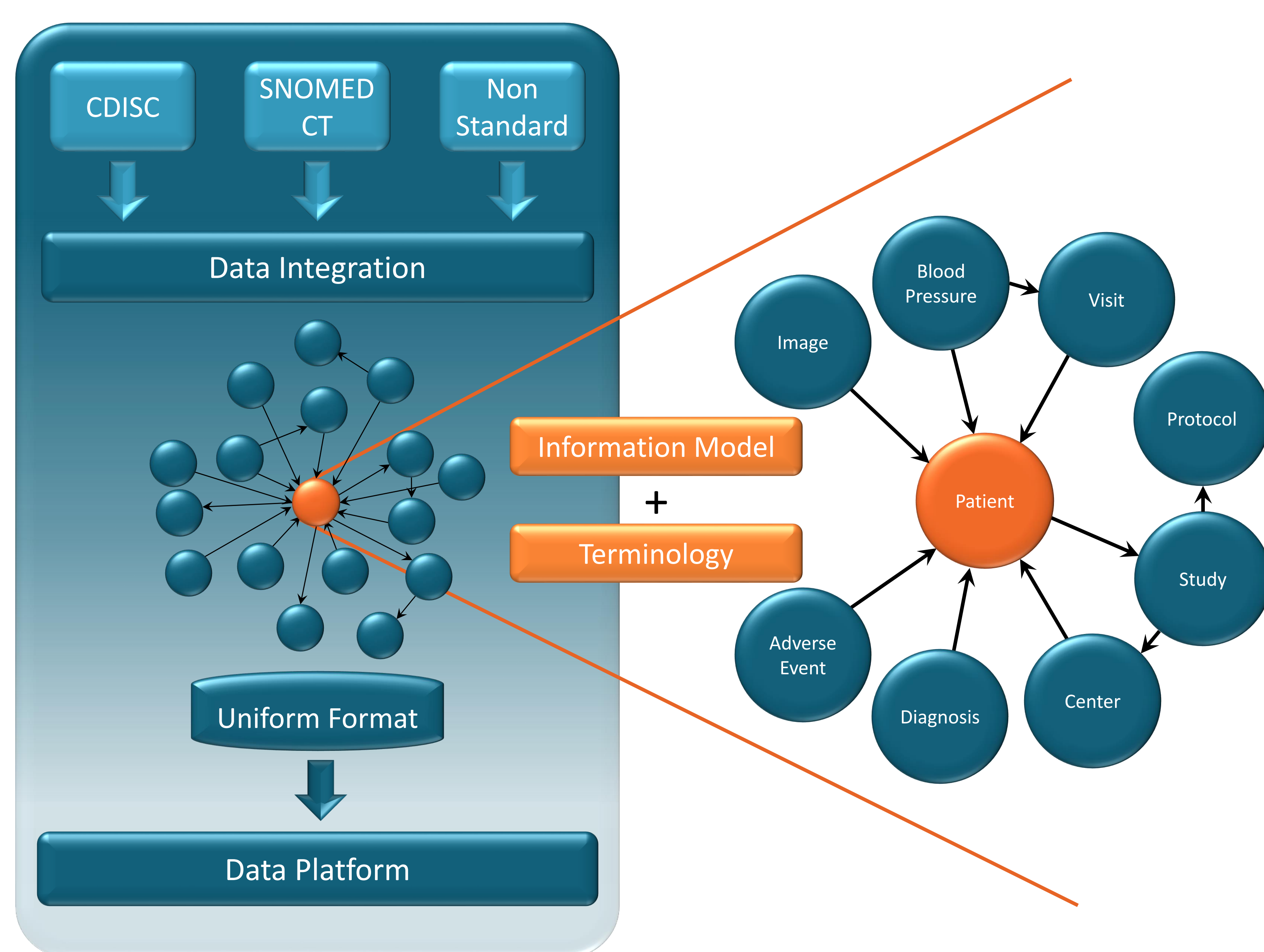


Figure 3. The data integration process with the use of a reality based information model and well-defined terminology, which can be indexed to a data platform for knowledge creation.

Easy Access – Bridging the Gap between User and Data

Data Integration

Data integration is the process in which heterogeneous data from different sources (data sets, documents, images etc.) is combined to a stable and uniform data format. This format has to describe both the data structure, for example how data is decomposed, and the terminology, including both data labels and data contents.

Information Model

A way to achieve an unbiased data format is to use an information model that mimics the real world behind the collected data.

Capish has developed an information model that is constituted of small, self-contained units of information, linked together with relations (Figure 3). These units with relations are beneficial for:

- **Data integration** – all types of information and combinations can be handled
- **Patient privacy** – easy to hide units with personal information
- **Intellectual property** – easy to hide units with confidential information

Terminology

Standardized terminology in output regardless of input is necessary to achieve full integration. Capish is using a well-defined terminology on top of the information model, with the possibility to include existing standards and terminologies.

File Format

One of the key features of the Capish® information model is that all the small information units are stored as simple documents in the non-proprietary language standard XML, making the information easy to archive and retrieve.

Data Platform

The XML-files are indexed to the Capish® data platform, which is designed to give the user advanced search possibilities and the ability to choose centrality of the information model (reflective logic).

Data Navigation

While the information model and the data platform form the foundation, an easy to use, interactive tool to access the integrated data is also necessary.

Capish has developed a browser-based application (Figure 4) that allows physicians, researchers or anyone comfortable within a given knowledge domain to take full advantage of the integrated data:

- **Search** – ideal for free or structured text searches on the entire information base
- **Analyze** – approach the data from different angles, identifying relevant details and get results fast
- **Visualize** – drill down from interactive graphs to individual patient information
- **Explore** – navigate the data via the relations without structural knowledge of the database

The user gets unlimited access to the available data and can access the information from a population or individual point of view, depending on the question.

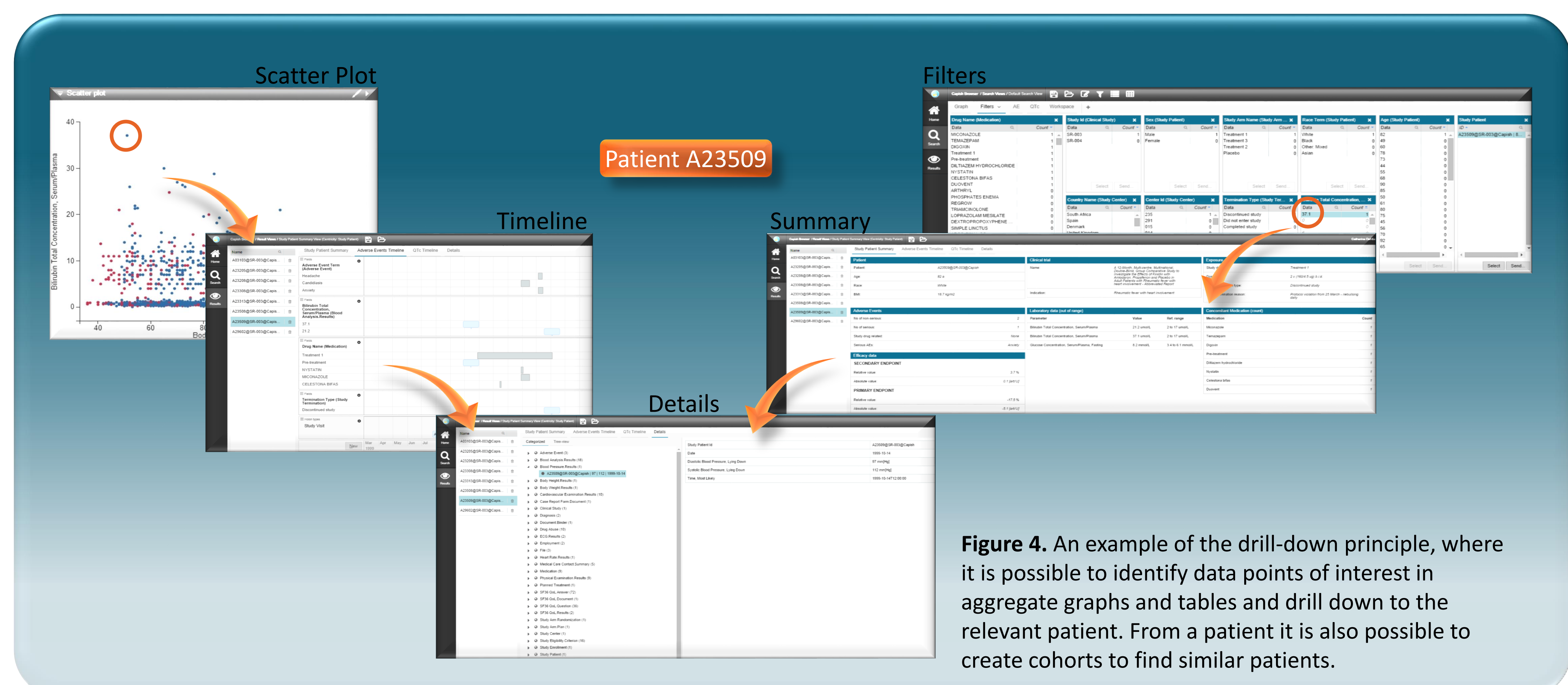


Figure 4. An example of the drill-down principle, where it is possible to identify data points of interest in aggregate graphs and tables and drill down to the relevant patient. From a patient it is also possible to create cohorts to find similar patients.

Conclusion

Data transparency should give the user the ability to access and make sense of data. To achieve this, data has to be well-defined and in a consistent format. A solution that has proven to be useful in the integration of very disparate data, is to use an information model that is built on the actual reality where the data was created. This way to integrate data into information allows people to create knowledge in an easy way on their own, without the need of database knowledge.

About Capish

Capish is a company developing software solutions for the management, exploration and visualization of data. By mapping data to a structured conceptual framework, Capish® Solutions can effectively harmonize, integrate, store and deliver analysis-ready data.

Our mission is to provide the user with easy access to data that can be clearly interpreted and fully explored, without the need of programming or database understanding. The introduction of Capish® Solutions as a central information platform allows data and results to be shared and communicated easily.

Capish
Stortorget 9
211 22 Malmö
Sweden
www.capish.com

