1 Introduction: What Are the Right Questions to Ask for a Human-Centric View of KDubiq?

The relevance of a human-centric view for KDubiq in particular or knowledge discovery (KD) in general arises from the comprehensive nature of KD. KD is more than the application of algorithms – it encompasses the whole process of turning data into knowledge: business/application understanding, data understanding, data preparation, modelling, evaluation, and deployment [1]. People play a pivotal role in this process: people create data, data and knowledge are about people, and people are (or some of them should be) the ultimate beneficiaries of the discovered knowledge. As shown in the previous chapters, ubiquity adds more dimensions to KD; this chapter analyzes how these dimensions affect the user perspective of KDubiq systems.

In order to take the people-centric view we aim for, we consider results from two areas of Computer Science and Cognitive Science that strive towards a human-centric view of information systems. Human-computer interaction (HCI) is the study of interaction between people and computers, which occurs at the user interface (hardware and software). Cognitive modelling (CM) is the use of computational methods to model mental representations and processes and the behaviour resulting from them. Many HCI decisions depend on the cognitive processes that the user goes through or is meant to go through during interaction; in this sense, cognitive modelling is an essential part of HCI.

A full picture of HCI and CM for KDubiq may be obtained by considering the impact of each feature of KDubiq for each stage of KD.

The features of KDubiq to be investigated are (as defined in Chapter 1 Section 2): the presence of a population of devices with computing capabilities (in the following: device ubiquity), an extreme distribution of data (data ubiquity), a continuous flow of incoming data (data flow), and a rich semantics of data due to the heterogeneity of the involved agents. As in Chapter 4 it is helpful for our purposes to distinguish between sources of heterogeneity; we will refer to

* With contributions from the WG6 members:

www.cs.kuleuven.be/~berendt/HCI-ubiq
computing heterogeneity (e.g., of sensors) and people heterogeneity. In line with
the present chapter’s people-centric view, we will go beyond Chapter 4 in our
treatment of people heterogeneity: whereas that chapter states that people are
heterogeneous, we ask in what ways they are heterogeneous and which specific
effects this has on KDubiq. Privacy will also be investigated here, in a people-
centric view that complements the data-centric one of Chapter 5.

The KD stages to be investigated are in principle all the standard stages of KD
(in the sense of CRISP-DM as used above). However, the CRISP-DM stages
structure the process of KD by the activities of “analyst users”, whereas most
of the people interacting with a KD or KDubiq system will be “end users” for
whom the application is made (patients in Chapter 1, Section 6.1, music lovers in
Chapter 1, Section 6.2, car drivers and pedestrians in Chapter 5, Section 2, etc.).
They generally perceive only a coarsened version of the CRISP-DM stages, which
can be described as data collection, data processing, information presentation
and interaction in all stages, and evaluation. Prior to these, and in cases of good
software engineering, end users are also the (knowing or unknowing) subjects of
an earlier phase: that of requirements analysis. We integrate this stage to ensure
that a comprehensive view of users, other stakeholders, and their respective
interests is taken.

Since KDubiq applications are a type of ubiquitous-computing applications,
many observations made in the present chapter also hold true for ubiquitous
computing in general. Nonetheless, we aim to emphasize HCI/CM challenges
specific to KDubiq (as opposed to UbiComp in general) by our foci on and
within these stages: data collection (KD applications need more data than many
non-KD-computing applications), data processing (for example, the inferencing
inherent to KD generates more privacy problems than the “mere” data storage
and retrieval of most non-KD-computing applications) and the impact this has
on information presentation / interaction (for example, predictability and trans-
parency become specific usability issues when one deals with an “intelligent”,
adaptive system).

A people-centric perspective is needed when constructing a KDubiq system
because the KDubiq features (as defined in Chapter 1, Section 2) affect all phases
of knowledge discovery. The present chapter is a short overview of key challenges,
structured by a matrix organisation: What is the impact of KDubiq features on
the KD stages? The chapter concentrates on the most relevant combinations of
KDubiq features and KD stages (see Table 1).

The chapter thus aims at a systematic, “meta-level” analysis of the HCI/CM
challenges faced by the emerging field of KDubiq; rather than at giving an
overview of current – and conceivably soon outdated – concrete usage inter-
faces. We will illustrate the identified challenges, throughout the chapter, by
references to the application example ACTIVITY RECOGNITION (see Chapter 1
Section 6.1).

An outlook and visions for the future conclude the chapter.