Chapter 58
Human–Data Interaction in Healthcare

Federico Cabitza
Università degli Studi di Milano-Bicocca, Italy

Angela Locoro
Università degli Studi di Milano-Bicocca, Italy

ABSTRACT

In this chapter, we focus on an emerging strand of IT-oriented research, namely Human-Data Interaction (HDI) and on how this can be applied to healthcare. HDI regards both how humans create and use data by means of interactive systems, which can both assist and constrain them and the operational level of data work, which is both work on data and by data. Healthcare is a challenging arena where to test the potential of HDI towards a new, user-centered perspective on how to support and assess "data work". This is especially true in current times where data are becoming increasingly big and many tools are available for the lay people, including doctors and nurses, to interact with health-related data. This chapter is a contribution in the direction of considering health-related data through the lens of HDI, and of framing data visualization tools in this strand of research. The intended aim is to let the subtler peculiarities among different kind of data and of their use emerge and be addressed adequately. Our point is that doing so can promote the design of more usable tools that can support data work from a user-centered and data quality perspective and the evidence-based validation of these tools.

INTRODUCTION

Twenty-five years ago, medical informatics was defined as "dealing with the storage, retrieval and optimal use of biomedical data" (Shortliffe et al. 1990). At that time, little emphasis was put on the practices of data production, that is on how medical practice, and single stories of illness, care and recovery are represented, accounted and "datafied" in some valid and reliable manner. However, these practices, which include policies, rules, habits, conventions, tools and techniques, have always been intertwined with and affected by the available ITs, as well as by the expectations of the stakeholders on how to make sense

DOI: 10.4018/978-1-7998-1204-3.ch058

and use of health-related data. Different perspectives on these expectations, and on what valuable health data are, lead to manifest chasms between primary use and other uses of health information, as often discussed in the specialist literature (Fitzpatrick and Ellingsen, 2013). To try to cross these chasms, we need to create the suitable language to describe the differences and give some operational definitions.

We distinguish between three different macro-types of data and the related processes in which these data are either produced, processed or consumed: namely, primary data, which come from a broad range of sources and are produced within a caring process to make its unfolding seamless and smooth (Berg, 1999); secondary data that are derived from the primary data for purposes different than care, like accounting and medical billing (Abdelhak, Grostick, & Hanken, 2012); and tertiary data, that are produced from the secondary data for any unanticipated need of the potential consumers of health services (see Figure 1).

To illustrate this tripartition, an analogy from the agriculture domain can be drawn (Locoro, 2016): primary data are like the produce of the land, which farmers grow for themselves as well as the external market. Secondary data are the product of a transformation of these primary data, like the one performed in food industry where vegetables are cleansed and chopped. Tertiary data are further transformed from secondary data to make them more easily consumable, that is suitable for and conveyed to a broader population of consumers in terms of information services, like fresh-cut vegetable products can be seen as the service to have vegetables already ready-to-eat.

The definitions mentioned above shed light on the relationship between data and their uses, which cannot be overstated. The tripartition that we propose reflects the different uses and practices in which data are produced and consumed and it calls for a specific area of research focusing on how to support people in interacting with their data of concern and in gaining insight from them: Human-Data Interaction (HDI).

BACKGROUND

As rightly noted by Hornung et al. (2015), the expression HDI is neither specific nor new. On the one hand, humans have always interacted with data, that is they have inscribed, read and communicated data in terms of simple marks, more complex symbols and any kind of meaningful signs[1]. On the other hand, this interaction has predominantly been mono-directional until interactive systems have allowed users to get access to data according to their peculiar and situated needs. In the computing literature, the expression has been used since the 1990s, e.g., by Kennedy et al. (1996), to address the needs of the users in the context of data intensive settings and hence "the problem of delivering personalized, context-aware, and understandable data from big datasets" (Cafaro, 2012). More recently a number of authors have tried to define Human-Data Interaction, by focusing on different aspects: Elmqvist (2011) and Cafaro (2012) focused on the material and embodied aspects of interaction; Crabtree and Mortier (2015) prefer to assimilate HDI to "the management and use of personal data in society at large" where *personal* here means that data are either "about individuals" or are delivered in terms of "personalized experiences". In particular, Elmqvist (2011) defines HDI as "the human manipulation, analysis, and sense making of large, unstructured, and complex datasets" and Cafaro sees HDI systems "as technologies that use embodied interaction to facilitate the users' exploration of rich datasets". We agree with Hornung et al. (2015) and Crabtree and Mortier (2014) in regard to the fact that "HDI does not only refer to embodied interaction, but to all kinds of interaction" and that "HDI should investigate 'data that affects people' ".