# Chapter 10
# A Medical Data Trustworthiness Assessment Model

**Bandar Alhaqbani**
*Queensland University of Technology, Australia*

**Colin J. Fidge**
*Queensland University of Technology, Australia*

## ABSTRACT

*Electronic Health Record systems are being introduced to overcome the limitations associated with paper-based and isolated Electronic Medical Record systems. This is accomplished by aggregating medical data and consolidating them in one digital repository. Though an EHR system provides obvious functional benefits, there is a growing concern about reliability trust (trustworthiness) of Electronic Health Records. Security requirements such as confidentiality, integrity, and availability can be satisfied by traditional data security mechanisms. However, measuring data trustworthiness is an issue that cannot be solved with traditional mechanisms, especially since degrees of trust change over time. In this chapter, a Medical Data Trustworthiness Assessment model to assist an EHR system to validate the trustworthiness of received/stored medical data based on who entered the data and when is presented. The MDTA model uses a statistical approach that depends on the observed experiences available to the EHR system. In order to provide an accurate trustworthiness estimate for historical medical data, a time scope around the time when the data was entered was used. This scope enables the model to capture the dynamic behavior of the data entry agent's trustworthiness. To conduct this assessment medical metadata is used to extract information about the medical data sources (e.g. timestamps, and the identities of healthcare agents and medical practitioners) and, thereafter, this information is used in a statistical process to derive a trustworthiness value for the medical data. The result can then be expressed in the displayed health record by manipulating the EHR's metadata to alert the medical practitioner to possible trustworthiness problems.*

DOI: 10.4018/978-1-60960-174-4.ch010

## INTRODUCTION

Electronic Health Records can enable efficient communication of medical information, and thus reduce costs and administrative overheads (Blobel, 2004; Gunter & Terry, 2005). However, to achieve these potential benefits, the healthcare industry needs to overcome several significant obstacles, in particular concerns about the trustworthiness (reliability) of EHR medical data. Trustworthiness is a crucial factor that has a strong effect on how medical practitioners use data (Iakovidis, 1998). This concern is raised because EHR data is typically composed from different healthcare providers' Electronic Medical Record systems, from paper-based medical reports, and from referrals that patients get from those healthcare providers who do not have an EMR system or an electronic connection with the EHR system. Furthermore, by using an EHR system, a medical practitioner will thus be exposed to historical medical data with varying levels of reliability; the data might originate from a healthcare organization that does not satisfy patient safety requirements, e.g. one which is known to habitually enter inaccurate or incomplete data, or be entered by a medical practitioner who fails to satisfy medical guidelines, e.g. someone who is known to violate medical procedures. As a consequence, the trustworthiness of EHR data depends on the trustworthiness of its sources.

In general, in order to measure the trustworthiness of an agent, reputation systems (Xiong & Liu, 2003, 2004) provide an accumulative trustworthiness measure of an agent where all past experiences and/or feedback about the agent are combined. Most reputation systems are built to assess the trustworthiness of an agent at the present time. In other words they predict the expected future behavior of an agent based on its current trustworthiness. However, they do not provide a way to assess an agent's trustworthiness at a particular time in the past. Evaluating the trustworthiness of past data entries is crucial in the healthcare domain because an EHR combines *historical* medical data.

To illustrate this requirement, consider the following example. Assume that in year 2009 EHR system *A* received two medical reports, Patient *Y*'s diagnosis and Patient *Z*'s prescription, that were created by Dr *X* in 2000 and 2005 respectively. The EHR system maintains a database where it stores its observed experiences with external agents. It uses an eBay-like (Schneider et al., 2000) reputation system (though this is not an appropriate mechanism as we will see in the following section) in which it records the number of observed positive and negative experiences with an agent per annum and uses this to calculate a cumulative trust measure (Figure 1). In this case, these positive and negative experiences are generated from previously evaluated medical entries that were created by Dr *X*. Correct diagnoses and accurately following medical procedures are examples of positive experiences whereas misdiagnoses, incomplete or careless data entry, and failure to follow medical procedures are negative events. Figure 2 represents the observed trustworthiness of Dr *X* that EHR system *A* maintains over time. Now, let's see how EHR system *A* will evaluate the trustworthiness of the two received medical reports.

In current reputation systems, the calculated trustworthiness value for Dr *X* in the year 2009, i.e. 0.48, will be used as the trustworthiness of the two medical records, however this is an inaccurate measure because it represents Dr *X*'s expected *future* behavior instead of his behavior at the time the records were created. From Figure 2, we notice that Patient *Y*'s diagnosis was created at a time period when Dr *X* was evaluated to be trustworthy, whereas Patient *Z*'s prescription was written during a time period when Dr *X* was believed to be untrustworthy. Therefore, assigning the trustworthiness value that is calculated in year 2009 to these two medical records is inappropriate due to the fact that trustworthiness is a dy-