TITRE : Lowering the medical data production load in hospitals by challenging the institutional demand

Introduction

The data production load on healthcare professionnals has become a sensitive issue in France. The non-stop increasing and in-depth variable demand is one of main cause of this increase. In response to this phenomenon, the *nouveaux recueils* program has been rolled out at the Technical Agency for Information on Hospital Care (ATIH), with the aim of automating the collection and transmission of hospital health data. The project is structured around eleven major projects, sponsored by the French health ministry. The program has been introduced into the short-term hospitals financing reform project.

In particular, the program aims at conceiving and deploying data management instruments for a foreseen healthcare data governance function. To do so, the ATIH has launched a healthcare data demand mapping project. The goal is to list all the medical information data requested by national healthcare agencies and characterize the datasets and variables.

Méthod

Two main tables have been built. The first one aims at listing the datasets requested by the institutions. The second one is a list of every variable requested for each dataset. We identified the volume of variables and datasets demanded, tested whether there are redundant demands and evaluated the volume of data produced directly from patients records. From this first word, we conceived a data production load evaluation method using the Manhattan distance tool. The calculation is based on (1) the distance between the information source software and the data input one, and (2) the volume of necessary operations to transform the medical information in the structured requested data. The database has been gathered into a PowerBI file, which has allowed the ATIH to conceive and build primary scorecards and data demand evaluation tools.
The map and scorecards are still on the build and is on constant evolution and a maintenance process has been introduced within the ATIH.

Résults

To date, 141 datasets and 1968 variables have been identified, which are not exhaustive. For each of these variables, several qualifiers have been identified: requesting institution, healthcare field, scope, level of the variable, frequency of collection, etc. Most datasets are composed of less than 40 variables.
The ATIH sought to qualify the main software source of the variable, with many questions raised as sources are very heterogeneous, due to the degree of complexity of hospital information systems. We also tried to identifiy the critical variables, meaning the variables that are necessary for achieving the goal of dataset (financing, hospitals quality certification mainly). Among the identified variables, 741 are considered critical for the achievement of the dataset objective.

Discussion/conclusion

The process of listing all datasets and variables implies many unanticipated difficulties such as dealing with the heterogeneity of the data management performance in the requesting institutions and identifying the critical variables in each datasets. These are technical issues that imply coordination and transaction costs. But the implementation of a global data demand management approach remain the main issue of the project.