

A federated multilevel data infrastructure to enable interoperability among clinical data of different dialysis facilities

[Domenico Vito]

Abstract—Clinical data has opened a great deal of possibility for modern medicine. The benefits could be more if data are shared across health centers. To share data interoperable platforms are needed. Interoperability stands in the ability of different information system to exchange data. It can be foster at different levels, starting from simple technical one, related to the connection of the systems to the practical, by which data can be finally shared for elaboration purposes by different users. Today interoperability is highly requested in several departments of healthcare organizations. This is particularly the case of the dialysis facilities, where the strong lack of standards reflects in a huge gap in having interoperable systems. This work present an effort to overcome the problems of interoperability of data related to dialysis units. A Federated Database System (FDBS) approach to construct a common data repository. The repository has been built by the Dialysis Data Infrastructure (DDI), a unique multilevel standardized data structure supported by the Dialysis MATlib (DM), an embedded Matlab® library, that's able to convert, harmonize and query the raw data coming from the dialysis treatment units into a common interoperable format. The DDI and DM has been developed in the context of the Italy-Switzerland cooperation project INTERREG DialysisIS, and currently contains 846 dialysis session recorded referred to 130 patients. The complete dataset is actually used both by clinicians and researches for data analysis and research inquiries. Even if is at first stages of development the system represents a promising step toward the interoperable use of data in dialysis facilities

Keywords—Dialysis, interoperability, federated database, data mining

I. Introduction

The use of medical data has introduced a great innovation in clinical practice. The possibility to gather and analyze accurate information at the right time from different sources as public healthcare information systems opened the possibility to better orientate correct clinical assessments and related decisions. Furthermore, if information is shared across laboratories, and hospitals, patients could benefit of the same focused treatments even in different clinical centers. Data sharing can also increase public trust in clinical trials and conclusions

derived from them by lending transparency to the clinical research process [2]. Data from medical sources are voluminous, but they come from many different sources, not all commensurate structure or quality [3]. While the data gathering at the hospital level is highly keen to provide detailed information about clinical parameters, very often this effort have the intrinsic limit of not being standardized among the different health care providers. The lack of standardization represents a great barrier to share data and this problem is very common in multicentric initiatives where different health information systems are involved. Each center in fact, could have various licensed software to produce clinical data and export it. The export formats should be different, and thus the final data cannot be easily merged in unique platforms. The integration of different data sources in common structures is highly connected to the concept of interoperability. Interoperability can be defined as "the ability of different information technology systems and software applications to communicate, exchange data, and use the information that has been exchanged" [1]. Interoperability allows the different health information systems to work together and could bring several benefits to effective delivery of healthcare, individuals and communities. Through health care information exchange and interoperability indeed, clinicians everywhere can have a longitudinal medical record with full information about each patient. Conversely patient will have better information about their health status across organizational boundaries [4]. The need for interoperating systems is evident in several department of healthcare organizations. This is particularly the case of the dialysis facilities, where a huge gap on interoperable information systems is still present compared with the other medical fields.

The different dialysis machine vendors in fact developed software for physician offices, and even more for hospitals often so customized that did not allow data interexchange between different clinical centres [5,6]. In this work we present an effort to overcome the problems of interoperability of data related to dialysis units. Our solution proposes a Federated Database System (FDBS) approach [8] implemented on the Dialysis Data Infrastructure (DDI), a unique multilevel standardized data structure supported by the Dialysis MATlib (DM). The DM is an embedded Matlab® library, that's able to convert, harmonize and query the raw data coming from the dialysis treatment units into a common interoperable format.

The DDI and DM has been developed in the context of the Italy-Switzerland cooperation project INTERREG DialysisIS, that involves two reseach units of Politecnico di Milano (the Life Support System Unit and the Department of Management), and 4 dialysis units between Italy and

Domenico Vito
Politecnico di Milano, P.zza Leonardo Da Vinci 32
Italy

Switzerland namely Ente Ospedaliero Cantonale di Lugano (EOC), Switzerland, S.C. Nefrologia e Dialisi dell'A.O. della Provincia di Lecco, Italy Servizio di Nefrologia e Dialisi dell'A.O. di Como, Italy and Servizio di Nefrologia e Dialisi dell'A.O. di Varese, Italy. The project is focused on the customization and standardization of dialysis treatment, in order to improve the patient's medical and social recovery. On pursuing this goal, the project provides the construction of a common web platform to share clinical protocols, patient data and results among the involved clinical centers.

The platform makes the basis for a collaborative effort among clinicians and dialysis units in order to grant the possibility to the patient to perform dialysis therapies in a personalized way, even in different geographically distant places. The formatter will need to create these components, incorporating the applicable criteria that follow.

A. *The problem of data integration*

The creation of a shared platform between different health centers is a common problem in clinical data management. Most of the times, this problem is highly related to the lack data interoperability. Standing to the classification of Wang et al.[7], data interoperability could be implemented at different levels. The basic ones are the technical and syntactic interoperabilities. The first is related to the physical connection and network connectivity among different data sources. The second indeed defines the structure or format of data exchange. If technical interoperability can be usually considered to be covered by the presence of a diffuse internet connection, the syntactical one is several times not granted on health data. As the clinic information systems produce data in different formats and extension, they have to be converted into a common standard format before being integrated in unique common structures. An effective shared data platform should definitively grant higher levels of interoperability. In particular these are the semantic interoperability, by which a common meaning of the data is defined and the pragmatic interoperability that allows a common utilization of the exchanged information by different users, with an high degree of consistency and easiness in updating. Such simple considerations show how the process of data integration from different data sources is not trivial and needs a conceptual model to be properly addressed. In computer science a conceptual model can be defined as “the abstract and simplified representation of a system for some specific purpose by languages, figures, tables, or other suitable artefacts”[7]. The conceptual model that has been adopted to deal with data integration in the Dialysis Project has been the Federated Database System [8]. The term Federated Database System (FDBS) was coined by Hammer and McLeod (1979)[9] and Heimbigner and McLeod (1985)[10]. It stands for a collection of cooperating but autonomous component database systems (DBSs)[8]. Since its introduction, the term has been associated to several different DBS architectures. A key characteristic of a federation, however, is the cooperation among independent systems. with controlled and sometimes limited integration of autonomous DBSs. Standing to its description it comes intuitive how the FDBS logic can be used to define the structure of a common repository for the data

recorded in different dialysis units. These data have different and peculiar features and act as independent datapools. Their integration into a common storage infrastructure should be managed into different steps.

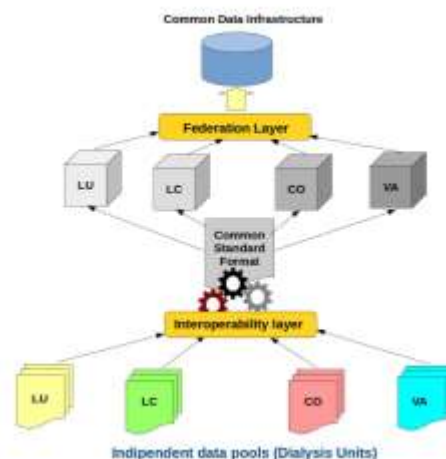


Figure 1. Common Storage infrastructure in FDBS logic

First of all, an interoperability layer should work to convert into the *common standard format* all the not homogeneous data coming from different source. As a result of this process, a converted database, associated to each dialysis unit is obtained. Secondly, as each centre has its own database, a federation mechanism should acts to merge all the data into a common data infrastructure. Figure 1 resume in a figure the concept of the common storage infrastructure in FDBS logic.

II. Materials and methods

A. *Data Collection protocol*

The shared platform has been based on data referred to standard hemodialysis sessions and hemodiafiltration. The data were collected either from monitor sensors (clinical prescription, machine's sensors data, bio-impedanzimeter (hydration status), hemogasanalyzer (blood electrolytes and main catabolite haematochemical data), and hospital information systems. Clinical history and registry data have also been recorded for each subject. Blood pressure measurement has been performed during the treatment by a pressure sleeve on the right arm of the patients. Enrolled patients were higher than 18, undergoes dialysis therapy since at least 3 months, and have an artero-venous fistula as vascular access (AVF).

B. *FDBS logic into practice: DB-Dialysis and Dialysis MATlib*

The FDBS has been traduced into practice by the implementation of the Dialysis Data Infrastructure (DDI) as the common data repository for all the clinical centres.

The interoperability and federation mechanisms among data have been realized through the Dialysis MATlib (DM), a group of Matlab® script and m-files that work on the DDI in order to convert, harmonize, query, merge and mine the raw data coming from the dialysis units. Figure 2 resumes the global architecture of DDI and its interaction with the DM.

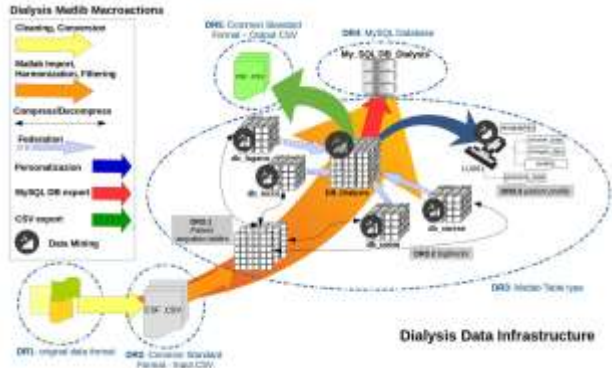


Figure 2. General Framework of the Dialysis Data Infrastructure and Dialysis MATlib

The Dialysis Data Infrastructure has been designed to be a flexible tool for both clinician and researchers to share the clinical data of the Dialysis project. Accounting for their peculiarities, it furnishes the possibility to handle the data coming from the specific dialysis unit separately and in a customized way. Even though, these data can be finally gathered into a common repository, obtaining a federation structure. The DDI also allows the extraction of patient's specific data in order to perform personalized analysis on the data. As depicted in Figure 2, the DDI provides 5 levels of data representation (DR1, DR2, DR3, DR4, DR5). The first two levels are directly related to the primary sources of data coming from the clinics. In particular level DR1, represent the raw data files directly transmitted by the dialysis units. These data are usually exported from proprietary software's or even reported manually by the dialysis unit operators: so they are saved in files that are highly uneven in formats. In the physical implementation of DDI the primary files of DR1 are stored and constantly uploaded on an online ftp repository. The repository file system is organized reflecting the organization in centres as in Figure 3.

Each dialysis unit is associated to a unique folder in the file system.

Each one of these folders contains also a CSF-CVS sub-folder, that's the site for the *Level DR2 data* files. *Level DR2* files are obtained by the conversion of original raw files into a common standard format (CSF).

In CSF data are collected in tables, which will constitute a *common database schema* (CDS) for the upper levels data structures.

In the CDS, each table groups information depends on the source of acquisition, as described in Table 1.

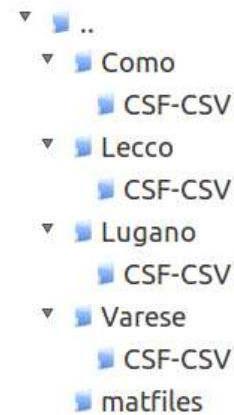


Figure 3. General ftp repository file-system

TABLE I. CSF TABLES SCHEMA

DR2 Table Name	Acquisition Source	Description
Anagraphics	HIS*, Clinical Records	Anagraphic/Anamnesic data
Session_info	HIS, Clinical Records	Prescriptions on the treatment within the session
Hemogas_data	Hemogas Analyzer, Hospital Laboratory	On-line hematochemical data
pressure_data	Pressure Sleeve	Pressure Sleeve on-line acquisitions
Dialysis_machine_data	Dialysis Machine	Dialysis Machine on-line acquisitions
bcm_data	Bioimpedenziometry	Patient hydration Status got by bioimpedenziometry
dialisate_data	Clinical Records	Prescriptions on Dialisate Composition
Dbag_data	Dialisate bags datasheets	Dialisate bags initial composition
Events_register	Clinical Records	Sessions event register

*HIS= Hospital Information System

after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. The conversion between DR1 and DR2 represents a first step of data interoperability.

In particular, interoperability is realized here on the existences of a common acquisition protocol between the dialysis units. The common acquisition protocol drove the logic on how the CDS has been designed.

In level DR2 data are saved in the specific folder of the ftp-repository as CSF-CSV (Common Standard Format CSV) input file, that are .CSV files in row-column format with point-comma column separator.

Typically each CSF-CSV corresponds to one of the tables described in Table 1.

Each table collects all the acquisitions referred to a specific patient.

In order to grant the referential integrity, tables are filled following some few basic principles:

- *each patient* is identified by a unique patient ID of 5 alphanumeric characters: the first 2 are a couple of letters that identifies the clinical centre, the last 3 are digits representing a progressive number
- *each dialysis session*, for each patient is identified by a progressive number

The CSF-CSV files are named in patient ID-oriented way as follow:

<nametable_patientID>.csv

(1)

The CSF-CSV input files constitute a consistency layer for all the DDI. The entire further layer, in fact is implemented on the basis of the CSF-CSV. For example data in level DR3 were obtained importing, filtering and harmonizing the CSF-CSV into Matlab® (Figure 1).

In particular, *level DR3* reports data in the Matlab® workspace as table type variables, available since the R2013b version **Error! Reference source not found.** of the software. They are usually used to store heterogeneous types of data coming from a text file or a spreadsheet in a unique tabular format with rows and columns.

In *level DR3* table type variables has been used to reproduce the structure of the CDS in the Matlab® environment.

In order to easily handle data analysis at different levels of aggregation, 3 types of DR3 representation has been implemented. The first one is the DR3.1. Here data are represented in *patient single tables*. The *patient single table* follow the CDS structure and will allow iterative operations, referring to the data of a specific acquisition source at the patient's level.

Since they represent the more proximal import of the CSF-CSV input files into Matlab® workspace, the patient single tables inherit the same patientID-oriented organization and naming of level DR2.

As depicted in Figure 1, patient single tables can be gathered in huger structures by the compress Dialysis MATlib macro-action. These huge structures are the DR3.2 type, also called *big-blocks*. DR3.2 provides each single dialysis unit of a unique database obtained by the aggregation of single tables of all the patients treated in the same unit (db_lecco,db_lugano,db_como,db_varese). The aggregation allows to maintain the CDS but not the patientID organization. Nevertheless the operation can be inverted by the decompress Dialysis MATlib macro-action. The DR3.2 data structures can be specifically used for analysis properly targeted to the clinical centres.

Finally, the single dialysis unit databases can federate into a main central data-storage (DB_Dialysis) that collects all the acquired data. DB_Dialysis can be used for general undifferentiated inquiries.

For high personalized purposes indeed, there is the DR3.3 type, or *patient profile*: here data are represented in a compact patient data struct, that contains all the tables all the CSD in one Matlab® variable. The patient ID identifies each patient data struct, and each table will be accessible by the dot notation. DR3.3 format will facilitate personalization allowing an easy querying of patient specific data. All the data at level DR3 are saved in .mat files. The files are stored in the common ftp repository in a proper folder (Figure 3).

The ultimate levels of data representation in the DDI are the DR4 and DR5: they are both obtained from the DB_Dialysis by dedicated export actions. In particular DB4 level stores data in the MySQL_DB_Dialysis, a MySQL® database, underpins a web platform. The DB5 indeed allows to having all the DB_Dialysis tables as .CSV output files in the common standard format with dot-point as column separator.

C. *Dialysis MATlib*

The Dialysis MATlib is a customized library, implemented in Matlab® code, to manage data in the framework of the Dialysis Data Infrastructure. It achieves the interoperability of raw data coming from the clinical units, allowing converting, harmonizing and querying them in a versatile way. Specifically Dialysis MATlib has been created with the Extract, Transform, and Load (ETL) methodology logic.

The ETL methodology is frequently used in database operations, mainly in data warehousing and refers to three simple processes [19]:

-*extract*, data from heterogeneous multiple data sources

-*transform*, data into a proper storage format

-*load*, the transformed data into a target database with a defined structure

Dialysis MATlib realizes these processes exploiting the features of the relational tables and their conversion into the Matlab table type variables. The table data type indeed, allows storing variables that are either string or numbers in a unique rows and column-oriented structure. Through a dot notation is therefore possible to easily access to table rows and column, making data manipulation more simple and direct. Furthermore using the vector and matrix transformations, that are typical of the Matlab® environment, errors correction, standardization and data format conversion can be massively performed using just coding language instead of complex SQL queries. Dialysis MATlib functions are divided into 8 groups that correspond to the macro-actions on the DDI showed in Figure 1. The description of the groups is reported in Table 2

The Macroaction of the Dialysis MATlib covers all the data workflow within the DDI: they practically act as the layers of the FDBS logic described in Figure 1. More specifically, macroactions 0 and 1 can be associated to the interoperability layer. Interoperability is here obtained directly by the Matlab® code that realizes differentiated operations depending on the data input received in order to produce output tables in a standard format.

Macroactions 2 and 3 indeed act as federation layers. They coordinate the merging operations of dialysis unit's databases into a common data repository, the DB_Dialysis.

TABLE II. DIALYSIS MATLIB FUNCTIONS GROUPED BY MACROACTIONS

Macroaction No.	Name	Descriptions
Anagraphics	HIS*, Clinical Records	Anagraphic/Anamnesic data
Session_info	HIS, Clinical Records	Prescriptions on the treatment within the session
0	Cleaning, Conversion	Import data from CSV (DR1, DR2) to table variables.
1	Matlab import, Harmonization, Filtering	Conversion in DR3 sub formats
2	Compress/Decompress	Correction of transcription errors, customization
3	Federation	Filters the main dataset from spurious and not consistent data
4	Personalization	Statistical Analysis Feature extraction
5	MySQL Export	Export data into a .CSV file to be filled into the Dialysis MySQL-DB (DR4)
6	CSV Export	Export data into a generic .CSV file
7	Data Mining	Allows the analysis of data through mining techniques

III. Results

The implemented Dialysis MATlib is able to manage the collection of data coming from 4 different dialysis units. This means that the system can coordinate 4 different hospital data pools. As the patient's data are coming from different devices, the Dialysis MATlib is capable to interface several heterogeneous data sources. In particular the library can merge in a unique repository the data furnished in proprietary formats of 2 of the main dialysis machine vendors available on the market and installed in the involved dialysis units. Furthermore it handles to convert in a defined common standard, the data produced by 4 different types of hemogas analyzer and 4 different professional bioimpedentiometers. Finally, it can gather anamnesis and anagraphic data coming from 4 different Hospital Information Systems. All the data are stored into the framework of the Dialysis Data Infrastructure, which currently contains the acquisition of 846 dialysis sessions performed on a total of 130 patients. The number of sessions and patients by centers are reported in Table 3.

TABLE III. DISTRIBUTION OF SESSION RECORDED PER CLINICAL CENTERS

Clinical Center.	No of Patients	No of dialysis session recorded
Ente Ospedaliero Cantonale di Lugano	20	.216
S.C. Nefrologia e Dialisi dell'A.O. della Provincia di Lecco	50	300
Servizio di Nefrologia e Dialisi dell'A.O. di Como	30	180
Servizio di Nefrologia e Dialisi dell'A.O. di Varese, Italy.	30	180

The complete dataset is actually used both by clinicians and researches for data analysis and research inquiries.

A statistical study for clinical indexes extraction has been already conducted **Error! Reference source not found.**

It has exploited the FDBS logic of the DDI, being performed in differentiated way by dialysis units. The study is still ongoing and the DDI can be used to produce further works.

IV. Discussion

The DDI and the Dialysis MATlib constitute a versatile tool to manage the clinical data of the Dialysis Project. The combination of these two elements has been used to overwhelm the hard data interoperability problems persist among the data collections of the involved dialysis units. Currently, the whole system can perform 4 levels of interoperability by the definition of Wang et al.[6]. Starting from the practical interoperability given by the diffuse internet connection, the presence of a common data format provides syntactic and semantic interoperability. The Dialysis MATlib indeed sets the features for a common use of the data, enabling practical interoperability. Practical interoperability has been particularly important in the Dialysis Project. Since the study involved different clinical units, it was central to furnish an instrument that should share analysis, mining and data management functionalities among heterogeneous users in different clinical centers. The practical interoperability has been powered by the the federated structure of the DDI.

It allowed the use of data at various levels of aggregation, from center specificity to patient personalization, widening the possibility of data exploitation. Actually, the DDI and Dialysis MATlib are physically based on a central ftp repository that is accessible from remote stations by a controlled accreditation in order to grant data privacy and security. A key element in the implementation of practical interoperability has been the definition of macroactions on the different datapools. The macroactions can be considered the factual implementation the interoperability and the federation layers described in Figure 1. They are performed by functions that has been coded in Matlab®.

It can be surely argued that Matlab® is not the ideal environment to deal with large data repositories and even more that today there are several softwares more specific to support data mining [21]. Furthermore Matlab® is a proprietary software, and need a specific licence to be installed and distributed on several computers: this element of course is a barrier toward the portability of the shared platform. Nevertheless, this decision has been done to exploit the potentials of Matlab® to treat heterogeneous data in a vector-matrix logic. Once data are translated into workspace variables, they will be available indeed to be elaborated by Matlab® toolboxes. Currently, Matlab® is providing several toolboxes for large dataset management and exploratory analysis [22] like the Machine Learning, the Database and the SVM Toolbox. As indeed, the Dialysis MATlib provides a standardized input workflow, the DDI can be widened in every moment with new data. This represents a great potential, by which the computational limits stand in the maximum dimension allowed for a Matlab® variable to the RAM memory. The limit depends on the specific machine architecture and it is 2 GB running on 32bit and from 500 to 4 TB on 64bit respectively if the OS is Windows 7 or the desktop version of Windows Server [23]. However, the limit could be outmatched in Matlab® by “MapReduce” and “memmap” functions [23]. These functions are suitable for analyzing large datasets that cannot fit computer's memory. Moreover, MapReduce allows can connect to the well-known big-data platform Hadoop [23]. Even some counter backs could exist, the use of MATLAB® as main developing environment presented several advantage. Since the Dialysis project has the prevision to continue, the choice of the software has been strategic for mining data in the federated structure of the DDI.

v. Conclusions

Interoperability is a great challenge in the spreading field of health data. The challenge is particularly captivating on the dialysis field, where there is still a wide gap in the definition of common exchange protocols and shared format on the produced data. The work has presented an effort to constitute a common data repository between multiple dialysis units. Since each of this unit provides data in peculiar formats, standing as an independent data pool, a shared platform has been proposed, based on a common data format and a federated database logic. The platform has been factually implemented through the Dialysis Data Infrastructure, a flexible data architecture for data provision at different levels of aggregation, and the Dialysis MATlib, an embedded Matlab® library that allows to multiple functionalities for data management and mining. Currently the combination of the DDI and DM gathers data from 4 different dialysis units, which stand in a total of 846 dialysis session recorded. The complete system has been also used for statistical analysis offering the possibility of center specific and patient focused analyses. The platform grants several levels of interoperability, from syntactical to practical [7]. Even if the platform is at an early stage of development and it's not fully portable it has been demonstrated to be a useful tool to sharing clinical and research data.

Acknowledgment

This work was funded by INTERREG IT-CH EU Program: Dialysis (Dialysis therapy between Italy and Switzerland)-ID 33570710

References

- [1] HIMSS “Dictionary of Healthcare Information Technology Terms, Acronyms and Organizations”, 2nd Edition, 2010, Appendix B, p190
- [2] S. Olson, A.S. Downey, “Sharing clinical research data” [Internet]. 2001. Available at: <http://goo.gl/DiTmvv>.
- [3] K. J. Cios and G. William Moore, “Uniqueness of medical data mining,” *Artif. Intell. Med.*, vol. 26, no. 1–2, 2002 pp. 1–24.
- [4] D.J. Brailer. “Interoperability: the key to the future health care system. *Health Aff (Millwood)*”, 2005, Suppl Web Exclusives.
- [5] J.H. Sadler, “Computers in Hemodialysis Facilities, *Hemodialysis Horizon*”, Washington, 2006.
- [6] J.A. Sargent, “Identifying the Value of Computers in Dialysis: Business Applications”, *Nephrology News & Issues* June, 2001, pp 19-23.
- [7] W. Wang W, A. Tolk A, W. Wang, “The Levels of Conceptual Interoperability Model: Applying Systems Engineering Principles to M&S”, *Proc 2009 Spring Simul Multiconference*, 2009, p. 9.
- [8] A.P. Sheth, J. Larson, “Federated database systems for managing distributed, heterogeneous, and autonomous databases”, *ACM Comput Surv*, 22(3),1990, pp. 183–236.
- [9] M. Hammer, D. McLeod, “On database management system architecture”, *Tech. Rep.* 1979.
- [10] D.Heimbigner ,D. McLeod, “A federated architecture for information management,” *ACM Trans. Inf. Syst.*, vol. 3, no. 3,1985, pp. 253–278,
- [11] E. Morris, L. Levine, C. Meyers, P. Place, D. Plakosh, “System of systems interoperability (SOSI), final report”. *Software Engineering Institute, Carnegie Mellon University: P Pitts- burgh*, 2004.
- [12] D. Vito, G. Casagrande, G. Cappoli, C. Bianchi, G. Pontoriero, C. Schoenholzer C., “A predictive index of intra-dialysis IDH A statistical clinical data mining approach”, 2015, pp 109–13.
- [13] B. Blobel, “Architectural Approach to eHealth for Enabling Paradigm Changes in Health”, *Methods Inf Med.*, 49(2), 2010 ,pp. 123–134.
- [14] J. Kalina, “Robustní analýza obrazu obliceje pro genetické aplikace. *EJBI* “,6(2),2010,cs95–cs102. Available from: <http://www.ejbi.eu/articles/201012/47/2.html>
- [15] I. Drăgan, R.D. Zota, “Big data Migration for Telecom Business Support Systems”, 2015, pp. 126–9.
- [16] Intel whitepaper, “Extract, Transform, and Load Big Data with Apache Hadoop”, 2021.
- [17] Mathworks, “Global Optimization Toolbox: User's Guide (r2011b). Retrieved November 10, 2011.
- [18] E. Naka and A. Bitri, “A Survey of Classification, Clustering and Association Techniques in some Data Mining Software”, 2014, pp. 27–38, 2014.
- [19] W. L. Martinez and A. R. Martinez, “Exploratory Data Analysis with MATLAB”, 2005.
- [20] VV, AA., “Large Files and Big Data [Internet].”, Available from: <http://it.mathworks.com/help/matlab/large-files-and-big-data.html>

About Author (s):



Domenico Vito, PhD. Specialized in bioinformatics especially in machine learning and statistical methods. Since 2014 he is active in EU project regarding the OneHealth approach and the connection with different partners