

A Digital Twin Platform for Diagnostics and Rehabilitation of Multiple Sclerosis

Dessislava Petrova-Antonova[1(✉)], Ivaylo Spasov[2], Iva Krasteva[1], Ilina Manova[2], and Sylvia Ilieva[1]

[1] GATE Institute, Sofia University, Sofia, Bulgaria
{d.petrova, sylvia}@fmi.uni-sofia.bg,
iva.krasteva@gmail.com
[2] Rila Solutions, Sofia, Bulgaria
{ispasov, ilinam}@rila.bg

Abstract. 30 million people across Europe are affected approximately from rare diseases. The brain disorders (including, but not limited to those affecting mental health) remain a major challenge. The understanding of mental disorders' determinants and causes, processes and impacts is a key to their prevention, early detection and treatment as well as factor for good health and well-being. In order to improve the health and disease understanding, a close linkage between fundamental, clinical, epidemiological and socio-economic research is required. Effective sharing of data, standardized data processing and the linkage of such data with large-scale cohort studies is a prerequisite for translation of research findings into the clinic. In this context, this paper proposes a platform for exploration of behavioral changes in patients with proven cognitive disorders with a focus on Multiple Sclerosis. It adopts the concept of a digital twin by applying the Big Data and Artificial Intelligence technologies to allow for deep analysis of medical data to estimate human health status, accurate diagnosis and adequate treatment of patients. The platform has two main components. The first component, provides functionality for diagnostics and rehabilitation of Multiple Sclerosis and acts as a main provider of data for the second component. The second component is an advanced analytical application, which provides services for data aggregation, enrichment, analysis and visualization that will be used to produce a new knowledge and support decision in each instance of the transactional component.

Keywords: Big Data · Cognitive assessment · Digital twin · Multiple Sclerosis

1 Introduction

The research of cognitive diseases in patients with a number of neurological disorders is gaining increasing social significance. The affected persons are often of working age and timely diagnosis is an important factor for proper treatment and prevention. The most of the approaches for testing of patients are based on paper tests and involved a highly specialized medical staff. The processing and analysis of paper results is difficult and time consuming. In practice, each therapist works in isolation with a particular group of patients and limited healthy controls for comparison. These circumstances do

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12249, pp. 503–518, 2020.
https://doi.org/10.1007/978-3-030-58799-4_37

not allow accumulation of data with different format and origin, as well as searching for certain patterns and dependencies in cognitive deficits. Although the development of automated versions of the paper tests is in place, they primarily serve as storages of results without further analysis of disorders. Some of them are limited to statistical processing of results without prediction and prescription possibilities and others use machine learning techniques to predict probable outcomes only for individual patients.

At the same time, the rapid growth of data, leading to the so-called Big Data phenomenon, affects all domains of everyday life, ranging from user-generated content of around 2.5 quintillion bytes every day [1] to applications in healthcare [2], education [3], knowledge sharing, and others. In all domains, data is the key, not only by adding value and increasing the efficiency of existing solutions but also by opening new opportunities and facilitating advanced functionalities supporting timely and informed decisions. Given the data value for applications in different domains there is an urgent need for solid end-to-end, data-driven and data-oriented holistic platforms that provide a set of mechanisms for runtime adaptations across the complete path of data and lifecycle of domain services. Such platforms adopt the concept of digital twin – a digital profile of the physical world that helps to optimize its performance. The digital twin is based on massive, cumulative, real-time, real-world data across an array of dimensions. The data is used for building digital models based on advanced artificial intelligence and machine learning algorithms that may provide important insights, leading to actions in the physical world such as a change in human health and well-being. The Big Data and Artificial Intelligence technologies applied for digital twinning of patients allow deep analysis of medical data to estimate human health status, accurate diagnosis and adequate treatment leading to so called precision medicine.

This paper proposes an architecture of a digital twin platform for exploration of behavioral changes in patients with proven cognitive disorders with a focus on Multiple Sclerosis (MS). The platform identifies the causes and trends that lead to a deepening of the long-term cognitive decline. It allows for assessment of cognitive status in a timely manner and suggest appropriate preventive as well as rehabilitation actions of cognitive disorders. The main contribution of the platform are as follows:

- Understanding the determinants of MS (including nutrition, physical activity and gender, and environmental, socio-economic, occupational and climate-related factors) through digital twin modelling of patients;
- Better understanding the cognitive diseases and thus improving diagnosis and prognosis for patients with MS through application of digital twin models;
- Improving the collection, enrichment and use of health data related to the cognitive diseases;
- Adoption of advanced machine learning and artificial intelligent methods for precise diagnosis and rehabilitation of MS;
- Improving both scientific and medicine tools through development of a technological platform for automation of digital twin modelling of patients.

The rest of the paper is organized as follows. Section 2 briefly describes the current state of the research on the problem area. Section 3 summarizes the requirements to the platform, while Sect. 4 is devoted on its architecture. Section 5 describes the automations of the tests for diagnosis and rehabilitation of MS. Finally, Sect. 6 concludes the paper and gives directions for future work.

2 State of the Art

A review of the current state of the research on the problem area was performed, covering the computer implementations for evaluation of patients with mild and moderate cognitive disorders such as MS, Alzheimer's disease, early dementia, mild cognitive impairment, etc. The following state-of-the art articles are explored:

- A systematic review of the diagnostic accuracy of automated tests for cognitive impairment [4] – The systematic survey includes studies from January 2005 to August 2015 providing evaluation of diagnostic accuracy of automated tests for the assessment of mild cognitive impairment or early dementia. Articles in English are included. The review examines 11 tools for automated testing of cognitive disorders.
- Status of computerized cognitive testing in aging: a systematic review [5] – The survey includes studies published by 2007 that assess or detect changes in age-related cognitive status, early dementia, and mild cognitive impairment via computer-based tests. Each automated test is evaluated against availability, validity and reliability of tests, range and use. The review includes 11 tests.
- Computerized cognitive testing for older adults: a review [6] – The review examines 17 tests for assessment and screening of cognitive status with publications available by 2012. A comparative analysis of tests according to the hardware, mode of execution and administration, runtime and supported cognitive domains has been performed. The review also examines the psychometric properties of the tests and shows that the most of the tests are well documented in terms of validity and reliability, but the other properties vary.
- Computerized cognitive testing for patients with multiple sclerosis [7] – The review focuses on the available automated tests of multiple sclerosis. The following tests are included: Automated Neuropsychology Assessment Matrix (ANAM), the Mind streams Computerized Cognitive Battery (MCCB), the Amsterdam Neuropsychological Tasks (ANT), the Cognitive Stability Index (CSI), the Cognitive Drug Research (CDR) battery.

The selection criteria for the software solutions included in the review are as follows:

- Provide implementation of tests in cognitive domains affected by MS;
- Validated for use in patients with MS or used in mild cognitive impairment such as: LCU, early dementia, Alzheimer's disease, Parkinson's disease, brain injury, etc.;
- Solution documentation is available at the time of review;
- The implementation of the solution and the tests are in English.

The cognitive domains affected in patients with multiple sclerosis are: Information processing, Visual-spatial memory, Auditory-verbal memory, Expressive language, Executive function, and Visual-spatial information processing.

Tables 1 and 2 present a comparative analysis of the software solutions considered for assessing cognitive status by the following characteristics: application area (1), type of device (2), way of interaction with the device (3) patient's performance mode (4), number of tests and batteries (5), additional factors (6), output data (7), maintained cognitive domains (8), and diseases (9).

Table 1. Software solutions for assessing cognitive status (A).

	ANAM Tests	BrainCare	CANS-MCI	CANTAB Insight
1	Research software, Medical software	Research software, Medical software	Medical software	Medical software
2	PC, Notebook	Device with a browser	Windows based	iPad
3	Mouse/Keyboard	Mouse/Keyboard	Touchscreen	Touchscreen
4	Alone in a controlled environment; Preliminary training is required	Alone with technical assistant; Allowed in a home environment	Alone – technical assistant starts the test execution	Alone in a controlled environment; Audio instructions available;
5	22 tests/3+ standard barriers; battery configuration is supported	10 tests for mild cognitive impairment	8 tests/1 battery	3 tests/1 battery
6	Part of the tests: Demographic/History Module, Mood Scale, etc.	2 scales – for mood and for nervousness and anxiety	Definition of questions for determination of factors that influence the test execution	Mood assessment (the Geriatric Depression Scale or GDS-15)
7	Automated generation of report for the specialist; Data access for exploration and download	Automated generation of report for the patient and the specialist	A neuropsychologist from the company prepares a report	Automated assessment – the norm is calculated based on the gender, age and education of the patient
8	Attention, Concentration, Response Time, Memory, Information Processing Speed, Decision Making, and Execute Function	Attention, Visual-spatial information processing, Execute function, Memory, Auditory-verbal memory	Memory, Execute Function and Information Processing	Executive function, Processing speed, Attention, Working memory,Episodic memory
9	8 diseases	4 diseases	4 diseases	Assessment of cognitive status without concrete diseases

The Automated Neuropsychological Assessment Metrics (ANAM) platform includes several tools, based on 22 tests and behavioral questionnaires for cognitive assessment [10]. Several standard batteries, including predefined tests are supported such as General Neuropsychological Screening battery (ANAM GNS), Traumatic Brain Injury (TBI) и ANAM-MS. New test batteries for specific purposes can be created. The ANAM questionnaires assess additional factors such as demographics, mood scale, etc.

Table 2. Software solutions for assessing cognitive status (B)

	CANTAB connect research	Cognigram - cogstate	Cogstate research	CAMCI
1	Research software	Medical software	Medical software	Research software, Medical software
2	Touchscreen device; Device with a browser	Touchscreen device; PC	Touchscreen device; PC	PC, Notebook
3	Touchscreen/mouse/Keyboard	Touchscreen/mouse	Touchscreen/mouse	Mouse/Keyboard
4	Technical assistant is required for particular tests; Remote execution is allowed; Audio instructions available	Alone in a clinical or home environment with technical assistant	Technical assistant is required for particular tests	Alone in a controlled environment; Preliminary training is required
5	16 tests/17 standard batteries; battery configuration is supported	4 tests/Cogstate Brief Battery	11 tests/7 standard batteries	22 tests/3+ standard barriers; battery configuration is supported
6	Mood assessment	N/A	N/A	Part of the tests: Demographic/History Module, Mood Scale, etc.
7	Automated assessment – the norm is calculated based on the gender, age and education of the patient	Automated generation of report for the specialist	Automated generation of report for the specialist	Automated generation of report for the specialist; Data exploration and download
8	Attention, Memory, Execution function, Emotional and Social Cognition, Psychomotor speed	Visual-motor control, Associative training, Psychomotor function, Executive function, Attention, Verbal training, Visual training	Visual-motor control, Associative training, Psychomotor function, Execution function, Attention, Verbal training, Visual training, Working memory, Emotion recognition	Attention, Concentration, Response Time, Memory, Information Processing Speed, Decision Making, and Execute Function
9	20+ diseases	40+ diseases	40+ diseases	7 diseases

BrainCare applies a battery of tests, which map patient capabilities and results across seven cognitive areas: Memory, Executive function, Attention, Visual spatial, Verbal function, Problem solving, and Working memory [11]. The tests are divided in two groups – mild tests and moderate-severe tests. The mild tests include: Verbal memory test, Non-verbal memory test, Go-NoGo response inhibition test, Stroop test, Visual spatial processing test, Verbal function test, Staged information processing speed test, Finger tapping test, "Catch" game and Problem-solving test. The moderate-severe tests cover assessment of orientation to time and place, language skills, non-verbal memory, similarities and judgement, reality testing, spatial orientation and execution function (Go-NoGo basic). Unlike the mild tests, only one moderate-severe test (i.e., GoNoGo Basic) is interactive. For all other tests, responses are entered by the test supervisor rather than by the patient.

The Computer-Administered Neuro-Psychological Screen for Mild Cognitive Impairment (CANS-MCI) platform automates 8 tests for assessments of mild cognitive injuries in 3 cognitive domains. The assessment includes measures of free and guided recall, delayed free and guided recognition, primed picture naming, word-to-picture matching, design matching, clock hand placement and the Stroop Test. The CANS-MCI tests are applied to establish baseline and longitudinal measures of cognitive abilities that are relevant to a number of medical conditions and their treatments, such as early onset Alzheimer's disease, drug and alcohol rehabilitation, concussions incurred in sports, cancer treatments (e.g. "chemobrain"), MS, Lupus and Parkinson's [12].

The Cambridge Cognition is a leading company in the field of neurology. CANTAB Insight is an analytical assessment tool to enable quick and accurate measurement of brain function across five cognitive domains, namely executive function, processing speed, attention, working and episodic memory [13]. It automates 3 tests: Paired Associates Learning for assessment of visual memory and new learning, Spatial Working Memory for assessment of retention and manipulation of visuospatial information and Match to Sample Visual Search for assessment of attention and visual searching, with a speed accuracy trade-off. Tests are adaptive, so testing will end once a patient reaches tests' limit for the number of attempts for their age, gender or level of education. CANTAB Insight includes an optional mood assessment (the Geriatric Depression Scale or GDS-15). CANTAB Connect Research is a precise and reliable research software providing sensitive digital measures of cognitive function for all areas of brain research. It covers five key domains: attention, memory, executive function, emotion and social cognitions and psychomotor speed. CANTAB Connect Research supports 16 tests, which can be combined in different batteries, and delivers 17 featured batteries.

Cogstate Cognigram provides a battery with 4 tests (Cogstate Brief Battery) for assessment and monitoring of patient's cognitive state in 4 cognitive domains [14]. The tests are based on a single playing card stimulus, which is presented in the center of the device screen. Cogstate Research delivers 11 tests, which can be combined in different batteries [15]. Each test has been designed and validated to assess specific domains including psychomotor function, attention, memory, executive function, verbal learning and social-emotional cognition. All tests are cultural and language independent and can be performed in a clinical or in a home environment.

Computer Assessment of Memory and Cognitive Impairment (CAMCI)-Research is a customizable battery of computerized tasks to assess cognitive performance [16]. The CAMCI-Research battery includes 9 behavioural tasks. A series of self-report questions are provided to gain information from the patients regarding factors that could affect their performance (e.g., perceived memory loss, alcohol use, depression, anxiety, etc.). The CAMCI-Research software does not offer a medical diagnosis and can be used only for research, investigational or educational purposes.

3 Requirements Specification

The requirements of the platform are closely related to the implementation stages of the digital twin, shown in Fig. 1. The *Create stage* include building of Patient Information Model (PIM) by collecting data from different sources. The data can be classified in two categories: (1) operational data related to results from MS tests; and (2) external data related the patient cognitive status, behavior, etc. The *Interact stage* provides real-time, seamless, bidirectional connectivity between the physical patient and its digital twin (virtual patient) through the platform. The *Aggregate stage* covers data ingestion into a data storage and data pre-processing such as cleaning, consistency checking, linking, etc. The data may be augmented with information for other patients with similar diagnosis. The *Analyze stage* is based on variety models of virtual patient that are built on top of PIM. The main goal of this stage is to produce new knowledge that drive the decision-making process. Artificial intelligent methods and cognitive computing are applicable on that stage. The *Insight stage* visualizes the insights from analytics as 3D views, dashboards and others. It aims to provide evidence about areas for further investigations. The *Decision stage* applies the new knowledge to the therapy of the physical patient in order to produce an impact of the digital twin. The decisions lead to precise diagnosis and timely and adequate rehabilitation.

The operational data will be obtained through automation of the MS tests. Several interviews with two clinical specialists with expertise in MS are conducted to understand the domain and to collect the initial set of requirements for software implementation of the tests for assessment and rehabilitation of MS. As a result, a description of diagnostics and rehabilitation processes of MS is obtained, which shows how the platform will be used by the clinicians and how the patients will interact with the platform. The functionality of the first version of the platform is defined to cover the initial set of requirements, including automation of the following tests:

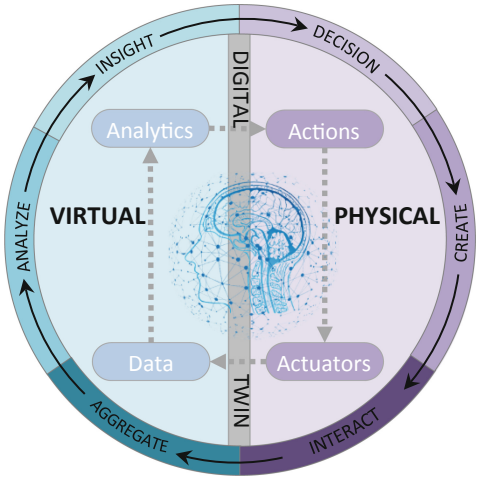

Fig. 1. Digital twin concept.

- Beck Depression Inventory (BDI-II) test;
- BICAMS – Symbol Digit Modalities Test (SDMT), Brief Visuospatial Memory Test (MVMT) and California Verbal Learning Test (CVLT);
- Rehabilitation tests.

The platform has two main components. The first component provides functionality for Create, Interact and Decision stages of the digital twinning. It is developed as a transactional application, called CogniSoft, which has multiple instances in different neurological departments. The second component covers all stages and will be developed as an advanced analytical application using Big Data and Artificial Intelligence technologies. It will provide services for data aggregation, enrichment, analysis and visualization that will be used to produce a new knowledge and support decision in each instance of the transactional component.

4 Platform Architecture and Technologies

As was mentioned so far, the platform consists of two components that will be implemented separately. In the paper, they will be called applications, although each of them can be considered as a stand-alone platform.

4.1 Analytical Application

The architecture of the analytical application is shown in Fig. 2. The analytical application relies on the Hadoop Distributed File System (HDFS). It consists of five layers: (1) Data collection, (2) Data ingestion, (3) Data storage, (4) Data processing and analytics, and (5) Data insight. The primary sources of data are provided by the transactional applications in the neurological departments. The patients' data is enriched with Electronic Health Records (EHRs), open clinical datasets, data form social network and other external applications.

- The data ingestion is implemented by four web services as follows:
- Edge service – ingests large amounts of data that comes from file servers, such as web-logs, DB dumps, etc.
- PULL service – pulls data from external applications and APIs.
- PUSH service – provides a gateway for external applications to push data.
- Direct ingestion service – pulls data directly from databases of the internal transactional applications.

The PULL and PUSH services transmit data to the Apache Kafka cluster, which is a distributed publishing/subscribing message system of one or more brokers [9]. For each existing topic, the brokers manage zero or more partitions. The publisher connects to Kafka cluster and sends request to check which partitions exists for the topic of interest and which nodes are responsible for those partitions. Then the publisher assigns messages to partitions, which in turn are delivered to the corresponding brokers. The subscriber is a set of processes, which cooperates each other and belong to a same consumer group. The consumers in a given group are able to consume only from the partitions that are assigned to that group.

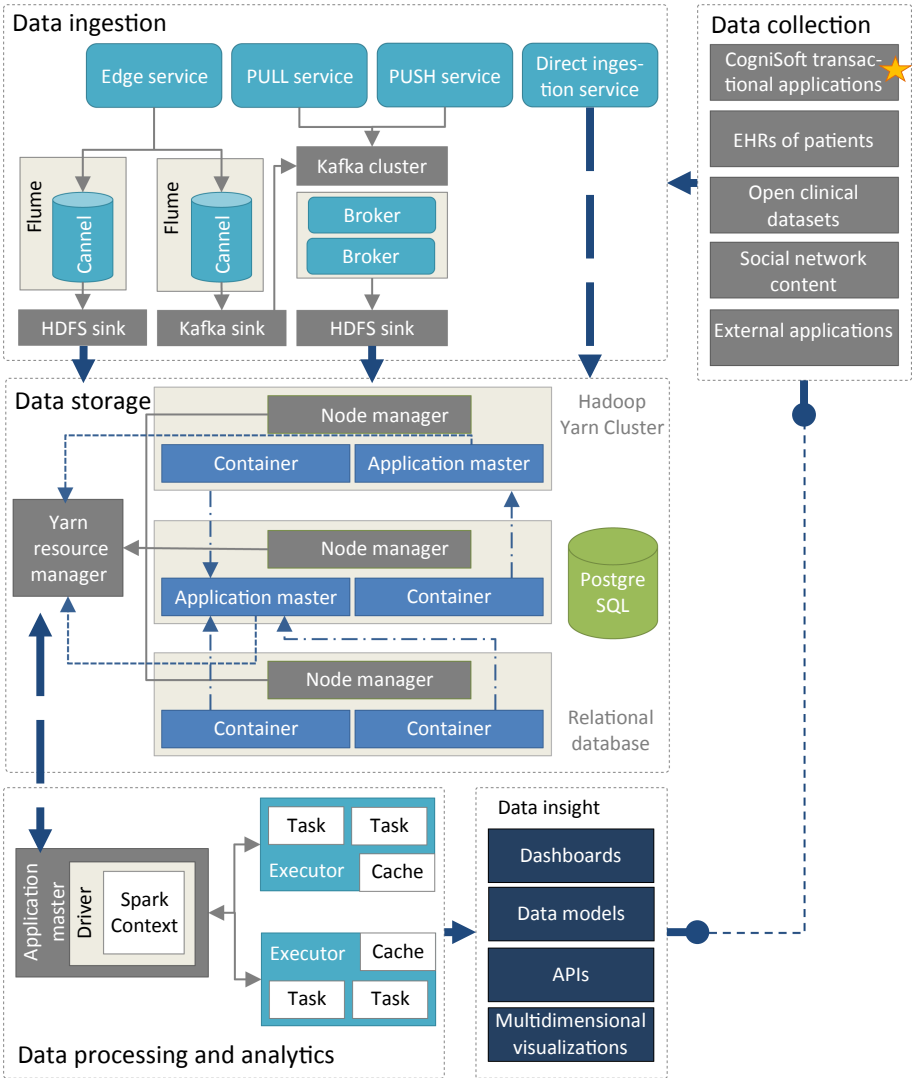

Fig. 2. Architecture of the analytical application.

The Edge service transmits data to the Apache Flume Service, which is a distributed, reliable service for ingestion of large amounts of data in different file system. The Flume Agent sends messages from the source to the sink. Separate sinks are used to ingest data in different file systems. In the context of Kafka, the Flume is integrated to stream data in Kafka topic with a high speed, e.g. it is used when a heavy velocity data comes into place.

YARN enables the platform to perform operations by using a variety of tools like Spark for real-time processing, Hive for SQL, HBase for NoSQL and others. It

allocating resources and scheduling tasks. The Resource manager runs on a master daemon and manages the resource allocation in the cluster. Each Node manager runs on a slave daemon and is responsible for the execution of a task on every single data node. The Application master manages the job lifecycle and resource needs of individual applications. It works along with the Node manager and monitors the execution of tasks. The Container packages the resources including CPU, RAM, HDD, Network, etc. on a single node. Since, YARN processes the data in separate containers, which are logically units consisting of task and resources, the deadlock situations that appear typically in the first version of Hadoop, are minimized [8].

Spark delivers five types of components to the platform. The driver is a central coordinator that split the workload in to smaller tasks and schedules their execution on the executors. The driver passes the executors' requirements to the Spark application master, which negotiates for the resources with the YARN resource manager. The executors themselves and the Spark application master are hosted by the YARN containers. The Spark context allows Spark driver to access the YARN resource manager and keeps track of live executors by sending heart beat messages regularly. The executors perform the assigned tasks on the worker nodes and return the result to the Spark Driver.

4.2 Transactional Application

The architecture of the CogniSoft transactional application follows the Service-Oriented Architecture (SOA) paradigm. Its implementation includes several software modules, described in this section.

Architecture of CogniSoft Transactional Application. The architecture of the CogniSoft transactional application is shown in Fig. 3.

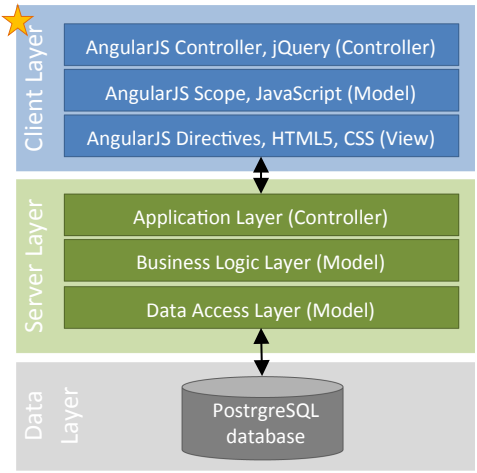

Fig. 3. Architecture of the CogniSoft transactional application.

The *Client layer* provides a web-based user interface (UI) for both patients and clinicians. It is implemented using Angular web application framework, which is embedded in HTML to create dynamic responsive web applications. The AngularJS supports the MVC architecture and thus it is smoothly integrated with the server layer. The two-way data-binding provides automatic synchronization between the model and the view, which are always synchronized. In this way, the user always sees an up-to-date view of the model.

The AngularJS scope objects refer to the application model. They are arranged hierarchically to mimic the DOM structure of the application. The AngularJS controllers augment the AngularJS scope objects by setting up their initial state and adding behavior to them. The role of the AngularJS directives is to attach a specific behavior to the DOM elements or to transform them and their children.

The *Server layer* provides a set of Application Programming Interfaces (APIs), implemented using Spring framework. It is developed using the Spring web application framework that fully support Representational State Transfer (REST) architecture. Thus, the components of the platform are implemented as a set of REST services. Since Spring is realized in a modular fashion, the developers are able to pick the modules that are relevant to their server development.

The *Data layer* provides data repository of the patients' data and their results from diagnosis and rehabilitation tests. It is built on PostgreSQL Relational Database Management System (RDBMS). Different instances will be created, working in different modes (transactional or analytical). The transactional databases and analytical database use a unidirectional master-slave replication.

Modules of CogniSoft Transactional Application. The CogniSoft transactional application is implemented in a modular fashion. It consists of 7 main modules, describes in this section.

The *Security module* is responsible for the security and privacy issues. Since the application works with sensitive data, along with the anonymization, a data separation technique is applied. For example, the user profiles are separated by their personal records. All APIs in the server layer relies on the Spring Security library, while the client layer uses Angular Security. Angular implements a lot of best practices and built-in protections against the most popular web application attacks. The data access is controlled not only on the application level, but additional measures are provided on a network level. The data is stored on machines, which are part of the internal network. External users can access the system by sending requests to a web server behind a proxy firewall using SSL cannel. The architecture of the application is stateless, meaning that the authorization is passed through a standard JWT in the HTTP header. The header is issued during authentication and contains assertions signed by the server.

The *User module* is responsible for users' roles and profiles. The users' roles are the same for the client and the server layers. The *Administrators* can access all system elements, including the Audit module, and are allowed to create nomenclatures and to perform CRUD operations on every types of objects. The *Clinicians* can administrate the patients' records and are allowed to create groups of patients, versions of the tests and group of tests in batteries, to assign and monitor the execution of batteries. The *Patients* can execute tests, which are assigned to them and eventually to access the results from the execution. The *Controls* can execute tests, which are assigned to them.

The *Nomenclature module* provides functionality for definition of system nomenclatures, which are multilingual. All nomenclatures are stored in a single table in the database, but they are grouped based on the type and the language.

The *Personal records module* implements functionality for administration of the personal records of users' roles clinicians, patients and controls. The information stored for each role is different. For example, the patient's personal record contains information about the disease, nationality, education, affiliation and other classification information, while the clinician's personal record keeps information about the specialty, medical center or hospital and participation in public healthcare projects.

The *Disease module* keeps track of the patient's disease progression. The corresponding data is stored in a separate JSON file associated with the patient's personal record. The *Test module* implements the diagnostics and rehabilitation tests. Currently, the functionality for the BDI-II and BICAMS battery is available. The clinicians are able to change the tests at the diagnosis stage as well as at the rehabilitation stage. The assignment of a test to a patient requires a new record to be added in the table "Assignment" of the database. Each test assignment associates a set of tests (test battery) with a group of users. The test execution is recorded in the table "Execution" of the database. Each executable test includes a static part, stored in the table "Header" of the database, which is common for all tests of the same type. The dynamic part of the executable test defines how the test will be performed.

The *Audit module* performs audit on the system actions such as login and logout as well tracks the changes of every valuable object. Information about the users responsible for the creation and last modification of the audited records are stored.

5 Test Automation

The BICAMS provides a cognitive assessment for MS, covering the most vulnerable neuropsychological domains of mental function – speed of information processing, episodic memory, visual-perceptual functioning, attention. They are automated in the current version of the CogniSoft transactional application.

5.1 Symbol Digit Modalities Test

The patient is required to perform a symbolic substitution with a corresponding digit of a given 9-digit code (key) for a period of 90 s. The test is based on the paper version using:

- Key field – it consists of 9 symbol-digit pairs that illustrate the correspondence between a symbol and a digit (1 to 9). A standardized predefined character pool is used.
- Work field – contains 6 couples of rows (consecutively arranged tables with 15 columns and 2 rows each). Each top row contains characters from the represented key in random order. Each bottom row is blank and must be filled in by the patient with the corresponding (key) digits. The patient is expected to complete the task as quickly as possible. The correspondences should be filled in consecutive order.

The adapted computer version of the test consists of the following steps: (1) Instruction, showing text or audio directions for execution of the test; (2) Demonstration, showing a short video that demonstrates the test execution; (3) Training, which is one time entering the data in the work field; (4) Execution, which is the actual test execution; and (5) End, which finishes the test and shows an encouragement message.

The test execution involves sequentially filling in the numbers on the displayed screens (standard 6 screens). The final screen ends with the field where the patient can fill in any additional information about the particular circumstances surrounding the test execution and his performance. At the discretion of the therapist, information related to performance outcomes is displayed or not to the patient. This includes: execution time, number of correct answers, problematic key symbols (recurring errors), etc.

Figure 4 shows how the SDMT looks like in the CogniSoft transactional application.

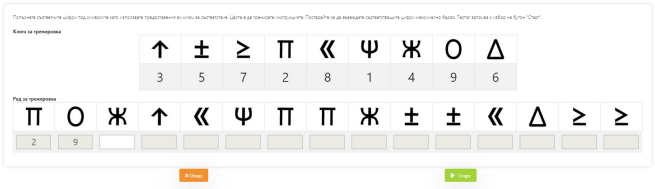

Fig. 4. UI of the SDMT.

A rehabilitation game is implemented on the same principle as the SDMT by replacement of the symbols with pictures. The configuration parameters allow the clinician to change the size and contents of the key, the size of the work field, the number of training screens depending on the patient's specific needs.

5.2 Brief Visuospatial Memory Test

The CogniSoft transactional application adapts the paper version of the BVMT. It provides specially designed forms containing 6 predefined geometric shapes (Table 2 × 3), as is shown in Fig. 5.

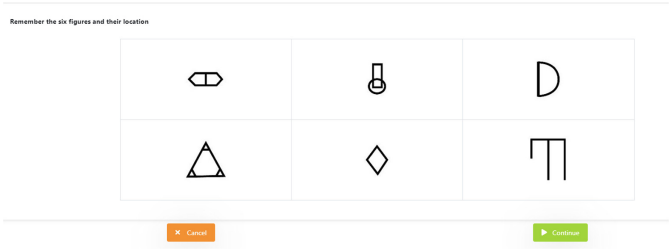

Fig. 5. UI of BVMT.

The test consists of the following main steps:

- Memorizing: For 10 s a table 2 × 3 is shown with validated figures from the predefined pool.
- Reproducing: Two tables are shown. One is empty 2 × 3 table and the second one is with size 3 × 4. It is filled with 12 figures. Among them in a random principle are included the 6 figures from the previous screen (validated) and 6 figures similar to those (not validated). The aim is with drag and drop the patient to fill the empty table with the correct figures based on what is memorized from the screen.

This process is repeated 3 times and the best performance is taken into consideration. The adapted computer version of the test consists of the following steps: (1) Instruction showing text or audio directions for execution of the test; (2) Execution, which is the actual test execution, consisting of 2 parts: memorizing and reproducing, described above; and (3) End, which finishes the test and shows an encouragement message.

The results are based on the number of correct figures and their correct positioning in the resulting table. In a rehabilitation mode, other pools of figures are used, and also the configuration parameters can vary such as storage time, table size, etc.

5.3 California Verbal Learning Test

The CVLT is implemented using three tables of words. The table with the "validation" words includes 4 categories of words such as sport, paper products, geographic objects and sweets. The second table contains words that are similar to those in the first table and falls in the same categories. The third table contains words that falls in categories that are different from ones of the validated words. During test execution 8 words are randomly selected form the first table, 6 words are randomly selected form the second table and 2 words are randomly selected form the third table.

The test consists of a series of 5 consecutive attempts, during which the same validated 16 words belonging to the selected categories are read. The patient should try to memorize them and then indicate them. After each attempt, statistics are kept on the number of memorized words and the categories to which they refer. The expectation is that after every attempt, the number of correct answers will increase.

The computer version of the CVLT adapts the paper version and consists of the following steps: (1) Instruction, showing text or audio directions for execution of the test; (2) Demonstration, showing a short video that demonstrates the test execution; (3) Execution, which is the actual test execution, consisting of five attempts; and (4) End, which finishes the test and shows an encouragement message.

The patient recognizes the words by pressing the selection buttons labeled with the words heard. The recognition is performed in two steps corresponding to two different screens. The final screen ends with possibility for the patient to fill in additional information about the particularities and circumstances surrounding the tests execution and performance. At the discretion of the therapist, information related to the results of the performance is displayed or not displayed: time of implementation, number of correct answers, problem categories.



Fig. 6. Rehabilitation game based on the CVLT.

A rehabilitation game, shown in Fig. 6, is implemented by adding relevant pictures on the buttons in addition to the words. There are also variations in configuration related to the number of validated words used, recognition screens, number of repetitions.

6 Conclusions and Future Work

The preventive and therapeutic approaches, tailored to patient requirements, need a personalized medicine, which early detects diseases. It is a societal challenge to adjust to the further demands on health sectors due to the ageing population. The effective healthcare requires improved decision making in prevention and in treatment provision as well as identification and support of best practices.

The digital twin platform described in this paper provides a practical first step towards application of the precision medicine in MS. The platform's architecture and technologies are described with a focus to its two main components:

- CogniSoft transactional application, which automates tests for diagnosis and rehabilitation of MS such as BICAMS tests;
- Advances analytical application, which provides services for data aggregation, enrichment, analysis and visualization that will be used to produce a new knowledge and support decision in each instance of the CogniSoft transactional application.

Being web-based, the CogniSoft transactional application guarantees maximal access to its functionality in different neurological departments. It is the main data provider for the analytical application, which is not currently developed. Once integrated in a common digital twin platform, the clinicians will be able to perform efficient diagnosis, prognostication, and management decisions for individual patients with MS. In the presented implementation clinical questions that deal with the shorter-term management of patients are covered. The patient state assessment and rehabilitation, covered by the CogniSoft transactional application, help in intervention planning in the hospital setting. Assessment of real prognostic performance needs a longitudinal set-up, which is the subject of ongoing research activities and the focus of the analytical

component of the platform. The potential of the platform is wider and could be expanded to provide a long-term decision support and evidence-based outcome predictions in MS. The clinicians will be able to quickly interpret the patients' data and to view the probable outcomes form rehabilitation, based on past patients' data in their neurological departments, EHRs from other systems, open clinical databases, etc.

Acknowledgement. This research work has been supported by CogniSoft "Information System for Diagnosis and Prevention of Multiple Sclerosis Patients" project, funded by the Program for Innovation and Competitiveness, co-financed by the EU through the ERDF under agreement no. BG16RFOP002-1.005, GATE "Big Data for Smart Society" project, funded by the Horizon 2020 WIDESPREAD-2018-2020 TEAMING Phase 2 programme under grant agreement no. 857155 and CogniTwin "Digital twin modelling of patients with cognitive disorders" project, funded by the Bulgarian National Science fund, under agreement no. KP-06-N32/5.

References

1. IBM. Bringing big data to the enterprise. http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html. Accessed 20 Jan 2020
2. MIT Technology Review. https://www.technologyreview.com/business-report/data-driven-health-care/free/. Accessed 20 Jan 2020
3. van Rijmenam, M.: Four Ways Big Data Will Revolutionize Education. https://datafloq.com/read/big-data-will-revolutionize-learning/2016. Accessed 20 Jan 2020
4. Aslam, R.W., et al.: A systematic review of the diagnostic accuracy of automated tests for cognitive impairment. Int. J. Geriatr. Psychiatry **33**(4), 561–575 (2018)
5. Wild, K., Howieson, D., Webbe, F., Seelye, A., Kaye, J.: Status of computerized cognitive testing in aging: a systematic review. Alzheimers Dement. **4**(6), 428–437 (2008)
6. Zygouris, S., Tsolaki, M.: Computerized cognitive testing for older adults: a review. Am. J. Alzheimers Dis. Other Dement. **30**(1), 13–28 (2015)
7. Lapshin, H., O'Connor, P., Lanctôt, K.L., Feinstein, A.: Computerized cognitive testing for patients with multiple sclerosis. Mult. Scler. Relat. Disord. **1**(4), 196–201 (2012)
8. Saraswat, H., Sharma, N., Rai, A.: Enhancing the traditional file system to HDFS: a big data solution. Int. J. Comput. Appl. (0975 – 8887) **167**(9), 12–14 (2017)
9. Introduction to Kafka. https://gist.github.com/kbeathanabhotla/0183e312307b05835c74. Accessed 14 Jan 2020
10. Vincent, A.N., Roebuck-Spencer, T., Fuenzalida, E., Gilliland, K.: Test–retest reliability and practice effects for the ANAM general neuropsychological screening battery. Clin. Neuropsychol. **32,** 1–16 (2017)
11. Doniger, G.M.: NeuroTrax Computerized Cognitive Tests: Test Descriptions, NeuroTrax Corporation, June 2014
12. Real-world Applications of Cognitive Assessment. https://screen-inc.com/uses-of-cognitive-tests/. Accessed 20 Jan 2020
13. CANTAB Insight. https://www.cambridgecognition.com/products/digital-healthcare-technology/cantab-insight/. Accessed 20 Jan 2020
14. Congstate Cognigram. https://www.cogstate.com/category/cognigram/. Accessed 20 Jan 2020
15. Congstate Academic Research. https://www.cogstate.com/academic-research/. Accessed 20 Jan 2020
16. CAMCI-Research. https://pstnet.com/products/camci-research/. Accessed 20 Jan 2020