

Medical 4.0: Medical Data Ready for Deep and Machine Learning

Natalia Labuda¹, Tomasz Lepa², Marek Labuda² and Karol Kozak^{1,3,4*}

¹Clinic for Neurology, Carl Gustav Carus Campus, Technische Universität Dresden, Fetscherstr 74, D-01307 Dresden, Germany

²Digital 4.0 Sp.z.o.o., ul. Smęgorzyńska 26, 80-298 Gdańsk, Poland

³Fraunhofer IWS, Winterbergstr 28, D-01277 Dresden, Germany

⁴Medical University Wrocław, Wybrzeże L Pasteura 1, PL-50-367 Wrocław, Poland

Abstract

Digitization is not an end in itself, but rather a means to an end. Ultimately, the point is to continue improving care provision in Europe, creating more efficient medical processes and achieving additional benefits for health care system. Medicine represents the next frontier for the Internet of Things: Four digital evolutions. Doodle is a famous process and database and works best for coordinating teams of people. Doodle radically simplifies the process of collecting data for scheduling events, meetings, appointments. However, the need to collect and share available scientific data on single PC or server from multiple project partners is a challenging process. This manuscript presents Medical 4.0, a doodle alternative for scientific data, presenting how radically the platform contributes to the four digital evolutions in medicine.

Keywords: Digitization; Doodle; Medicine

Introduction

In current migration situation more people in Europe than ever before are receiving regular healthcare. But managing this onrush of literally millions of new patients has forced the hospital and industry to confront significant challenges. Digital transformation enables the continual build-out and extension of services and data, while coordinating services across the care continuum to support population health.

Decentralization and complexity of the digitally captured medical data will consequently be followed by a decentralization of devices for new therapies, use of tablets, 3D technologies and computers in hospitals. Today's patient by patient entry in spread sheets will no longer be practicable with the further increase of data complexity. Single point data management must be replaced by a system of rules which the partial entry points will follow flexibly within the limits specified for them and the current situations signalled by the other partial data entry. The user-side data entry interface required careful design decisions to provide users with a clear and concise interface, process and workflow. In next step it is necessary for exploring patient data with deep and machine learning methods (Weka [1], R-Project [2], SVMlib [3], TensorFlow [4], Caffe [5], Apache Singa [6], Torch [7], Theano [8]) available to the general data management platforms through various implementations. Machine learning methods have been already applied to a broad range of areas within medicine for predicting and classifying medical data. The aim of the four digital evolutions [9] in medicine (Figure 1) is to exploit the potential resulting from:

- The extensive use of the in house-telemedicine products,
- The integration of medical processes with public health processes,
- Full implementation of decision supports systems,
- Therapy personalization,
- The digital mapping and virtualization of the real world and
- The opportunity to create "smart" solution for clinical research studies by using case report forms (CRFs) standards.

Data collection from medical research projects is a critical phase in current digital era, which leads to generation of high quality, reliable

and statistically sound data from studies. Electronic Data Capture (EDC) software can meet the need for faster and more reliable collection of eCRFs. Many software tools are available for data management. Commonly used tools are oracle clinical [10], macro [11], rave [12] and eClinical Suite [13,14]. In terms of functionality, these software tools are more or less similar and there is no significant advantage of one system over the other. Additionally, university hospitals or research institutions use custom-made EDC tools to suit their operational needs and procedures. Among the open source tools, the most prominent ones are LabKey [15], TrialDB [16] and Caisis [14], Catalyst Web Tools [17] from University of Washington Learning and Scholarly Technologies, OpenClinica from Akaza Research [18] and REDCap from Vanderbilt University and the REDCap Consortium [19]. Doodle [20] is a famous tool best for coordinating teams of people. With the app, people can control event data capture, set out time and date options for a meeting and then poll a group to capture time slots working best for them. Doodle radically simplifies the process of scheduling events, meetings, appointments. Doodle is an Internet calendar tool for time management and coordinating meetings. Meeting coordinators (administrators) receive e-mail alerts for votes and comments. M4.0 (<http://medical40.org/>) is like a doodle but optimized for complicated scientific research projects following doodle fundamentals: design, collect, distribute report. M4.0 is covering:

- Data capture calendar, data capture process,
- Facilitating the creation of patient-driven research studies that focus on optimizing healthcare and/or dealing with disease,
- Secure patient data exchange between health care centres and consortium partners,

***Corresponding author:** Karol Kozak, Clinic for Neurology, Carl Gustav Carus Campus, Technische Universität Dresden, Fetscherstr, Dresden, Germany, Tel: 035183391-3717; E-mail: Karol.Kozak@uniklinikum-dresden.de

Received November 14, 2017; **Accepted** November 28, 2017; **Published** December 06, 2017

Citation: Labuda N, Lepa T, Labuda M, Kozak K (2017) Medical 4.0: Medical Data Ready for Deep and Machine Learning. J Bioanal Biomed 9: 283-287. doi:10.4172/1948-593X.1000194

Copyright: © 2017 Labuda N, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

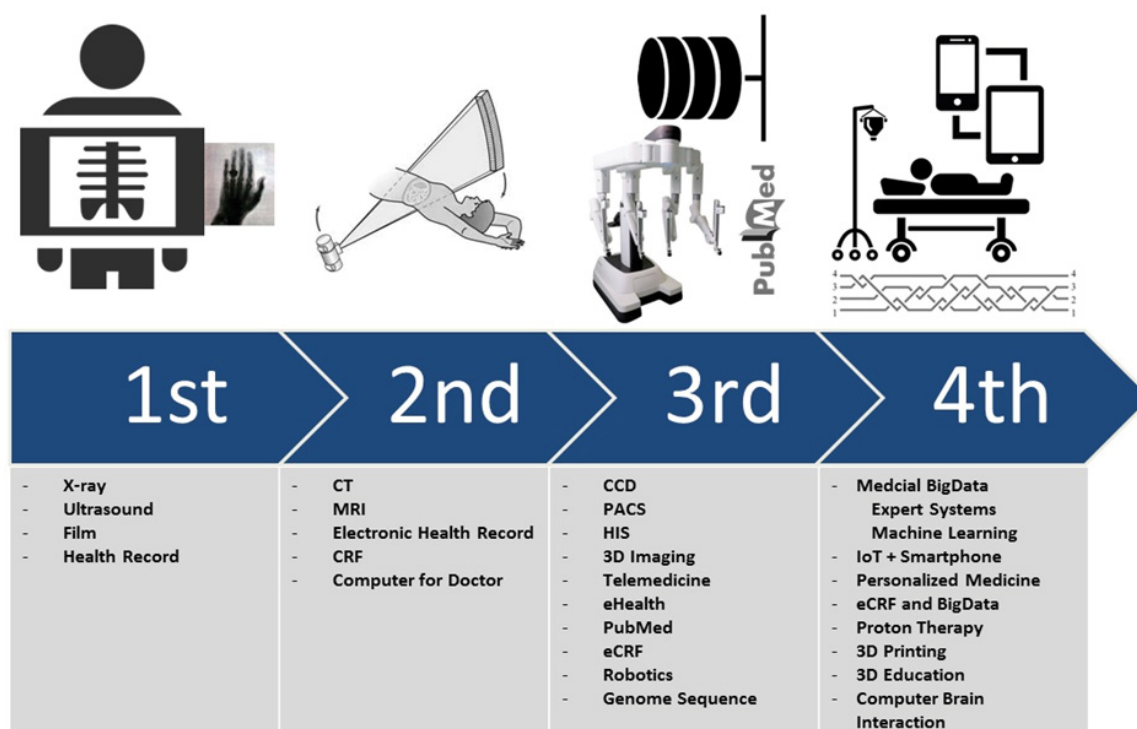


Figure 1: Digital evolution in medicine. Medicine is entering four digital evolutions.

- Deliver medical data ready for deep and machine learning,
- Deliver integrated deep learning frameworks into medical platforms.

M4.0 is thus not a parallel world to doodle, but rather an important trajectory on the way to the future of digital healthcare and four digital evolutions.

Methods

What's the significant difference between Doodle and Medical 4.0? While Doodle is an appointment setting process, controlling automatically schedule of appointment for survey members, M4.0 handles both automated inbound scheduling (like Doodle) but may capture any type of scientific data optimised for medicine, including project management in scheduling fashion, reminders, reporting and exporting. Here's a feature-by-feature breakdown of M4.0 and Doodle (Table 1).

Everyone involved in clinical research understands the importance of saving time and the need to do more with fewer resources. Changing working on studies from paper or excel based methodology to M4.0 model can help solve common challenges of data capture, such as:

- Efficiently managing time spent on data collection and review,
- Ensuring compliance,
- Minimizing data entry errors,
- Finding what you need, when you need it.

Functionality	Doodle	Medical 4.0
Send invitation	V	V
Design survey	V	V
Different data types		V
Flexibility on data forms		V
Report: decision from survey	V	V
Data inside of owner		V
Support images		V
Simplicity	V	
Intelligent data fields		V
Data audit		V
Email reminder	V	V
Mobile devices	V	V
User privileges management		V
User comments for survey	V	V

Table 1: Functionality comparison: Doodle vs. medical 4.0.

Data is typically recorded into the system by import masks and saved in a medical standard: Electronic case report form (eCRF).

eCRF designer

When creating eCRFs in M4.0, there are design options to choose from that allow optimization of forms for data entry and export and replicate complex assessments such as questionnaires. Saving eCRFs in a shared library allows for re-use of common eCRFs. When building eCRFs, edit checks can be programmed to help prevent data entry errors.

Data entry validation

Real-time checks on data are performed as it is entered into the system, based on the edit checks that were specified into the eCRFs.

Query management

M4.0 provides streamlined communication between study coordinator, data managers and for example patient care personal. The system has auto-generated queries, as a result of programmed edit checks and the ability to manually add queries. All queries need to have a response and be resolved before the data can be locked.

Audit trail

M4.0 supports validation and compliance with 21 CFR Part 11 including the appropriate technical controls delivering an audit trail that tracks the history of an eCRF. This includes a track of who made a change, the date and time a change was made, the old and new values entered and the reason the change was made.

Reporting and exporting

Once data is ready to be pulled out of the system, there are easily accessible and configurable reports and exports to statistical and machine, deep learning software applications.

Location of M4.0

Proper security practices must be instituted on the network and server(s) hosting M4.0 and also within the M4.0 software itself. This level of security exists in order to protect the data stored within M4.0.

M4.0 is cloud-hosted with data entry performed on a web-based interface inside of medical institution (Figure 2). Being in the cloud means easy access from anywhere with a browser and connection to the internet. Each user account in the system has a unique username, password and designated permissions. This limits what each person can see and defines the actions they can perform within the system. M4.0 stores patient data and all system and project information in multiple relational database tables within a single PostgreSQL database [21], which is an open source database. The framework used for back end and front end of M4.0 is written in C#, which is widely used, robust, object oriented programming language. For security the institution installing M4.0 will store all data on its own servers or in case of individual investigator it is possible to install M4.0 on single PC. All patient and project data are located at the local institution and no data will leave that institution is ever transmitted at any time by M4.0 from that institution to another location.

Security mechanisms

To protect and secure the medical data stored in M4.0 storage and databases, the system applies various mechanisms to protect against malicious users who may attempt to identify and exploit any security vulnerabilities in the system. Medical hospitals install on M4.0 server operating system enterprise-level internet application security scanners, to scan and test M4.0 security and its ability to withstand various methods of attack. The team behind M4.0 is deploying constantly updates of: database system (SQL Server) [22], web server (IIS) [23]. NET libraries, web browser and operating system itself. M4.0 system for data security reason is having webserver install in public zone and

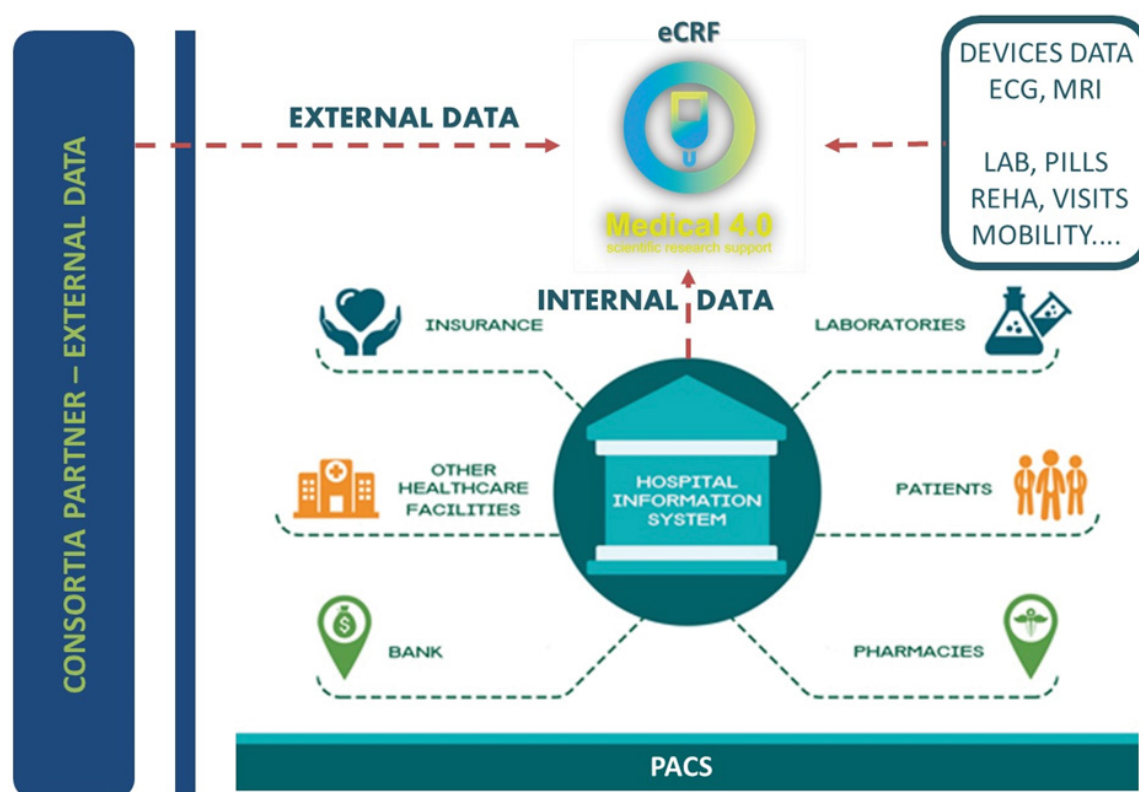


Figure 2: M4.0 is secure database located inside of medical institution with secure web ports and protocols to outside users.

Study	Description	Type of data capture	Amount of patients	Amount of questions in eCRF	Amount of parameters from medical devices and eCRF	Amount of modules in eCRF	Medical experts and centers	Location
Duchenne	The project monitors Duchene therapy and provides guidelines for personalized therapy.	Inside of hospital, home care	72	253	341	21	neuropediatric, rare diseases, rehabilitation, cardiology	Medical University of Gdańsk, Poland
Parkinson	The study evaluate nerve injury for Parkinson patients	Inside of hospital, home care	26	67	5	7	Neurology	Dresden University Hospital, Germany
PolSenior	Multidisciplinary projects on ageing examining medical, psychological and economical aspects of aging in Poland. Budget over 3 mln EUR, 6000 patients, 1500 patients in 8 regions will be examined by geriatricians.	Inside of hospital, home care	6000	712	2484	32	cardiology, geriatric, oncology, neurology	Medical University of Gdańsk, Poland
NoMED Cardiology	NoMED-AF study is the detection of atrial fibrillation in patients who are most likely to develop this type of arrhythmia, i.e. 65+ who are not aware of heart arrhythmia.	Inside of hospital, home care	3000	436	1115	33	Cardiology, geriatric, Rehabilitation,	Medical University of Gdańsk, Poland

Table 2: Piloting clinical research studies using M4.0 for electronic data capture.

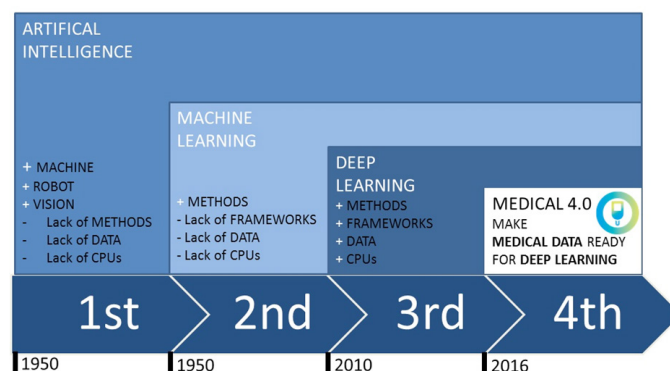
database itself with data installed in secure intranet zone inside of hospital. The connection between database and webserver is secure tunnelled with high level security mechanisms. To specifically protect against Cross-Site Request Forgery (CSRF), which is another method of attack, M4.0 utilizes a “nonce” (a secret, user-specific token) on every web form used in the application.

Results

From 2015 to 2017, the Medical 4.0 project group in Dresden and Gdansk, supported by the Digital 4.0 foundation, developed M4.0 and distributed the database at different University Hospitals in Poland and Germany. This software allowed the input of patients’ personal data, a structured data from different medical experts, clinical examination, clinical scores and treatment details. The M4.0 project was developed to provide scientific doodle tool for research teams for collecting, storing and disseminating project-specific clinical and translational research data. The first idea for Medical 4.0 project was initiated and installed at Medical Faculty, Technische Universität Dresden in June 2016. M4.0 was successfully deployed as a toolset and workflow for providing secure, web-based data collection services. Informatics investment has varied during the time of operation, but has typically not exceeded one person performing new module development work, supporting investigators and working with project consortium partners. M4.0 was developed created to support clinical research studies. In 2016, we launched a pilot studies from clinical research projects: Duchene, Parkinson and Seniority. We supported in the version 1.5, four pilots clinical research studies (Table 2).

M4.0 as source for digital evolution

Machine learning, deep learning and artificial intelligence (ML, DL and AI) are related technologies which have been evolved in history of digitalization. Clearly, they are very important and complex processes that are evolving very quickly. And now we enter Medical 4.0 System, in which computers, data, analytics and databases will come together in an entirely new way, with telemedicine connected remotely to computer systems equipped with machine and deep learning algorithms that can help in therapies with very little input from medical staff.



The emerging field of connected or digital healthcare is rapidly becoming a reality and has the potential to wedge itself into a staid system that has been averse to change.

While healthcare and life sciences organizations now have access to unprecedented amounts of information that can revolutionize treatment, it's still challenging to turn that deluge of data into personalized, actionable insights.

Medical 4.0 is a hospital intranet-based digital platform that:

- Optimizes data acquisition and integration (data flow),
- Delivers actionable, contextual and regulatory-compliant insights into the workflow of health and care professionals (insight flow),
- Creates an outcome-driven process enhanced by point-of-care automation across the care ecosystem (care flow),
- Focuses on the holistic view of the patient to support value-based care and remuneration models (financial flow).

People want to be better informed today and take more responsibility for their own health. Modern and sensible database solutions for collecting data from studies can make a significant contribution to development of new personalized therapies in medicine.

Discussion and Conclusion

The right database system for clinical research can save a significant amount of time by reducing inefficiencies and facilitating better

communication. User can build eCRFs, enter data, resolve queries and lock data all in one system. The sooner database lock occurs at the end of a study, the quicker the availability of the data for analysis. With M4.0, eCRFs are safely stored and backed up in a cloud-hosted system. With a library of standard eCRFs, it's easy to reuse them across multiple protocols. An M4.0 system offers hospitals flexible navigation, providing a database for deep learning, while using less excels tables and paper. The entire health and wellbeing landscape is being transformed by disruptive database innovations. The disruptive evolution 4.0 in databases and digital healthcare are both driving and enabling the changes in healthcare. They are creating a vision of the future and helping to realise that vision. Finally, making transparent the large-scale aggregation of anonymised clinical data is improving outcomes and facilitating best practices across systems.

References

1. Witten I, Frank E, Hall M, Pal C (2016) Data Mining: Practical Machine Learning Tools and Techniques. 4th edn. Morgan Kaufmann.
2. R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
3. Chang CC, Lin CJ (2011) LIBSVM: A library for support vector machines. ACM Trans Intell Syst Technol 27: 1-27.
4. Abadi M, Barham P, Chen J, Chen Z, Davis A, et al. (2016) TensorFlow: A system for large-scale machine learning. USENIX Assoc, pp: 265-283.
5. Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, et al. (2014) Caffe: Convolutional architecture for fast feature embedding. Proceedings of the 22nd ACM International Conference on Multimedia. ACM.
6. Ooi BC, Tan K, Wang S, Wang W, Cai Q, et al. (2015) SINGA: A distributed deep learning platform. ACM.
7. Collobert R, Koray K, Farabet C (2011) Torch7: A matlab-like environment for machine learning. Ed: BigLearn, NIPS Workshop. Idiap Publications.
8. Bastien F, Lamblin P, Pascanu R, Bergstra J, Goodfellow I, et al. (2012) Theano: New features and speed improvements. NIPS 2012 deep learning workshop.
9. Chakravorti B, Ajay B, Ravi, Chaturvedi RS (2017) 60 Countries' Digital Competitiveness, Indexed. Global strategy.
10. <http://www.oracle.com/us/products/applications/health-sciences/e-clinical/clinical/index.html>
11. <https://www.elsevier.com/solutions/macro/macro-overview>
12. <https://www.mdsol.com/en/products/rave>
13. <http://www.omnicomm.com/>
14. <http://www.caisis.org/>
15. <http://www.labkey.com/>
16. <http://ycmi.med.yale.edu/trialdb/>
17. http://www.washington.edu/ist/web_tools
18. <https://www.openclinica.com/>
19. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, et al. (2009) Research Electronic Data Capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 42: 377-381.
20. www.doodle.com
21. <https://www.postgresql.org>
22. <https://www.microsoft.com/en-us/sql-server/>
23. <https://www.iis.net/>