Implementing a Web-Based Architecture for DICOM Data Capture in Clinical Trials

Daniel Haak, Charles E. Page, Thomas M. Deserno

Department of Medical Informatics, Uniklinik RWTH Aachen
dhaak@mi.rwth-aachen.de

Abstract. Medical imaging plays an important role in clinical trials providing qualitative and quantitative findings. Patient's data in studies is captured in electronic case report forms (eCRFs) instead of paper-based CRFs, which are provided by electronic data capture systems (EDCS). However, EDCS insufficiently support integration of image data into patient's eCRF. Neither interfacing with picture archiving and communication systems (PACS), nor managing of digital imaging and communications in medicine (DICOM) data in eCRFs is possible. Hence, manual detours for image data in study's data capture workflow increase errorproneness, latency, and costs. In this work, a completely web-based system architecture is implemented interconnecting EDCS and PACS. Our approach utilizes the open source projects OpenClinica, DCM4CHEE, and Weasis as EDCS, PACS, and DICOM web viewer, respectively. In the optimized workflow, user interaction completely takes place in the eCRF. DICOM data storage and retrieval is performed by middleware components hidden from the user, ensuring data consistency and security by identifier synchronization and de-identification, respectively. This shortens paths for image data capture in the workflow, reduces errors, and saves time and costs. Beside this, valuable data for further research is centrally and anonymously stored in a research PACS.

1 Introduction

Nowadays, medical imaging plays an important role in controlled clinical trials, performed in the process of developing novel drugs and medical devices. Image-based surrogate endpoints support qualitative and quantitative findings in clinical trials providing eligibility, efficacy, and security [1]. Nowadays, captured patient's data in such trials is stored and managed in electronic data capture systems (EDCS), which provide electronic case report forms (eCRFs) instead of the traditional paper-based CRFs. Automatic evaluation of data by range checks reveals errors directly during data entry improving data quality, saving time and costs.

OpenClinica is one of the world's most popular EDCS [2, 3]. This web-based application offers rich functionality for data storage and management. Since its release in 1993, digital imaging and communications in medicine (DICOM) [4] has established as the leading standard for storage and exchange of image data in

H. Handels et al. (Hrsg.), *Bildverarbeitung für die Medizin 2015*, Informatik aktuell,
DOI 10.1007/978-3-662-46224-9_48,

medical applications. Picture archiving and communication systems (PACS) are used for DICOM-based storage and communication of medical image data and metadata in hospitals for patient care, as well as for research purposes. However, EDCS such as OpenClinica lack in support of DICOM, neither integration of DICOM image data into eCRFs, nor communication between EDCS and PACS is possible.

A first step in connecting EDCS and PACS has been performed by van Herk et al. [5]. A PACS is queried from the eCRF for DICOM objects using the hospital's patient identifier. After de-identification, the DICOM objects are made accessible via the web access to DICOM objects (WADO) protocol [6], and linked to the eCRF. After this, eCRF and image data is transferred to a research server. However, in this solution, the DICOM objects must be available already in the PACS for eCRF integration.

In our previous work, various architectures for EDCS and PACS connecting system have been analyzed [7]. Based on these architectures, a criteria catalog of 30 requirements has been built and 25 DICOM viewer software projects evaluated in concerns of their suitability for the presented architectures. An optimal architecture has been determined [8], in which DICOM data is integrated into the eCRF together with usual patient's data, it is then transferred to an PACS, back-linked with the eCRF and visualized in a web-based DICOM viewer. DCM4CHEE[1] and Weasis[2] have been proposed as appropriate PACS and DICOM web viewer for this architecture, respectively.

In this work, a final and workflow-optimized architecture for integration of DICOM images into the electronic data capture process of clinical trials is presented.

2 Material and methods

2.1 Architecture

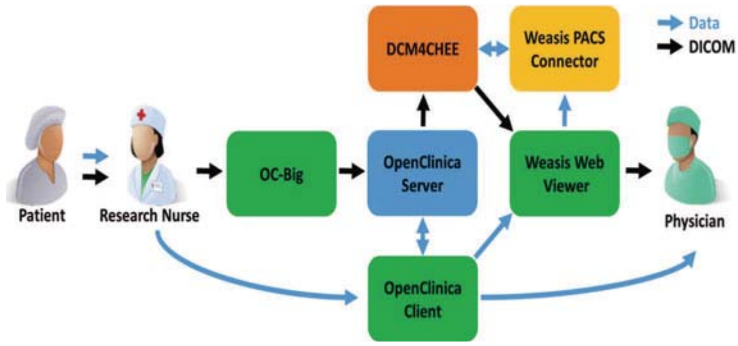
The architecture for a system connecting EDCS and PACS (Fig. 1) consists on client side of (i) OpenClinica client component; (ii) OC-Big, an OpenClinica extension for integration of binary large objects (BLOBs) and (iii) Weasis, a web-based DICOM viewer. The server side includes (iv) OpenClinica server component; (v) DCM4CHEE PACS and (vi) Weasis PACS connector.

2.2 Components

OpenClinica (Community Edition, Version 3.4) is an open source EDCS and clinical data management system. The web application allows design of studies with user-defined eCRFs and is used for data capture in multi-site clinical trials. OpenClinica follows industry standards and is approved by regulatory authorities such as the Food and Drug Administration (FDA). As usual for web applications,

[1] http://www.dcm4che.org/
[2] http://sourceforge.net/projects/dcm4che/files/Weasis/

Fig. 1. Architecture for EDCS and PACS connection and communication flows for data and DICOM.



OpenClinica consists of a server and a client component. OC-Big[3] is an extension for OpenClinica developed at our department [9]. The open source plug-in allows context-based integration of BLOBs into OpenClinica's eCRF.

DCM4CHEE (Version 2.17.2) is an open source DICOM Clinical Data Manager system consisting of an archive and image manager, and a PACS. The software is designed in a modular structure and provides various communication interfaces (e.g. for DICOM or HL7 communication).

Weasis (Version 2.0.2) is a open source web-based DICOM viewer, which is developed in Java. Weasis provides a wide range of image viewing tools (e.g. for measurements) and can be easily connected to PACS via WADO. Weasis PACS connector (Version 4.0) is an add-on for Weasis and allows invocation of the web viewer from a web context via the Java Network Launch Protocol (JNLP).

2.3 Workflow

Our system architecture is composed of two communication workflows through our architecture for storage and retrieval of PACS data (Fig. 2, top & bottom, respectively) which are triggered by user actions in the eCRF.

2.4 Storage

Typical patients data is stored as usual inside the eCRF on the OpenClinica client. For storage of patient's DICOM data, OC-Big [9] is used. Here, DICOM data, such as DICOM image (DCM) and binary image (JPG) data, document data (PDF) or compressed archives (ZIP, TAR, GZIP), can be selected for transfer. In addition, an OpenClinica identifier (OC-ID) is generated and additionally sent with the data. The OC-ID is constructed by patient and study identifiers in OpenClinica (OIDs), which are extracted from the selected eCRF. After successful transfer and reception of DICOM data by the OC-Big's server component,

[3] https://code.google.com/p/oc-big/

large binary object (BLOB) data is extracted and stored on server's file system. In the next step, the PHP script invokes a Linux shell (bash) script for PACS integration of DICOM image (DCM) and DICOM convertible (JPG, PDF) data. After a validation check of the system's environment (e.g. connection to PACS available, DICOM converter present), the extracted files are parsed and added to a list, consisting of DICOM and non-DICOM objects. During this, the StudyInsUID, which is mandatory for all DICOM objects, is included and all StudyInsUID validated regarding consistency. After this, an de-identification step is performed and all non-mandatory DICOM metadata removed from the header. Now, the PACS is queried via DICOM C_FIND command for the OC-ID and all listed files are stored via the C_STORE command into the PACS. During this, non-DICOM objects are converted into DICOM objects and a StudyInsUID added to the header. To ensure data consistency, in the next step this mapping is validated by an integrity check: the created list is compared with all DICOM objects which are mapped to the OC-ID identifier in the PACS. If this succeeds, a Weasis URL including the OC-ID is generated and stored into the eCRF by OC-Big.

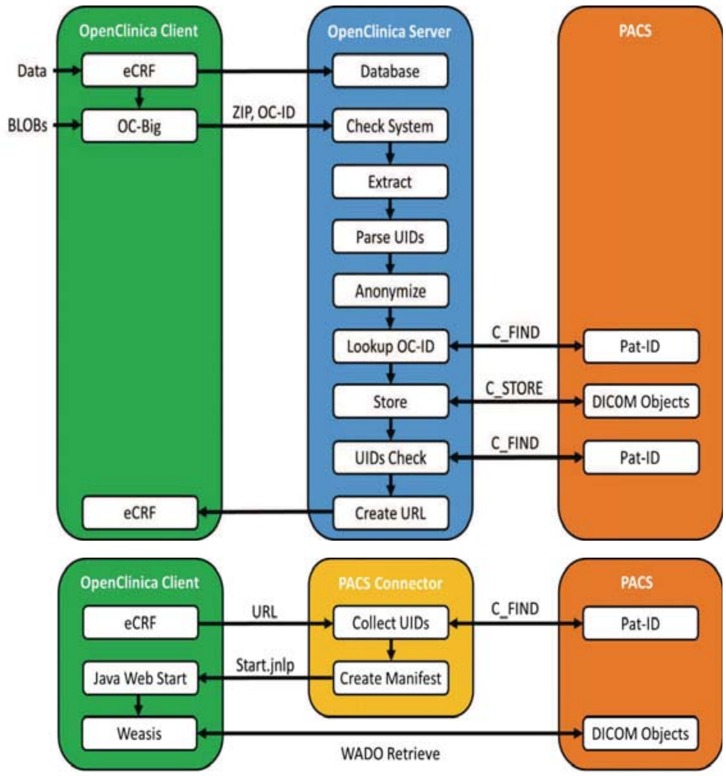


Fig. 2. Workflow of our implementation and involved components for DICOM data storage (top) and retrieval (bottom).

2.5 Retrieval

For retrieval of integrated DICOM data, a click on the Weasis URL embedded in the eCRF starts communication with the Weasis PACS connector. The connector analyzes the URL and fetches all DICOM UIDs from the PACS via C_FIND. Based on these UIDs a manifest of all DICOM objects listed in the PACS is created, written in a XML structure and encapsulated into a JNLP file, which is sent back to OpenClinica's client. Here, the execution of the JNLP file triggers the invocation of the Weasis web viewer via Java Web Start. Weasis now iteratively gathers all DICOM objects by their UIDs from the PACS via WADO and visualizes them in the web interface.

3 Results

Our resulting workflow is depicted in Fig. 3. For storage of DICOM objects in the eCRF, the data entry person (e.g. a research nurse) opens the eCRF of the patient in OpenClinica as usual and invokes OC-Big by a click on a button embedded in the eCRF (Step 1). In OC-Big, the user selects DICOM files from

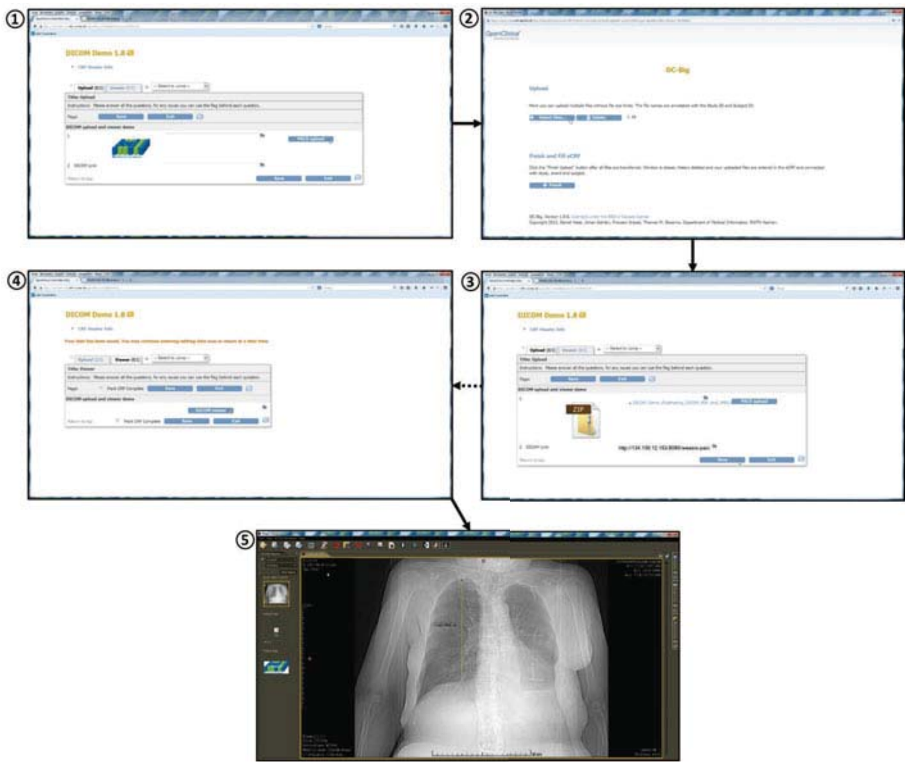


Fig. 3. Final workflow for study nurses using our implementation for DICOM data integration into EDCS.

the local file systems and starts the transfer (Step 2). After successful transfer of the DICOM data, a reference to the DICOM data is stored in the eCRF (Step 3). For retrieval of the DICOM objects, a button click in the eCRF starts Weasis web viewer (Step 4). After this, patient's DICOM data is available in Weasis for viewing or image interaction (e.g. measurements) (Step 5).

4 Discussion

In this work, a final architecture and implementation based on results of previous work for connection of EDCS and PACS has been presented. The architectures consist of multiple communicating components, whereof only two components (OpenClinica and Weasis DICOM viewer) are visible to the user. Both components are shared through web and accessible via modern web browsers, hence DICOM data can be stored and retrieved from everywhere (where an internet connection is available). Also hidden from the user, DICOM data objects and their metadata is automatically de-identified and synchronized and therefore, data consistency and security ensured, respectively. Nonetheless, all integrated data is stored only once in the PACS, which provides rich functionality for further data sharing. All used components are licensed as open source and affordable for low-budget clinical trials. However, DICOM data can only be visualized and annotations are currently not stored back, yet.

References

1. Miller CG, Krasnow J, Schwartz LH. Medical Imaging in Clinical Trials. London: Springer London; 2014.
2. Franklin JD, Guidry A, Brinkley JF. A partnership approach for electronic data capture in small-scale clinical trials. J Biomed Inform. 2011;44:103–8.
3. Leroux H, McBride S, Gibson S. On selecting a clinical trial management system for large scale, multi-centre, multi-modal clinical research study. Stud Health Technol Inform. 2011;168:89–95.
4. Mildenberger P, Eichelberg M, Martin E. Introduction to the DICOM standard. Eur Radiol. 2002;12(4):920–7.
5. van Herk M. Integration of a clinical trial database with a PACS. J Phys Conf Ser. 2014;489:012099.
6. Koutelakis GV, Lymperopoulos DK. PACS through web compatible with DICOM standard and WADO service: advantages and implementation. Conf Proc IEEE Eng Med Biol Soc. 2006;1:2601–5.
7. Haak D, Deserno TM. Integration of DICOM images into the electronic data capture workflow of clinical trials. Accepted on SPIE Medical Imaging 2015. 2015.
8. Haak D, Page CE, Deserno TM. Workflow-based integration of EDCS and PACS supporting image-based surrogates in clinical trials. Submitted to SIIM Annual Meeting 2015. 2015.
9. Haak D, Samsel C, Gehlen J, et al. Simplifying electronic data capture in clinical trials: Workflow embedded image and biosignal file integration and analysis via web services. J Digit Imaging. 2014;27(5):571–80.