Pilot 12: Radiology workflows

1. Key Information

1.1 Involved Partners

- Contextflow GmbH (CON)
- Atos Spain SA (ATOS)
- Medizinische Universität Wien (MUW)
- Servicio Madrileno de Salud (HUP)

1.2 Involved Countries

- Austria
- Spain

1.3 Keywords

- Clinical decision support
- Help to find the correct finding
- Finding similar cases
- Similarity search

1.4 Task Description

This pilot focuses on providing radiologists with relevant information during the reading of cases. During typical assessment of radiological imaging data, the radiologist parses the image, reports on findings, and in difficult cases, consults a range of sources to identify the finding, verify suspected findings, or to put the finding in the context of the disease. The prototype supports this by enabling radiologists to trigger search by marking a region of interest in the imaging data. The software then compares the marked patterns with a large database of cases, ranks those cases and shows the most similar ones. Furthermore, it provides a summary and scoring of the findings, and additional information is presented for supporting radiologists.

2. Building Blocks

2.1 Architecture

2.1.1 System Architecture

The Contextflow Big Data platform performs content-based image retrieval in the biomedical domain using deep-learning techniques such as Convolutional Neural Networks (CNN) or Recurrent Neural Networks (RNN) for data processing. Based on a query it ranks indexed examples corresponding to their similarity or relationship to the query, these ranked examples form the query result. When receiving the query, the platform can also anonymize the query data in the browser.

During retrieval, a user can use the system to search for relevant information that helps during the assessment, diagnosis, or reporting on an individual case. This information could for example be based on a query image of a computed tomography showing the lung of a patient, and a user indicated Region-of-Interest (ROI) in the image. If the ROI is placed on a region in the lung that exhibits emphysema, the system processes the image and ROI information, and could return a result consisting of cases containing emphysema, educational images illustrating typical characteristics of emphysema, research literature, or figures in research literature on emphysema. An overview of the platform is shown in Figure 1. It is possible to have a local installation and a cloud-based installation. For the local installation, all components in Figure 1 are within a hospital. For a cloud-based installation, the index and search engine are in the cloud and queried by anonymised queries.

The platform is deployed either via a browser or via an integration of an html based view in existing viewers.

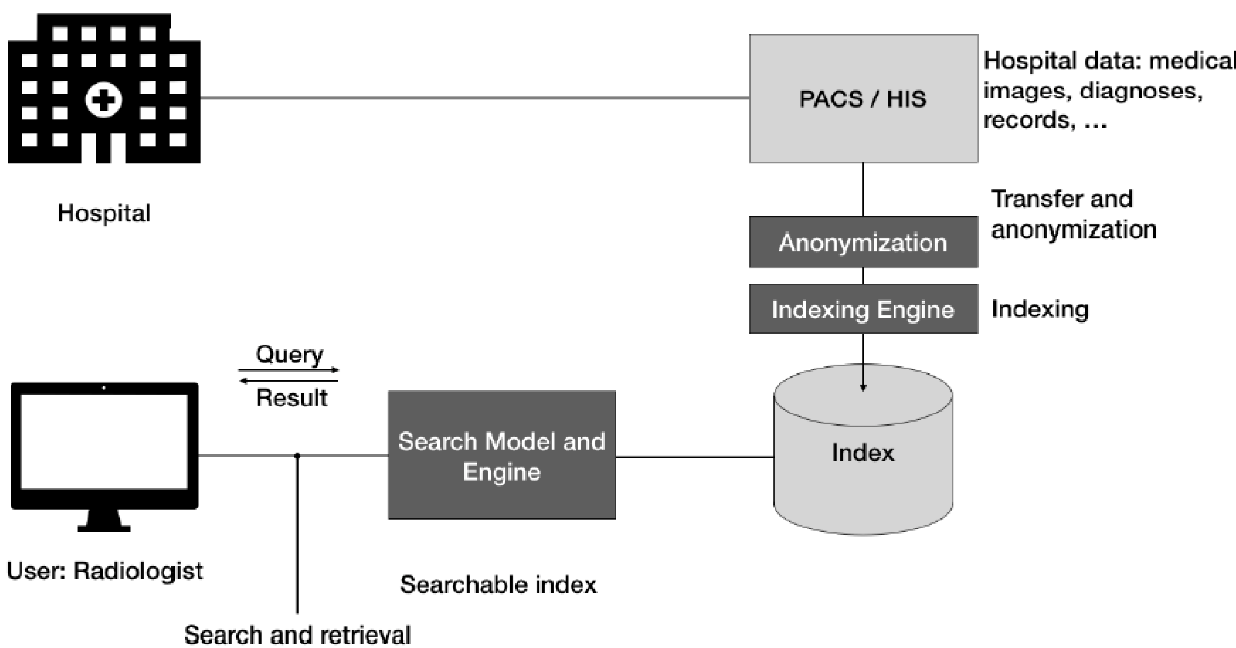


Figure 1 - Contextflow Big Data platform overview.

2.1.1.1 Indexing engine

The indexing engine is ingesting data such as images and textual information, and is using models trained prior to indexing to derive an optimal index and the necessary similarity relationships between imaging data or combined imaging and textual data across patients, cases, examinations, or other images.

The training of this indexing model increases the relevance of the similarity function in that it optimizes it to reflect disease related similarity across a training set of cases. For this it exploits linked imaging data, textual information (e.g. radiology reports) and structure information related to the disease and individual patient. During training the system is provided a large set of data to learn relationships and distance functions that reflect relevant similarity, or other relationships between images, or between patients that correspond to the images or textual data.

The system collects many (e.g., >1.000.000) localized appearance examples, and trains a Convolutional neural network (CNN) to derive an efficient and rich representation.

The core strength of this system is that this learning can happen in large part, or even only based on information collected from clinical routine, i.e., the need for additional expert annotation is minimized or removed entirely. The learned model can then be used to index a new (or the same) set of data. The Indexing engine receives a representation function, a distance function and a large set of image patches/images/image volumes that are to be indexed to allow efficient searching. It calculates the distance among all images, and prepares an index structure so that the distance between any new image and all of the indexed images can be calculated quickly.

2.1.1.2 Search model and engine

During search the user - a radiologist - inspects an image, and marks a region of interest to indicate a pattern that should serve as the basis for the search. This information is sent to the search model and engine as a query. Together with the query image the retrieval engine can receive the image information together with other data such as location or patient specific information.

The search process of the retrieval unit ranks the indexed images based on the query and enables filtering with query information. It provides the result either as output to the user at the user interface, or as an input to further processing.

The retrieval unit can perform multiple searches serving different users simultaneously, and also can search multiple indices.

2.1.1.3 User interface for radiologists

The Contextflow big data platform is used by radiologists reading individual patients' cases. They use the platform via their viewer or a browser based interface to obtain relevant information for the current case. The user interface enables the radiologists to (1) trigger a search by for instance marking a region of interest in the image or volume in front of them, and (2) view and explore the search results. The user interface enables the user to:

- Trigger a search based on a marked region of interest
- View search results and their statistical characteristics (e.g., findings)
- Group search results

- Explore search results by providing a detail view with which the user can inspect all cases in the result list, scrolling through the volume and visualizing the distribution of areas similar to the query.
- ● Visualization of reference information relevant for the diagnosis and differential diagnosis.

2.1.1.4 Anonymization Unit

Anonymization is a critical part of deployment and is performed in two alternative ways. Either anonymization is done on the hospital site, and only anonymized information reaches contextflow infrastructure. Or alternatively, the contextflow platform performs anonymization following standard specifications. This can happen for instance in the browser used for retrieval and as a graphical user interface.

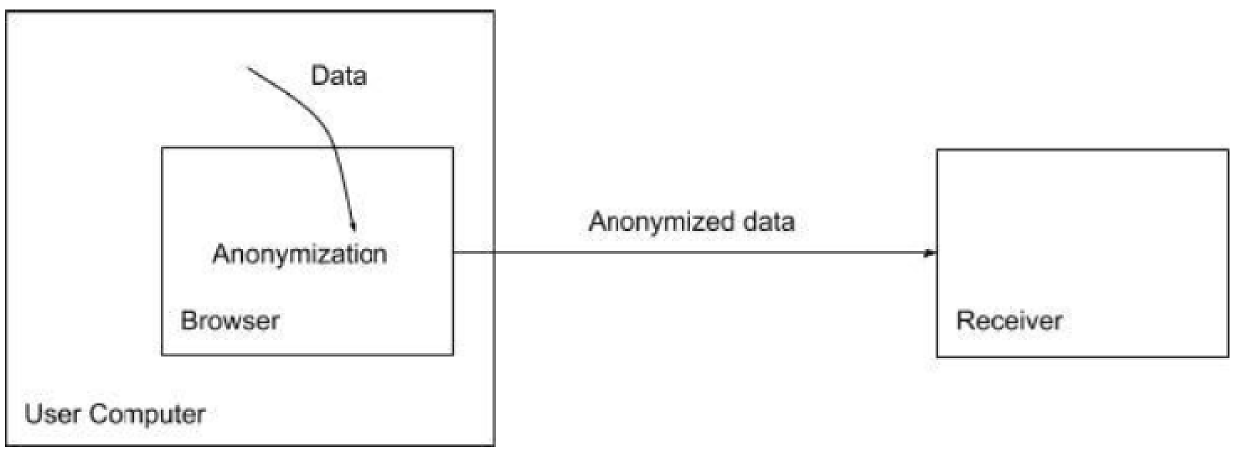


Figure 2 - Anonymization of query data in the browser before transfer.

Specifically, the anonymization engine performs anonymization of the data locally within the browser before uploading the data. The anonymization removes all features of the data that enable a linking to a specific person, and de-identifies the data through this process. The process includes steps such as:

- Removing fields in DICOM headers and headers of other medical document formats
- Removing personally identifiable information from text documents or structured data
- Removing image content that would enable identification.

2.1.2 Data Flow & Interoperability of services

Figure 3 shows the data flow of the system. Unanonymised data is processed exclusively within the hospital on the contextflow Local Server. Anonymised CT images returned as results by contextflow SEARCH are stored on the cloud. Queries to the search engine are done using abstract anonymised features extracted from the CT image under consideration - no image or patient data is transferred outside the hospital.

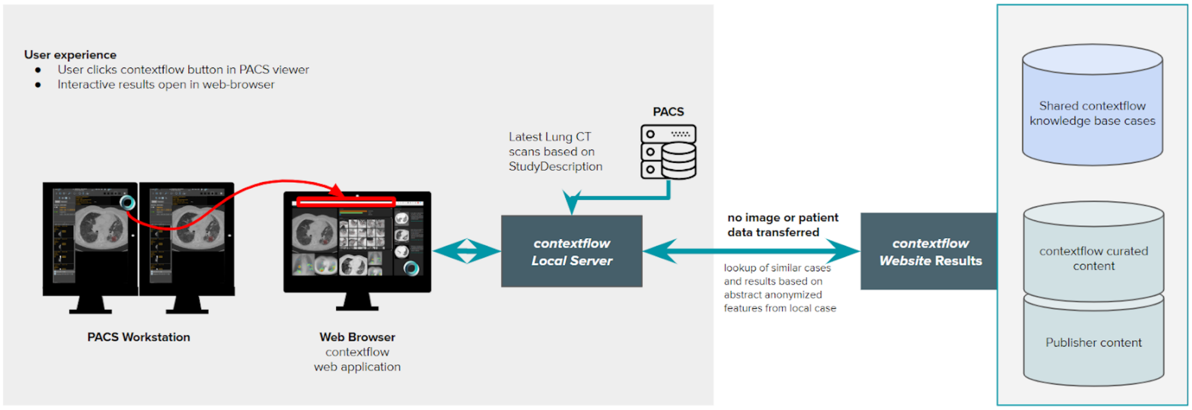


Figure 3 - Data Flow

2.1.3 Necessary Hardware

For the indexing, access to a server with GPUs is required. The query phase has been optimised to run on CPUs, so that it can be used on existing servers within hospitals, which often do not have GPUs installed.

2.1.4 Software Components

All of the main software components are proprietary to contextflow. For Machine Learning and Deep Learning, use is made of commonly available open source libraries.

2.3 Data Processing

2.3.1 Processing of large structured / unstructured data sources

2.3.1.1 Data Sources

Data Source	**Description**	**Acquisition**	**Characteristic (Size, Patients, Years, Origin/Region)**
MUW	Anonymised lung CT scans with report data	CT Scanners at the MUW	7600 lung CT scans of 5000 patients
MUW	Evaluation lung CT cases for the BigMedilytics	CT Scanners	108 study

	study	at the MUW	cases including diseased and healthy cases

multiple sources	**integration to data warehouse**	**data access**	**data stored in cloud**	**multi-party architecture**	**secure environment**	**transform raw / unstructured data**
yes	yes	de-identified data (e.g. images) will be processed	yes	no	no	yes

2.3.1.2 De-Identification and anonymisation

CT images used as training data by contextflow undergo an anonymisation process before leaving the hospital. As the final stage in the anonymisation process, each image is viewed by a person to ensure that no identifying characteristics remain in the image.

For the radiology reports, a text analysis tool is used that extracts anatomy and pathology terms, as well as negation. Only these extracted terms are transferred. This is a robust process as only information that is needed is extracted, as opposed to common anonymisation processes that attempt to remove sensitive information.

2.3.1.3 Acquisition

Data is acquired during clinical routine with heavily regulated image acquisition technology of different vendors (medical products). The data then typically resides in highly secured hospital data systems called picture archiving and communication systems (PACS) and internal databases.

The data forming the corpus is provided in bulk, after undergoing anonymization following standard (DICOM), and is transferred to the search database. These techniques are proprietary technology under the control of the hospital. Standards used involve DICOM image standards and transfer protocols.

2.3.1.4 Cleansing

Data cleaning consists primarily of selecting a subset of imaging data relevant for diagnosis, and for being included in the search corpus. Data such as navigator scans are excluded. Most of this is possible based on technical parameters.

Data cleaning was performed before transfer, using technical parameters of image acquisition such as resolution, or field of acquisition. This is exploiting DICOM standard encoded information. Sanity checks of regions imaged were performed using small vignet visualizations. Proprietary software was used.

2.3.1.5 Data Integration

Data integration encompasses the transfer of the data into an internal database, indexing of linked imaging and textual information, and preparation of search. Analysis is part of the indexing, and extracts searchable features that capture relevant information in the imaging data linked to disease and relevant radiological terms.

For analysis proprietary algorithms were used, including image parcellation algorithms, deep learning-based algorithms that extract features from imaging data, and segmentation algorithms that identify anatomical regions.

2.3.1 Multi-velocity processing of heterogeneous data streams

Does not apply

2.3.5 Complex real-time event detection

Does not apply.

2.3.5.1 Notifications

2.3.5.2 Situations of Interest

2.3.5.3 Event Processing

2.3.5.4 Event Sources

2.3.5.5 Evaluation

2.4 AI Components

2.4.1 Deep learning for multilingual NLP and image analytics

2.4.1.1 Natural Language Processing

The use of natural language processing (NLP) is applied in extracting structured information from radiology reports to create initial labels for training data. Anatomy and pathology terms as well as negation are extracted. Semantic information is mapped to terminologies or ontologies such as RadLex or MeSH. The extracted information is linked to images in the contextflow knowledge base.

Languag e	How will NLP support your pilot?	How will NLP help you to reduce costs?
German	NLP is important to use clinical routine data that comes with free text reports on the findings in images.	The NLP tool allows a rapid extraction of key information from radiology reports written in free text. In addition, this serves to anonymise the documents as only the necessary information is extracted from the reports. However, this process does make errors, in particular due to the sometimes complex style of writing in radiology reports, and it is necessary to have the results checked by radiology experts together with the images.

Which NLP tasks do you address?	Method	Softwar e framew orks	Vocabularies/corpus used	Describe your method in a few sentences.	Describe your corpus/training data?
Named entity recognitio n	However this is not focus of the develop ment in this project, we are using existing technolo gy in the company	tensor flow		Free text reports are parsed, words are mapped to a terminology and the terminology is used to extract location-finding pairs from the structured report	Radiology reports associated with the dataset obtained from MUW (described above)
Concept normaliza tion		tensor flow	Radlex, MeSH		

2.4.1.1.1 Evaluation

We have not done a formal evaluation of how well the software works, as we are using existing technology and employ radiology experts to correct the outputs.

2.4.1.2 Image Processing

Type	How will IA support your pilot?	How will IA help you to reduce costs?
Radiology	Image processing is crucial to perform the image pattern comparison, driving the image/case search during radiological diagnosis. DL is used to train it on the diseases relevant for the pilot, and to compare measures based on imaging data reflecting the disease specific appearance.	The BigMedilytics study in Pilot 12 showed that the average reading time by radiologists is 31% shorter when they are supported by the contextflow search engine. Reading time is the amount of time required for a radiologist to perform the analysis of an image and write the report. With this large reduction in reading time, it would be possible to read more images in the same time with the same number of radiologists, thus saving costs. A trend towards better diagnostic quality was also found in the BigMedilytics study.

Do you require	Which Image Processing tasks	Do you use public data	Describe your method in a few sentence.	Which training technique do you	Which pathologi	On which level does	What is the level of detail of your

regulatory approval?	do you address?	repositories? If yes, which?		use?	es are covered?	classification happen (volume, slice/img, or px-level)?	GT-annotations (volume, slice/img, or px-level)?
Yes, it needs CE certification (which has been obtained)	comparison of image content in medical imaging data (e.g. computed tomography volumes)	yes, for research we use available repositories e.g., challenges	The technique compares image content of a query ROI (region of interest) with a large number of image segment in the database (15Bn image segments, ~7000+ volumes). It identifies the closest matches and retrieves the corresponding cases.	we use mainly weakly supervised from radiological routine information, i.e., images together with report	lung diseases	on pixel level, and in a subsequent step, ranking is performed on a volume level.	volume level, and regions of interest for evaluation
	search based on a marked region of interest, to find image segments across a large database that carry similar patterns, and rank these cases according to similarity		During training we learn an image similarity function and index the imaging data to provide fast retrieval				

2.4.1.2.1 Evaluation

A study with radiologists was carried out, with 4 senior radiologists and 4 junior radiologists at the MUW. 108 study cases containing both diseased and healthy cases were used. Each participant reports: 27 cases without the CBIR-tool (Phase 1) and 27 cases with CBIR-tool (Phase 2). A participants reports a case only once (either without or with the system - never the same case with and without the system). The study showed that the average reading time by radiologists is 31% shorter when they are supported by the CBIR-tool. Reading time is the amount of time required for a radiologist to perform the analysis of an image and write the report. A trend towards better diagnostic quality was also found in this study.

2.4.2 Prediction Algorithms

Does not apply

2.4.2.1 Task

2.4.2.2 Data, Data Modelling

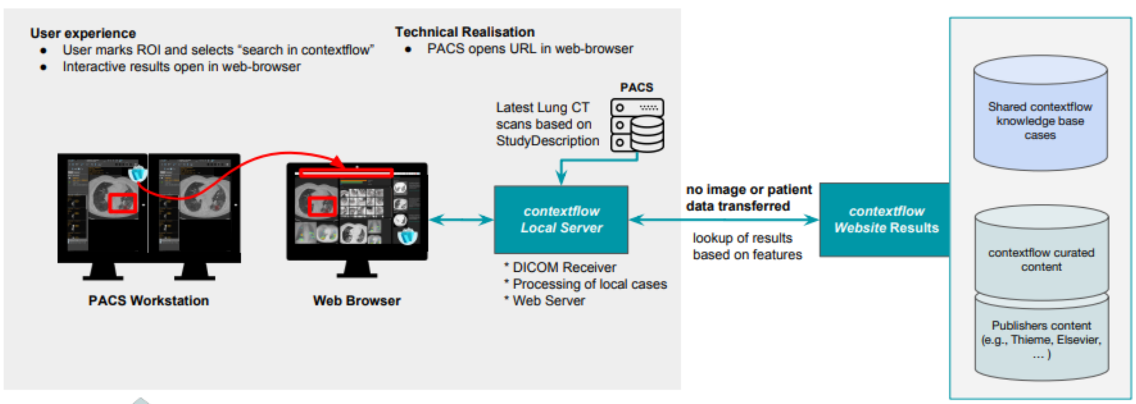
2.4.2.3 Features

2.4.2.4 Model

2.4.2.5 Evaluation

2.5 Security and privacy of data access and processing

The figure below shows the way the solution is devised.



The pilot deploys a local server at the user premises (typically the hospitals) and opens a connection to a centralized server where the search of the similar results is performed. Only abstract features extracted from the patient cases inside the hospitals are transferred outside the hospital to the search engine - it is not possible to infer any information about the initial image from the features. It is important to point out that no image or patient data is transferred outside the hospital.

2.5.1 Access Control

2.5.1.1 Authentication

It can be divided the type of access management control applied in these pilots into two different categories: those components which access sensitive data (radiological images) and the components without.

With regards to the components with access to sensitive data (within the grey square), it is worth highlighting that this access will be always executed in the hospital facilities and therefore into a controlled environment. These components will be fully integrated with the tools already in place and used by the health institutions as a brand-new functionality, hence the control access (Authentication and Authorization) will rely on the already used by these institutions.

Regarding the other components deployed outside of the hospital facilities it is worth mentioning that they will not access the sensitive information hence although access control is applied, they don't need to be so strict as the components described above.

2.5.1.2 Authorization

As described for authentication, authorisation makes use of existing processes in the hospital.

2.5.2 Data Protection

2.5.2.1 Data at rest

The sensitive data will always stay within the hospital facilities in a controlled environment. However, in any case the data store will be encrypted to guarantee that they will only be accessible with the appropriate authentication and authorization permissions.

2.5.2.2 Data in transit

All the communication between the different components involved in the pilot will rely on SSL/TLS protocol for the communications, hence the data will be encrypted, and therefore protected, in all the communications. Only anonymised data retrieved as search results is stored outside of hospital facilities. No sensitive data is transferred outside a hospital - the search query is sent in the form of abstract features.

2.5.3 Auditory and logs

2.5.3.1 System Auditory

This pilot is integrated with an already fully functional system within the health institutions, hence for the components with access to the sensitive data, the audit of the access to the services will rely on the services already in place. However, for the components which are outside of the hospital premises, although they don't access directly to sensitive data, will be monitored and audited in order to log all operations realized.

2.5.3.2 Services Auditory

2.5.4 Privacy measurements

2.5.4.1 Data Privacy Impact Assessment (DPIA)

The project was performed on retrospective anonymized data. Researchers did not have access to personalized information. The project and data transfer and processing as part of the project were assessed formally by the data clearing house of the MedUni Wien as part of the standard procedures in place to enable research.

2.5.4.2 Legal/Ethical process

The process to enable the study comprised both legal and ethical aspects.

Legal process: both a consortium agreement and a collaboration agreement were put in place to define the work, and use of data within the study.

Ethical process: before commencement of the study, a study protocol describing the research in detail was created and submitted to the Ethics Committee for approval. Only after approval did research commence. The ethics committee approval has to be renewed on a yearly basis, which has been done through-out the project to this date.

2.5.4.3 Processes for complying with the current legislation

The project partners have processes in place to ensure compliance with relevant legislation. Any data transfer has to be covered by legal agreements, a positive approval of the study protocol by the ethics committee, and a positive approval by the MedUni Wien data clearing house, that takes agreements, ethics approvals, and data specificiations into account, before approving transfer and processing. All of these steps were performed as part of the project.

2.6 Trustworthy AI

2.6.1 technology/user adoption and establishing trust

With the Pilot 12 solution, the radiologist remains in control and makes the final decision on what to write in the report on the case. The system makes suggestions on diagnoses, but, for each suggestion it makes, it shows a short summary document listing the key differential diagnoses and indications for that suggestion. The radiologist is therefore also reminded on what to check in order to confirm the suggestions from the system. The system therefore aims at a collaboration between radiologists and AI instead of replacing radiologists by AI. This approach is appreciated and assists in establishing trust and subsequent adoption.

2.6.2 ethical principles

- respect for human authority

As described in Section 2.6.1, the system makes suggestions while the radiologist makes the final decision on what to write in the report on the case. The radiologist is also able to ignore the suggestions made by the system if necessary.

- prevention of harm

As the system does not make unsupervised decisions, there is no possibility of a decision by the system harming a human.

- fairness

The system is trained on data from multiple sources, from both public repositories as well as from hospital partners. Efforts are continuously underway to expand the data sources. By collecting data from a wide variety of sources, we try to avoid biases in the data that could arise if data is collected from a small number of sources.

- explicability

When making suggestions on diagnoses, the system always shows summary documents listing indications and differential diagnoses for that suggestion. While this is not an explanation for what happens in the machine learning part of the system, it allows radiologists to rapidly verify the plausibility of the suggested diagnoses.

2.6.3 key requirements

- Human agency and oversight

The system does not make unsupervised decisions. The radiologist decides on what to write in the report on the case and decides which, of any, of the suggestions of the system are used.

- Technical Robustness and safety

The system has been certified as a medical device in the EU. To achieve the certification requires the development and adoption of processes for software development, error fixing, software update, etc., which guarantee the robustness and safety of the software.

- Privacy and data governance

Ethical approval is obtained for the use of the data. Sensitive data stays in the hospital, governed by existing data access policies and procedures. Only anonymised data is transferred to the company. Anonymisation procedures including human oversight are defined at the data providers.

- Transparency

Radiologists undergo training on the system and its capabilities and limitations before using the system. For each suggestion made by the system, a summary document on indications and differential diagnoses allows the radiologist to check the plausibility of the suggestion.

- Diversity, non-discrimination and fairness

The system is trained on data from multiple sources, from both public repositories as well as from hospital partners. Efforts are continuously underway to expand the data sources. By collecting data from a wide variety of sources, we try to avoid biases in the data that could arise if data is collected from a small number of sources.

- Societal and environmental well-being

The system is designed to support the work of radiologists and therefore lead to faster and better reporting on cases. In particular for smaller regional hospitals, the system has the capacity to improve diagnoses by providing indirect access to knowledge from experts at university hospitals, although this hypothesis remains to be tested. Overall, the system should improve the treatment of patients.

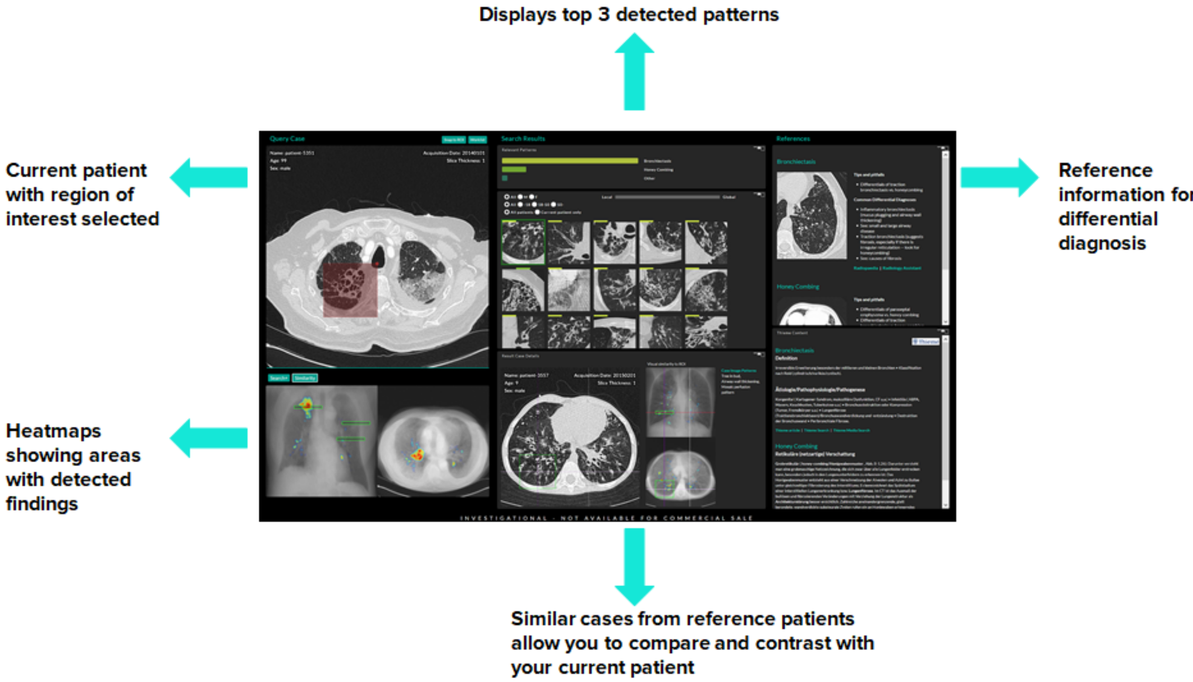
- Accountability

The first evaluation of the system in use by radiologists has been done in the BigMedilytics project and further evaluations are planned. These evaluations are in particular necessary for certification as a medical device. contextflow has a customer support process implemented by which customers can report unexpected or erroneous behaviour of the system, which will then be investigated by the contextflow support team.

2.7 System-Interaction

2.7.1 Human-Machine Interface / GUI

The following figure shows the system interface for radiologists and explanations of its components.



2.7.2 Education

Before the system begins to be used by radiologists in a clinic or hospital, a training workshop is conducted for the radiologists. All aspects of system use, as well as its capabilities and limitations, are covered in the workshop.

2.8 Other

2.8.1 Big Data Search Engines

3. Learnings

3.1 Challenges & Barriers

- Architecture

Even-though the computing environment in hospitals by now typically supports virtualization, deploying a system for complex computation on real-time diagnostic data to support assessment of individual cases, diagnosis, and treatment decision faces several challenges.

The integration is facilitated by the DICOM standard, enabling storage and communication of imaging data between hospital systems and deployed virtual machines. However, the DICOM[1] standard is complex and its implementations can have unexpected features, which can lead to delays in installing a system at hospitals.

Further challenges include the integration of systems in the workflow of clinicians, who are under tremendous time pressure, and therefore prefer single interaction points with software. This makes tight integration and collaboration among software vendors necessary.

While deploying systems on the cloud is a more economic means to support complex computation, as scalable infrastructure, and standardized, and unified security environments are available, a hesitancy to integrate cloud computing in daily processing of clinical data remains a relevant barrier.

- Processing of large structured / unstructured data sources

End users expect search results to be returned within a few seconds. As we are dealing with large 3D images, a large number of optimizations of the retrieval code were necessary to meet this requirement. This includes a smart scheduling of computation to enable preprocessing before user interaction whereever possible.

Hospitals generally do not have GPUs in their existing infrastructure, and standardized on-site GPU environements are scarcely available. This means that all processing has to be possible in a reasonable time on CPUs - an subopotimal requirement for deep learning algorithms, that had to be solved to obtain a workable solution for clinical deployment.

- Multi-velocity processing of heterogeneous data streams

N/A

- Complex real-time event detection

N/A

- Natural Language Processing

Despite efforts to move towards structured information, significant parts of record keeping and reporting in the clinical enviroment remains unstructured. This is a serious limitation to AI approaches, but can be partially remedied with help of natural language processing (NLP). While NLP systems have advanced the particular language, contextual meaning, and intent of

[1] DICOM is the international *standard* to transmit, store, retrieve, print, process, and display medical imaging information

clinical reports presents challenges to NLP algorithms, that result in limited extracted information from such reports.

- Image Processing

Challenges in medical image processing are manifold. Chief among them are the relatively high normal variability across individuals and anatomical locations, compared to often subtle signatures of disease. This makes sophisticated machine learning necessary to extract useful information.

Secondly, available data set sizes are much smaller than in standard image processing examples. As opposed to 10.000 examples, typical training set sizes often don‘t exceed a few hundred patients. This makes corresponding methodology necessary.

- Prediction Algorithms

N/A

- Security and privacy of data access and processing

Security is paramount, and access to hospital systems for installing and running on-site software, or for interfacing with systems outside of the hospital makes extensive preparation, process development, regulatory approval, and close collaboration with hospital IT providers and administrators necessary. Even with a signed data sharing agreement, it took a while to set up a data transfer pipeline.

- Trustworthy AI

Trustworthy AI is highly relevant in medicine. It is not only necessary to enable users to *double check* results of algorithms, but also to enable the integration of these results in more complex diagnostic processes.The technology - a search engine - evaluated as part of pilot 12 addresses this requirement as one of its central capabilities: it presents evidence for estimates in the form of comparable cases, and thus enables clinicians to assess estimate and evidence at once.

- System-Interaction
-

The COVID-19 pandemic delayed the evaluations of the system that we had set up with radiologists by 6 months.

3.2 Lessons Learned

- Architecture

It is necessary to involve the IT departments of hospitals deeply in preparations installation of software in a hospital.

It is a long journey, in terms of the amount of work that has to be done, from the prototype of the system to a fully functioning and certified system installed in the clinical routine.

Trusted - and possibly publicly controlled - cloud infrastructures are necessary to facilitate and accelerate the adoption of cloud based computing in health care. It is a key enabler of rapid technological advance.

- Processing of large structured / unstructured data sources

Processing load has to be carefully timed to exploit times between data acquisition and user interaction. Furthermore, aside from ML methodology, implementation details matter to enable practical use of ML in short time frames, and with limited computing hardware (no GPUs).

- Multi-velocity processing of heterogeneous data streams

N/A

- Complex real-time event detection

N/A

- Natural Language Processing

Specific training of NLP algorithms is necessary to enable the extraction of structured information from clinical data. At the same time rigorous quality control has to be implemented before information enters any model training or validation.

- Image Processing

The key ability of relevant methodology for image processing is to identify informative features and patterns associated with disease, but independent from natural variability, or even technological differences in data acquisition.

Deploying image processing algorithms is - at this time - dependent on implementation on limited hardware, typically not providing sufficiently standardized access to GPU computing resources. This requires specific considerations, and extensive optimization during the implementation of such algorithms.

- Prediction Algorithms

N/A

- Security and privacy of data access and processing

Security of the hospital computing environment is paramount, and interaction and collaboration with the responsible IT department at the earliest possible time is necessary.

Contracting can cause delays, and close and transparent communication are key to proceeding.

- Trustworthy AI

Trustworthy AI, and in particular the explainability of results are a key aspect of the system evaluated in Pilot 12. A key learning was that the need for being able to understand the algorithm results, and at the same time staying in control of the diagnostic process were consistent feedback throughout all interactions.

- System-Interaction

It is necessary to include new features into the existing workflow of the radiologists. Interactions that require switching to a different system are not well received. The time pressure requires a single point of interaction for the medical expert, and consequently tight integration of novel systems in the standard environment of the clinic.

This highlights the need for standardized interfaces, and integration capabilities, since the quality of services requires a reduction or removal of system barriers to deploy health care AI/ML/IT systems relevant for diagnosis. This is necessary to ensure that the best quality solutions reach the experts, and are not hindered by IT systems in place.

3.3 Main (quantifiable) achievements

The BigMedilytics study in Pilot 12 showed that the average reading time by radiologists is 31% shorter when they are supported by the contextflow search engine. Reading time is the amount of time required for a radiologist to perform the analysis of an image and write the report. This significantly surpasses the 20% reduction in time estimated as achievable at the beginning of the project. A trend towards better diagnostic quality was also found in the BigMedilytics study. A paper explaining these results in detail is under preparation.

The main technical achievements during BigMedilytics involved the integration of the contextflow search system with PACS (Picture Archiving and Communication Systems) as well as improvements to the search efficiency, search quality and user interface of the pre-existing contextflow search system.

4. Output

4.1 Papers

A paper presenting the results of the BigMedilytics study is under preparation.

4.2 Open Source & Resources (refer to ELG)

Due to commercial confidentiality, we are not able to make available any open source software or resources.

4.3 Demos

Below are links to videos of the contextflow system in action:
PACS integration: https://drive.google.com/file/d/1AJRkmfBADhNdVy4Kk6e2TXsujhIH96nL/view
System demo: https://drive.google.com/file/d/0B6YsWUeMdVrjUXQ1blliaGpKX0U/view
Technology explanation: https://drive.google.com/file/d/0B-uODI9ItjM5d0xnTmhkMGF4ZUk/view