

Aus dem Interdisziplinären Schlafmedizinischen Zentrum
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Möglichkeiten zur standortunabhängigen und
reproduzierbaren Biosignalanalyse in der
multizentrischen klinischen Forschung

zur Erlangung des akademischen Grades
Doctor rerum medicinalium (Dr. rer. medic.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Maximilian Beier
aus Berlin

Datum der Promotion
26.06.2022

Inhaltsverzeichnis

Abkürzungsverzeichnis	1
Kurzfassung	3
Abstract	4
1 Einführung	5
1.1 Virtuelle Forschungsumgebungen	5
1.2 Schutz von Patientendaten	6
1.3 Schlaf	6
2 Methodik	7
2.1 Anforderungen	7
2.2 Systementwurf	9
2.3 Klinisches Datenmanagementsystem	9
2.3.1 XNAT	9
2.3.2 European Data Format	10
2.4 Analyseumgebung	11
2.4.1 OpenStack und Docker	12
2.5 Benutzeroberflächen	12
2.5.1 Webtechnologien	13
2.5.2 Performance	13
2.5.3 Benutzerakzeptanztest	15
3 Ergebnisse	16
3.1 XNAT-Plugin	16
3.2 Normalisierung von EDF-Kanalbezeichnungen	17
3.3 Code-Upload	17
3.4 Daten-Upload	17
3.4.1 Performance-Tests	18
3.5 Analyse-Management	18
3.6 Inspektions-, Auswertungs- und Kollaborations-Anwendung	19
3.6.1 Parser	19
3.6.2 Visualisierer	19
3.6.3 Kollaborations-Erweiterung	20
3.6.4 Pseudonymisierung	20
3.6.5 XNAT-Integration	21
3.6.6 Performance-Optimierungen	21

3.6.7 Benutzerakzeptanztest	22
4 Diskussion und Ausblick	22
Literaturverzeichnis	26
Eidesstattliche Versicherung	31
Anteilerklärung	32
Ausgewählte Publikationen	33
Lebenslauf	76
Publikationsliste	77
Danksagung	78
Appendix	79

Abkürzungsverzeichnis

AASM	American Academy of Sleep Medicine
API	Application Programming Interface
CDMS	Clinical Data Management System
CMS	Code Management Service
CPU	Central Processing Unit
CSV	Comma-Separated Values
DGSM	Deutsche Gesellschaft für Schlafforschung und Schlafmedizin
DICOM	Digital Imaging and Communications in Medicine
DS-GVO	Datenschutz-Grundverordnung
ECAT	Dateiformat von Siemens-PET-Scannern
eCRF	electronic Case Report Form
EDF	European Data Format
EEG	Elektroenzephalografie
EKG	Elektrokardiogramm
EMG	Elektromyografie
EOG	Elektrookulografie
FAIR	Findable, Accessible, Interoperable, Reusable
FMP	First Meaningful Paint
GB	Gigabyte
GMDS	Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie
GPU	Graphics Processing Unit
HTML	Hypertext Markup Language
IT	Informationstechnik
JMS	Job Management Service
JS	JavaScript
JSON	JavaScript Object Notation
LKG	Landeskrankenhausesetz (Berlin)
LKHG	Landeskrankenhausesetz (Baden-Württemberg)
M	Mittelwert
MB	Megabyte
MD	Median
MRT	Magnetresonanztomographie
MTAF	Medizinisch-Technische Assistenten für Funktionsdiagnostik
n	Anzahl
N1	Non-REM-Schlafstadium 1 nach AASM
N2	Non-REM-Schlafstadium 2 nach AASM
N3	Non-REM-Schlafstadium 3 nach AASM

OCT	Optische Kohärenztomografie
PC	Personal Computer
PIN	Persönliche Identifikationsnummer
PSG	Polysomnographie
R	Range / Spannweite
RAM	Random-Access Memory
REM	Rapid Eye Movement
REST	Representational State Transfer
RTC	Real-Time Communication
SD	Standardabweichung
SPA	Single-Page Application
SSH	Secure Shell
SUS	System Usability Scale
SVG	Scalable Vector Graphics
TMF	Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V.
UI	User Interface
URL	Uniform Resource Locator
VM	Virtuelle Maschine
W	Wachzustand
XML	Extensible Markup Language
XNAT	Extensible Neuroimaging Archive Toolkit

Kurzfassung

Möglichkeiten zur standortunabhängigen und reproduzierbaren Biosignalanalyse in der multizentrischen klinischen Forschung

In multizentrischen, interdisziplinären Forschungsvorhaben ermöglichen sogenannte virtuelle Forschungsumgebungen gleichzeitige und standortunabhängige Zusammenarbeit, beispielsweise bei der Erfassung, Verwaltung und Auswertung von Forschungsdaten.

In Ermangelung einer passenden Lösung für die Schlafforschung wurde im Rahmen dieser Arbeit basierend auf Anforderungen von ForscherInnen, KlinikerInnen und VerfahrensentwicklerInnen eine cloudbasierte virtuelle Forschungsumgebung für Biosignaldaten entwickelt.

Den Kern bildet ein auf XNAT basierendes klinisches Datenmanagementsystem. Es ist mehrbenutzerfähig, erlaubt feingranulare Rechtevergabe und speichert sowohl Formular- als auch Signal- und deren Metadaten strukturiert ab. Über ein eigens entwickeltes Plugin werden Signaldaten im European Data Format (EDF) unterstützt und herstellereinspezifische Inkonsistenzen in EDF-Kanalbezeichnungen anhand einer umfassenden Liste normalisiert. Analysen werden automatisch bei Datenänderungen in XNAT angestoßen und über einen Job Management Service in einer auf OpenStack basierenden Cloud-Umgebung in Docker-Containern ausgeführt. Die Analyseergebnisse werden mit den Originaldaten verknüpft. Neue Verfahren können kontinuierlich hinzugefügt werden und unterliegen, um Reproduzierbarkeit zu gewährleisten, einer Versionskontrolle.

Multidimensionale Signaldaten lassen sich über eine Webanwendung einlesen, visualisieren, scoren, clientseitig pseudonymisieren und live mit einer weiteren Partei audiovisuell besprechen und auswerten. Über eine XNAT-Integration können die Daten in der Cloud-Umgebung analysiert werden und etwaige Ergebnisse werden direkt in die Oberfläche zurückgespielt. Da die Anwendung im Browser läuft, kann sie ohne Installation und spezielle Benutzerrechte ausgeführt werden.

Um die Benutzerakzeptanz der Anwendung zu testen, wurde eine Studie mit 16 Personen (14 Fachpersonal, 2 PatientInnen) durchgeführt. Nach einer Einführung in die Software füllten die ProbandInnen einen Fragebogen im dichotomen Antwortformat aus, über den die korrekte Funktionalität des Ein- und Mehrbenutzermodus geprüft und eventuelle Fehler dokumentiert wurden. Anschließend wurden die Zufriedenheit und, anhand des System Usability Scale (SUS), die Gebrauchstauglichkeit auf einer fünffach gestuften Likert-Skala angegeben. Alle ProbandInnen konnten die Anwendung verwenden, wobei zehn ProbandInnen nicht horizontal scrollen konnten, weil ihre Maus dafür nicht ausgelegt war, zwei ProbandInnen konnten den Mehrbenutzermodus nicht verwenden, was jeweils auf die Krankenhausfirewall zurückzuführen war und bei einem Probanden konnte aufgrund von Mikrofonproblemen der

Audiokanal nicht getestet werden. Es zeigt sich eine hohe bis sehr hohe Zufriedenheit mit den Teilaspekten der Anwendung und eine sehr hohe Gesamtzufriedenheit ($n = 16$; $MD = 5$; $R = 1$). Der aus den Ergebnissen des SUS errechnete Score ($MD = 86,3$; $R = 23$; $M = 87,2$; $SD = 7,7$) deutet auf eine gute bis exzellente Gebrauchstauglichkeit hin.

Im Sinne der Open-Science-Bewegung sind alle in dieser Arbeit vorgestellten Komponenten Open Source unter einer freien Lizenz. Dadurch konnten sie bereits in anderen Projekten weiterverwendet werden.

Abstract

Possibilities for Location-Independent and Reproducible Biosignal Analysis in Multicenter Clinical Research

In multicentre, interdisciplinary research projects, so-called virtual research environments (VRE) enable simultaneous and location-independent collaboration, for example in the acquisition, management and analysis of research data.

For want of a solution for sleep research, a cloud-based VRE for biosignals was developed based on the requirements of researchers, clinicians and developers.

The core is an XNAT-based clinical data management system. It is multi-user capable, allows fine-grained permission management and stores electronic case report forms as well as signal and metadata in a structured way. A plug-in supports biosignals in the European Data Format (EDF) and normalises inconsistencies in channel labels using a comprehensive list. Analyses are triggered when data changes in XNAT and executed via a job management service in an OpenStack cloud environment within Docker containers.

Analysis results are linked to the original data. New procedures can be added continuously and are version controlled to ensure reproducibility.

Via a web application multidimensional signal data can be imported, visualised, scored, pseudonymised, and audiovisually discussed and evaluated live with another party. Via an XNAT integration, the data can be analysed in the cloud environment and results are fed directly back into the interface. As the application runs in a browser, it works without installation and special user permissions.

To test the user acceptance of the application, a study was conducted with 16 participants (14 professionals, 2 patients). After an introduction to the software, a quantitative survey in dichotomous response format was used to check the correct functionality of the single and multi-user mode and to document any errors. Subsequently, satisfaction and, using the System Usability Scale, usability were evaluated on a five-point Likert scale.

All subjects were able to use the application, although ten subjects couldn't scroll horizontally because their mouse didn't support it, two subjects couldn't use the multi-user mode because of the hospital firewall, and one subject couldn't test the audio channel due to microphone problems. A high to very high level of satisfaction with the sub-aspects of the application and a very high level of overall satisfaction were found ($n = 16$; $MD = 5$; $R = 1$). The score calculated from the SUS results ($MD = 86,3$; $R = 23$; $M = 87,2$; $SD = 7,7$) indicates good to excellent usability.

All components presented in this thesis are open source under a free licence and are already being reused in other projects.

1 Einführung

Für die effiziente Durchführung klinischer Forschungsvorhaben ist die Unterstützung durch IT-Systeme heutzutage unabdingbar. Welche Software dabei zum Einsatz kommt, hängt meist von der Größe des Projekts ab, da der Betrieb und die Nutzung der Systeme selber Overhead erzeugen, also Arbeitsaufwand, der nicht direkt dem Forschungsvorhaben zuträglich ist. Kleinere Projekte, beispielsweise explorative Studien zur Erkundung einer Fragestellung und Entwicklung einer Hypothese, die über einen kurzen Zeitraum aus wenigen Personen mit kongruentem Kenntnisstand und starker Kommunikation bestehen, brauchen weniger Software. Textdaten wie Messwerte und Fragebögen werden hierbei häufig in Tabellenkalkulationsprogrammen wie Excel verwaltet und mit Massendaten wie Biosignalen oder Tomografien als Dateien auf dem lokalen Dateisystem gespeichert. Auswertungen der Daten finden manuell und ad hoc auf dem lokalen System statt.

Größere Projekte hingegen sind häufig mitarbeiterstärker, multizentrisch und interdisziplinär; das heißt, Personen verschiedener Fachrichtungen arbeiten in unterschiedlichen Rollen über mehrere Standorte hinweg zusammen. Dabei kommunizieren nicht mehr alle TeilnehmerInnen direkt miteinander und Personalwechsel sind wahrscheinlicher, was dazu führt, dass Wissen zunehmend fragmentiert.

Diese Umstände stellen besondere Herausforderungen an Forschungsinfrastrukturen.

1.1 Virtuelle Forschungsumgebungen

Um auf diese Anforderungen zu reagieren, haben sich sogenannte virtuelle Forschungsumgebungen etabliert. Sie ermöglichen standortunabhängiges, gleichzeitiges, kollaboratives Arbeiten mit geteilten Ressourcen und schaffen klare Teilprozesse entlang des Forschungsprozesses, um ihn transparent und nachvollziehbar zu gestalten [1]. Textdaten können meist über electronic Case Report Forms (eCRF) strukturiert erfasst werden [2]. Die eCRF etablierter Systeme eignen sich allerdings nur eingeschränkt, um parallel Tomografien oder Biosignale und

ihre Metadaten zu erfassen und zu speichern. Darüber hinaus bedarf es niedrigschwelliger Möglichkeiten zur Visualisierung von z.B. Tomografien wie Röntgen- oder MRT-Aufnahmen oder Biosignalen wie EEG, EMG oder EKG, da die visuelle Inspektion dieser Daten fundamentaler Bestandteil des Forschungsprozesses ist [3].

Bei der Entwicklung einer virtuellen Forschungsumgebung sollte hinsichtlich eines nachhaltig nachnutzbaren Forschungsdatenmanagements von Anfang an darauf geachtet werden, die FAIR Data Principles [4] umzusetzen. Diese fordern beispielsweise, dass Massen- und Metadaten miteinander assoziiert sind, eindeutig identifiziert und referenziert werden können und in freien, öffentlich zugänglichen Formaten gespeichert werden. Daten sollten dabei in einem definierten Schema vorliegen, in dem jeder Parameter mit unmissverständlicher Beschreibung, Datentyp, Kardinalität und Relation zu anderen Parametern definiert ist. Analyseverfahren sollten versioniert und ihre Ausführungsumgebung klar beschrieben werden, um die spätere Reproduzierbarkeit zu gewährleisten [5]. Analyseergebnisse sollten mit den Originaldaten verknüpft werden. Sobald sich Daten oder Verfahren im System ändern, sollten alle Analysen automatisch erneut ausgeführt werden, um Inkonsistenzen zwischen Ausgangs- und Ergebnisdaten zu vermeiden.

1.2 Schutz von Patientendaten

Herausfordernd ist auch der sichere und datenschutzkonforme Umgang mit Patientendaten [6]. Sollten diese beispielsweise aus dem Kliniknetz über das Internet übertragen werden, bedarf es eines Datenschutzkonzepts mit entsprechenden technischen und organisatorischen Maßnahmen zur Sicherung der Daten, z.B. Zugriffskontrollen auf den Zielsystemen und durchgehender Verschlüsselung bei Transport und Speicherung [7]. Für den Datenschutz gelten dabei (mindestens) die Datenschutz-Grundverordnung (DS-GVO), das Bundesdatenschutzgesetz und die Landesdatenschutzgesetze aller Projektteilnehmer – bei letzterem muss die jeweils strengste Regelung herangezogen werden. Beispielsweise erlaubt Baden-Württemberg über das Landeskrankenhausesgesetz die Übertragung von Patientendaten (§46 LKHG), Berlin hingegen nicht (§ 25 LKG). Das Berliner LKG schreibt auch vor, dass Patientendaten, sofern möglich, stets zu pseudonymisieren oder anonymisieren sind. Nur dann dürfen sie in einrichtungsübergreifenden Forschungsvorhaben und an Forschungsregister übermittelt werden. Nach Erwägungsgrund 26 der DS-GVO findet die DS-GVO allerdings keine Anwendung auf anonymisierte Daten [8, S. 5].

1.3 Schlaf

Der Fokus dieser Arbeit liegt auf der Unterstützung der Schlafforschung, insbesondere in der Verwaltung, Analyse und Visualisierung von in diesem Bereich üblichen Biosignalaufzeichnungen. Schlaf ist ein dynamischer Prozess, in dessen Verlauf verschiedene Biosigna-

le variieren. Die umfassende diagnostische Methode zur Messung von Schlafverlauf, -tiefe und -qualität ist die sogenannte Polysomnographie (PSG) [9]. Sie erstreckt sich in der Regel über den gesamten Nachtschlaf von etwa acht Stunden und enthält mehrere Biosignale wie Gehirnaktivität (EEG), Augenbewegung (EOG), Herzaktivität (EKG) und Muskeltonus (EMG). Das Handbuch der amerikanischen Akademie für Schlafmedizin (AASM) für das Scoring von Schlaf [10] empfiehlt für eine routinemäßige PSG mindestens 16 Biosignale zu messen, in der Praxis können Aufzeichnungen aber bis zu 40 verschiedene Signale enthalten [11]. Es existieren mehrere Regelwerke zur Benennung charakteristischer Signalmuster; die aktuell vorherrschende Nomenklatur der AASM unterscheidet zwischen wach (W), REM- (R) und Non-REM-Schlaf, der wiederum in Leicht- (N1, N2) bis Tiefschlaf (N3) unterteilt wird [12]. Dazu werden die aufgezeichneten Signale von Fachpersonal in 30-Sekunden-Fenstern, sogenannten Epochen, betrachtet und jeweils einem Schlafstadium zugeordnet. Die Interrater-Reliabilität dieser manuellen Auswertung liegt bei 82% [13, 14] – die Auswertung kann dabei durch den Computer unterstützt, aber noch nicht durch diesen ersetzt werden [15]. Zusätzlich können besondere Ereignisse markiert werden, z.B. kurzzeitiger Atemstillstand (Apnoe) oder Gliederzuckungen. Die Wechsel der Schlafstadien, die im Verlaufe des Schlafs auftreten, können grafisch als Hypnogramm dargestellt werden. Die im Hypnogramm sichtbaren Abweichungen von Normverläufen und die assoziierten Ereignisse können in der Diagnostik Hinweise auf Ursachen von Schlafstörungen geben. Da die Ursachen häufig vielfältig sind und komplexen Mechanismen der Schlafregulierung verschiedene medizinische Fachgebiete betreffen [16, 17], wird interdisziplinäre Kooperation zunehmend wichtiger [18].

2 Methodik

Die Evaluation und Auswahl geeigneter methodischer und technischer Ansätze zur Umsetzung der Forschungsplattform basieren auf den konkreten Anforderungen an diese.

2.1 Anforderungen

Aus Gesprächen mit WissenschaftlerInnen, medizinisch-technischen AssistentInnen und MedizinerInnen konnten folgende Anforderungen an die Plattform abgeleitet werden:

(a) *Electronic Case Report Forms*

Vordefinierte Prüfbögen sollten ausgefüllt und gespeichert werden können.

(b) *Arbiträre Daten*

Patientendaten sollten durch zusätzliche Erhebungen erweitert werden können.

(c) *Polysomnographien*

PSGs sollten als Dateien hochgeladen und gespeichert werden können.

(d) *Datenansicht*

Daten sollten eingesehen und durchsucht werden können.

- (e) *Zugriffsrechte*

Es sollte ein Zugriffsschutz existieren, der von DateneignerInnen feingranularer auf Eintragungsebene geregelt werden kann.
- (f) *Ortsunabhängigkeit*

Das System sollte ohne vorherige Installation auf beliebigen, mit dem Internet verbundenen Computern (Desktop- und Mobilgeräte) verwendet werden können.
- (g) *automatische Analysen*

Neu hinzugefügte Daten sollten ohne weiteres Zutun analysiert und die daraus resultierenden Ergebnisse mit den Originaldaten verknüpft werden.
- (h) *Analyseverfahren*

Neue Verfahren sollten hinzugefügt und auf bereits vorhandene Daten angewendet werden können.
- (i) *Performante Visualisierung*

Es sollte eine Benutzeroberfläche zur grafischen Darstellung von Signaldaten geben. Diese sollte unabhängig von der Größe der betrachteten Datei und auch während der Navigation durch die Darstellung performant bleiben.
- (j) *Scoring*

Die Benutzeroberfläche sollte sich an etablierten Anwendungen orientieren. Sie sollte die Auswertung einer PSG ermöglichen, insbesondere die Erstellung eines Hypnogramms und die Markierung von Ereignissen. Ereignisse sollten importiert und exportiert werden können.
- (k) *Live-Kollaboration*

Authentifizierte BenutzerInnen sollten interaktiv zusammenarbeiten und Daten austauschen können, ohne zusätzliche Tools oder Plugins installieren zu müssen.
- (l) *Schutz von Patientendaten*

Identifizierende Informationen sollten vor der Übertragung pseudonymisiert werden.
- (m) *Interaktion mit der Plattform*

Lokale Dateien sollten über eine Benutzeroberfläche in die Plattform geladen werden können und dort generierte Analyseergebnisse sollten automatisch in die Benutzeroberfläche zurückgespielt werden.

Neben den funktionalen Eigenschaften sind Usability und Verfügbarkeit die größten Herausforderungen hinsichtlich der Akzeptanz der Plattform. Alle potentiellen NutzerInnen stehen unter hoher Arbeitsbelastung und müssen ihre Aufgaben möglichst schnell erfüllen können. Arbeitsschritte, die zusätzlich hinzukommen, müssen sich daher reibungsarm in die Arbeitsroutine einfügen. Dazu gehört, dass Aufgaben an sich schnell erledigt werden können, als auch, dass die dafür benötigte Software auf möglichst vielen Rechnern zur Verfügung steht. Letzteres ist schwierig, da Systeme in Krankenhäusern aus Sicherheitsgründen meist unter eingeschränkten Benutzerrechten laufen, um die Integrität der Systeme und der darauf lie-

genden sensiblen Daten zu schützen, wodurch es dem Personal selten möglich ist, Software nachzuinstallieren [19]. Dadurch sind Spezialanwendungen, z.B. zur Betrachtung und Auswertung von PSGs, nur eingeschränkt verfügbar.

Eine weitere nicht-funktionale Anforderung ist Wartbarkeit. Forschungsprojekte sind meist zeitlich begrenzt und finanzielle Mittel können zwar für die Entwicklung von Infrastrukturen verwendet werden, nicht aber für deren Weiterbetrieb nach dem Projektende. Das führt dazu, dass Systeme, die nach Abschluss eines Projekts nicht mit möglichst geringem personellen Aufwand betrieben und gewartet werden können, tendenziell eingestellt werden.

2.2 Systementwurf

Aus den funktionalen und nicht-funktionalen Anforderungen wird die technische Herangehensweise abgeleitet und eine Systemarchitektur entworfen. In dieser teilt sich die Plattform in drei Hauptkomponenten auf: Datenmanagement, Analyseumgebung und Benutzeroberflächen. Das klinische Datenmanagementsystem (CDMS) übernimmt die Speicherung aller Nutzer- und Studiendaten, stößt bei Datenänderungen Analysen an und nimmt anschließend die Ergebnisse entgegen. Um Analysen auszuführen, werden die jeweils benötigten Daten und Verfahren an die Analyseumgebung übergeben, die nach Abschluss die Ergebnisse zurückgibt. Die Dateneingabe erfolgt über anwendungsfallsspezifische Benutzeroberflächen.

2.3 Klinisches Datenmanagementsystem

Im CDMS werden formularbasierte Studiendaten, PSGs und deren Metadaten strukturiert gespeichert und mit etwaigen Auswertungsergebnissen verknüpft. Das CDMS muss standortübergreifend erreichbar und dazu mit dem Internet verbunden sein. Um unbefugten Zugriff auf Daten zu verhindern, muss es zudem Benutzerdaten verwalten und Benutzerzugriffe autorisieren. Die verfügbaren Lösungen lassen sich in drei Kategorien einteilen: Proprietär bzw. Closed-Source, Open-Source und Eigenentwicklung [20]. Unter anderem, da laufende Kosten, besonders nach Projektende, vermieden werden müssen, können keine proprietären Systeme verwendet werden. Angesichts der Anforderungen an die Plattform und gegebener Projektmittel ist auch eine Eigenentwicklung nicht möglich. Es soll daher eine möglichst passende Open-Source-Lösung verwendet und gegebenenfalls angepasst werden. Die Wahl fiel dabei auf das Extensible Neuroimaging Archive Toolkit (XNAT).

2.3.1 XNAT

XNAT ist ein kostenloses und quelloffenes webbasiertes CDMS, das zur Unterstützung klinischer Studien in der bildbasierten Forschung entwickelt wurde und sowohl strukturierte Speichermöglichkeiten für eCRFs als auch in den Neurowissenschaften gängige Datenformate bietet [21]. In XNAT werden Daten in einer Hierarchie aus Projekten, ProbandInnen und

Experimenten organisiert. Dateien wie z.B. Tomografien und Signaldaten werden auf dem Dateisystem gespeichert. Metadaten hingegen liegen strukturiert in einer Datenbank, wodurch sich diese leicht durchsuchen lassen. Über die Web-Oberfläche von XNAT können Signaldaten hochgeladen und vorhandene Daten durchsucht, inspiziert und exportiert werden. Die FAIR Data Principles können dabei weitgehend unterstützt werden [5].

Bezüglich des Uploads gibt es nur für DICOM-Daten eine vereinfachte Lösung; alle anderen Datentypen müssen über einen längeren Weg hochgeladen werden. Soll beispielsweise eine neue Datei zu einem Probanden hinzugefügt werden, bedarf es mindestens 15 Interaktionen in Form von Button-Klicks und dem Eintragen von Bezeichnungen. Dabei kann aber z.B. nicht überprüft werden, ob ein Bezeichner schon vergeben ist und sollte das der Fall sein, muss der Schritt wiederholt werden. Der Prozess birgt zudem mehrere Fehlerquellen, durch die Daten mit falschen Zuordnungen gespeichert werden können und erfordert deshalb von BenutzerInnen ein tieferes Verständnis der XNAT-Datenstruktur.

XNAT bietet ein Benutzerverwaltungssystem mit feingranularer Zugriffskontrolle über die gespeicherten Daten. BenutzerInnen authentifizieren sich mit Zugangsname und Passwort. Es bietet zudem eine umfangreiche REST-API [22], wodurch es auch anderen Anwendungen möglich ist, mit XNAT zu interagieren. Zusätzlich enthält XNAT eine sogenannte Pipeline-Engine. Pipelines bestehen aus einer Abfolge von Aufgaben und werden in einer XML-basierten Datenstruktur beschrieben. Innerhalb von Pipelines ist es möglich, Programme aufzurufen und ihnen Parameter zu übergeben. Dadurch lassen sich mit Pipelines beliebige Datenanalysen definieren und flexibel starten.

Von Hause aus unterstützt XNAT Tomografie-Formate wie DICOM, Analyze und ECAT und lässt sich durch Plugins um weitere Formate erweitern. Dadurch können auch Polysomnographien unterstützt werden. Gängige PSG-Geräte nutzen meist herstellerspezifische Formate, die nicht offen und daher für die Forschung ungeeignet sind. Sie können aber Aufzeichnungen in das European Data Format (EDF) exportieren, weswegen EDF als Grundlage für die Plattform gewählt und dafür ein Plugin entwickelt wird [23].

2.3.2 European Data Format

EDF [24] (und sein Nachfolger EDF+ [25]) ist ein freies und offenes Datenformat und der De-facto-Standard zur Speicherung von Biosignalen. Eine EDF-Datei besteht aus drei Teilen: einem statischen Header, einem dynamischen Header und den Signaldaten. Der statische Header enthält allgemeine Informationen über die Aufzeichnung, Patientendaten und die Anzahl der aufgezeichneten Kanäle. Lediglich das Feld „local patient identification“ soll dabei identifizierende Informationen wie Name und Geburtsdatum des Patienten enthalten. Der dynamische Header enthält Informationen über jeden Kanal, z.B. den Namen und die gemessene Einheit. Darauf folgen bis zum Ende der Datei die Signaldaten, die als eine Sequenz aus

16 Bit großen Ganzzahlen gespeichert werden. Dazu werden der Aufzeichnung die Messwerte aller Kanäle, die innerhalb eines bestimmten Intervalls aufgezeichnet wurden, kontinuierlich in sogenannten Records angehängt. Dadurch sind die Signale eines jeden Kanals über die gesamte Datei verteilt. Abhängig von der Anzahl der Signale und ihrer Abtastrate sind Aufzeichnungen eines Nachtschlafs zwischen 30 MB und 2 GB groß, liegen in der Regel aber im Bereich von 300 bis 500 MB¹.

Im Zuge der manuellen Auswertung von PSGs werden alle Epochen gescored, was heißt, sie werden einem Schlafstadium zugeordnet und es werden besondere Ereignisse markiert. Für diese Informationen gibt es derzeit kein textbasiertes Standardexportformat. EDF+ erlaubt zwar ein „EDF Annotations“-Signal als separaten Kanal zum Speichern von Anmerkungen und Ereignissen, dieses ist jedoch in der Größe begrenzt, binär kodiert und an die womöglich große Signaldatei gebunden (die komplett eingelesen werden müsste, um alle Annotationen zu extrahieren), was die Arbeit damit umständlich macht. Um die Annotationsdaten in der Analyseumgebung zu speichern, wird JSON als textbasiertes Exportformat gewählt. Dadurch lassen sich Annotationen leicht zugreifen und durchsuchen. Eine Annotationsdatei enthält dabei genau ein Objekt, das wiederum eine Liste von Ereignissen als Schlüssel-Werte-Paare enthält. Der Schlüssel ist dabei jeweils die Unixzeit des Ereignisses in Millisekunden und der Wert ist ein String, der das Ereignis beschreibt.

Die EDF-Spezifikation definiert die Struktur der Metadaten und der Signale, macht aber bezüglich der Bezeichnung der Signale keine Vorgaben. Die Nachfolgespezifikation EDF+ führte Standardnamen und Namensschemata ein [27], diese sind aber nicht verpflichtend und in der Praxis wird die Standardbenennung selten eingehalten, weswegen über verschiedenen Geräte hinweg mehrere Varianten für das gleiche Signal gefunden werden können. Diese können sich leicht unterscheiden (z.B. „C4-A1“ = „EEGC4-A1“ = „EEG C4-A1“) aber auch stark (z.B. „LEGBEINLI“ = „EMG LAT“). In der klinischen Routine wird meist mit einem begrenzten Satz an Geräten gearbeitet, weswegen das Personal das Problem löst, indem es die verschiedenen Namen erlernt. Die Benennungsunterschiede sind aber bei der Arbeit in multizentrischen Studien mit einer deutlich größeren Menge an Geräten und vor allem in automatisierten Analysen problematisch.

2.4 Analyseumgebung

Für ForscherInnen ist es wichtig, dass Analysen so schnell wie möglich ausgeführt werden können, um Feedbackschleifen kurz zu halten und eine schnelle Iteration in der Weiterentwicklung von Verfahren zu ermöglichen. Da Analysen automatisiert z.B. beim Hinzufügen neuer Biosignale oder Analyseverfahren ausgeführt werden sollen, ist mit Lastspitzen zu rechnen und es erscheint nicht effizient, ForscherInnen eigene dedizierte Rechenumgebungen bereit zu

¹basierend auf Daten von <https://sleepdata.org/datasets> [26]

stellen, da diese bei Beanspruchung womöglich zu geringe Ressourcen hätten, um performant zu sein, sich die restliche Zeit aber im Leerlauf befänden und die vorhandenen Ressourcen ungenutzt blieben. Die Analyseumgebung wird deshalb vom CDMS entkoppelt, um dessen Verfügbarkeit im Falle von Lastspitzen nicht zu beeinträchtigen. Zudem werden die gegebenen Systemressourcen über eine Cloud-Umgebung abstrahiert, um sie effizient über alle NutzerInnen hinweg zu managen. Um das Risiko von Datenschutzproblemen zu reduzieren und um bezüglich der weiteren Entwicklung unabhängig zu bleiben, wird auf die Nutzung kommerzieller Cloud-Anbieter verzichtet und stattdessen eine eigene Cloud aufgesetzt.

2.4.1 OpenStack und Docker

Als Grundlage der Cloud-Umgebung dient OpenStack, ein etabliertes Open-Source-Projekt zum Betrieb von privaten Cloud-Infrastrukturen. OpenStack positioniert sich als Infrastructure-as-a-Service-Plattform, abstrahiert also physische Server und nutzt deren Ressourcen, um Services anzubieten, z.B. virtuelle Server. Diese verfügen über frei wählbare Rechenleistung und Speicherplatz, sind über ein von OpenStack verwaltetes Netzwerk zugänglich und können on-demand erstellt und auch wieder beendet werden. Dadurch kann die Last auf das System gleichmäßig und für die NutzerInnen transparent verteilt werden. Sollte sich im Laufe des Projekts abzeichnen, dass der Ressourcenbedarf das Angebot übersteigt, kann OpenStack horizontal skaliert werden, indem neue Rechner in das System aufgenommen werden. Dies ist im laufenden Betrieb möglich, die Arbeit der NutzerInnen bleibt davon also unbeeinträchtigt.

Neben virtuellen Maschinen, in denen jeweils ein komplettes Betriebssystem gestartet werden muss, bevor ein Analyseverfahren ausgeführt werden kann, unterstützt OpenStack auch Docker-Container. Diese nutzen Linux-Kernel-Virtualisierung, wodurch sie das bereits gestartete Betriebssystem des Wirt-Systems mitnutzen können. Dadurch sind sie schneller verfügbar als virtuelle Maschinen und haben einen geringeren Ressourcen-Overhead, beanspruchen also weniger Rechenzeit, die nicht direkt der eigentlichen Analyse zuzurechnen ist.

XNAT-Pipelines können nicht direkt Anwendungen in OpenStack ausführen, weswegen ein Service entwickelt werden muss, der als Vermittler zwischen beiden Systemen dient.

2.5 Benutzeroberflächen

Um anwendungsfallspezifische Oberflächen zu ermöglichen, werden CDMS und Benutzeroberflächen getrennt. Die Entscheidung leitet sich aus der Erkenntnis ab, dass das System von unterschiedlichen Benutzergruppen mit distinkten Anwendungsfällen genutzt werden wird. ForscherInnen brauchen beispielsweise eine Oberfläche zur Auswertung von Analyseergebnissen (Anforderung *d*). Dafür soll die Web-Oberfläche von XNAT genutzt werden. VerfahrensentwicklerInnen hingegen benötigen eine Oberfläche zur Integration neuer Analyseverfahren

(Anforderung *h*) und zur schnellen Evaluation dieser Verfahren mit Testdatensätzen. Zudem brauchen sie, zusammen mit KlinikerInnen, eine Oberfläche zur visuellen Inspektion, Auswertung und interaktiven Besprechung von PSGs (Anforderungen *i – m*).

Anwendungen zur Visualisierung von Biosignaldaten werden häufig in systemnahen Sprachen wie C oder C++ geschrieben, um performant zu laufen. Diese Programme müssen aber für bestimmte CPU-Architekturen und Betriebssysteme kompiliert und gegebenenfalls vor der Benutzung auf dem Betriebssystem installiert werden, wodurch sie nicht, wie in Anforderung *f* vorausgesetzt, ortsunabhängig nutzbar sind.

2.5.1 Webtechnologien

Eine Alternative stellen Webanwendungen dar. Auf quasi allen modernen Desktop-Computern und Mobilgeräten ist ein Webbrowser vorinstalliert. Ursprünglich darauf ausgelegt, Hypertext zu laden und darzustellen, wurden sie in den letzten Jahren darauf optimiert, JavaScript-Code sicher und performant auszuführen. Damit einhergehend etablierte sich das Konzept der Single Page Applications (SPA), bei denen nur noch ein einziges HTML-Dokument initial geladen wird, das wiederum den gesamten Anwendungscode nachlädt. Die Ausführung der Anwendung findet danach ausschließlich im Browser statt und Kommunikation mit dem Server ist nur noch für etwaigen Datenaustausch nötig. Moderne Browser-APIs ermöglichen es auch, lokale Dateien in einer SPA einzulesen und zu verarbeiten, wodurch es keiner Interaktion mit dem Server mehr bedarf. Dadurch lassen sich z.B. in einer EDF-Datei enthaltene Patientendaten lokal pseudonymisieren, bevor sie in das CDMS übertragen werden (Anforderung *l*). Die in allen modernen Browsern unterstützte WebRTC-API ermöglicht direkten Datenaustausch in Echtzeit mit weiteren Parteien, wodurch auch die Live-Kollaboration (Anforderung *k*) umgesetzt werden kann.

Zur Entwicklung der Oberflächen wird das Framework React verwendet. Damit können für die verschiedenen Anforderungen, wie z.B. Daten-Auswertung, Live-Kollaboration und XNAT-Integration, leicht separate Komponenten entwickelt werden. Durch die Aufteilung bleibt die Komplexität besser beherrschbar und die einzelnen Komponenten lassen sich anschließend je nach Bedarf kombinieren.

Die Umsetzung als Webanwendung hat allerdings den Nachteil, dass Anwendungsfälle wie das Einlesen und Visualisieren von Biosignalen weniger performant ausgeführt werden können, als es in nativen Anwendungen möglich wäre. Wie bereits erwähnt, ist die Performance der Anwendung aber essentiell für die Akzeptanz des gesamten Systems.

2.5.2 Performance

Wie in Anforderung *i* beschrieben, müssen die Signaldaten performant dargestellt werden. Performance umfasst dabei zwei Aspekte: Die Start- und die Aktualisierungsgeschwindigkeit.

keit. Eine übliche Metrik bezüglich des Starts der Anwendung ist der First Meaningful Paint (FMP), die Zeit vom Öffnen der Anwendung bis zum ersten relevanten Inhalt, was in diesem Fall der erste fertige Graph ist. Wir definieren eine FMP von ≤ 2 s als akzeptabel und eine FMP von 0 als ideal. Die Aktualisierungsgeschwindigkeit beschreibt die benötigte Zeit, um einen Graphen nach Benutzereingaben neu zu zeichnen, z.B. während des Scrollens auf der Zeitachse. Sie wird als Latenz bezeichnet. Wir definieren eine Latenz von ≤ 200 ms als akzeptabel. Als untere Grenze kann die Wiederholungsrate des Monitors genommen werden, die üblicherweise bei 60 Hz liegt, woraus sich eine ideale Latenz von 16 ms ergibt. Als kritisch für beide Metriken wird die Geschwindigkeit des Zeichnens der Graphen gesehen. Um diese zu optimieren, gibt es fundamental zwei Möglichkeiten: weniger Datenpunkte zeichnen und Datenpunkte schneller zeichnen. Hierzu wurde vorab untersucht, inwiefern sich die Daten verlustfrei komprimieren lassen und mit welchen Techniken Signaldaten möglichst performant im Browser visualisiert werden können [28].

Komprimierung

Gemeinhin ist die Zeitauflösung der Biosignale höher als die horizontale Auflösung des Monitors, auf denen sie betrachtet werden. Für EEG, EKG und EOG empfiehlt die AASM beispielsweise eine Abtastrate von 500 Hz, also 500 Messwerten pro Sekunde. Auf einem mit 1200 px auflösenden Monitor stehen bei Betrachtung einer Epoche (30 s) aber nur 40 px für die Daten einer Sekunde bereit. Es würden also pro Pixelreihe die Übergänge von 12 konsekutiven Messpunkten gezeichnet werden und sich dadurch zwangsläufig überlagern. Dasselbe visuelle Ergebnis ließe sich erreichen, wenn nur der Übergang vom kleinsten zum größten Wert dieser 12 Messpunkte gezeichnet werden würde. Ist die Bildschirmauflösung vor dem Einlesen der Signaldaten bekannt, könnte die in diesem Fall vorzuhaltende Datenmenge bei gleichbleibender Genauigkeit der Darstellung deutlich reduziert werden – im obigen Beispiel etwa um rund 83%. Neben der höheren Darstellungsgeschwindigkeit könnte die daraus folgende Entlastung des Arbeitsspeichers auch das Risiko eines Absturzes der Anwendung auf schwächeren PCs reduzieren. Allerdings wüchse damit auch die Komplexität, da z.B. bei Änderung der Größe des Browserfensters die Signaldaten neu verarbeitet werden müssten.

Performance-Tests verschiedener Darstellungstechniken

Die populärsten Techniken, Daten in Webanwendungen zu visualisieren, sind Scalable Vector Graphics (SVG) und das Canvas-Element. Letzteres gilt als performanter [29], was aber für den vorgestellten Anwendungsfall womöglich nicht zutrifft, da die Geschwindigkeit von der Anzahl der darzustellenden Elemente abhängt [30]. SVG könnte bei der Darstellung der Graphen schneller sein, da die Graphen meist dieselbe Anzahl an Datenpunkten enthalten und die dazugehörigen Elemente bei Änderungen wiederverwendet und verschoben werden könnten, statt sie jedes Mal komplett neu zu zeichnen, wie es mit Canvas nötig wäre.

Als Metriken für die Performance beider Lösungen dienen benötigte Rechenzeit und RAM-Nutzung. Als Simulationsgrundlage dient der finale Anwendungsfall, die Darstellung von Signalen als Zeitreihen. Dazu wurden drei Testfälle entwickelt: für das Canvas-Element, für SVG unter Verwendung des Path-Elements und für SVG unter Verwendung des Polyline-Elements. Jeder Testfall wurde zusätzlich in zwei Varianten aufgeteilt: eine für einzelne, unkomprimierte Werte und eine für Wertebereiche in Form von Min-Max-Paaren. Als Testdaten wurden im Voraus 1200 zufällige Datenpunkte pro Kanal generiert. Um etwaige unbekannte browserinterne Optimierungen abzuschwächen, wurden zwei verschiedene Datensätze generiert, aus denen für jede Testiteration eine neue zufällige Kreuzung erzeugt wurde.

Damit alle Testfälle automatisch, isoliert und reliabel getestet werden können, wurde eigens ein Testsystem basierend auf Puppeteer² entwickelt, das für jeden Testdurchlauf eine eigene Browserinstanz startet. Das Testsystem misst als Performance sowohl die Ausführungszeit von JavaScript, als auch die Zeit, die der Browser damit verbringt, Style-Änderungen zu berechnen, sie zu layouten und zu zeichnen. Jeder Testfall wurde für 8, 16, 32 und 64 Kanäle mit jeweils 100 Aktualisierungen getestet. Die Ergebnisse (siehe Abbildung 3 in [28]) zeigen deutlich, dass selbst unter der Möglichkeit, bereits gezeichnete Elemente wiederzuverwenden, Canvas durchgehend die schnellere Technik ist und die Geschwindigkeitseinbuße mit zunehmender Anzahl an Kanälen am geringsten ausfällt.

2.5.3 Benutzerakzeptanztest

Um Erkenntnisse über die Funktionalität und Gebrauchstauglichkeit der Inspektions-, Auswertungs- und Kollaborations-Anwendung zu bekommen, sowie den generellen Ansatz zu testen, die Anwendung als SPA umzusetzen und um Feedback von NutzerInnen einzuholen, wurde eine Studie zur Benutzerakzeptanz durchgeführt. Die grundlegende Frage war dabei: Funktioniert die Anwendung zufriedenstellend?

Studienteilnehmer

Die Anwendung zielt auf die Unterstützung interdisziplinärer Teams ab. Aus diesem Grund wurden ProbandInnen aus verschiedenen Fachbereichen mit Nähe zur Schlafforschung bzw. -medizin eingeschlossen. Insgesamt wurden 16 Personen befragt, sieben Frauen und neun Männer. Davon waren 14 Personen Fachpersonal und zwei ehemalige PatientInnen, da die Anwendung langfristig auch für diese Zielgruppe zugänglich sein soll. Die ProbandInnen hatten keine Vorerfahrung in der Nutzung der zu evaluierenden Anwendung.

²Ein Browser der standardmäßig ohne grafische Benutzeroberfläche startet und sich besonders für Test-Automatisierung eignet: <https://github.com/puppeteer/puppeteer>

Studiendesign

Die Daten wurden über eine quantitative Befragung erhoben. Die ProbandInnen befanden sich in verschiedenen Städten Deutschlands in einem frei gewählten Umfeld (meist daheim, teilweise in der Klinik) und haben an ihren gewohnten PCs gearbeitet. Sie bekamen initial telefonisch eine kurze Einführung zur Benutzung des Programms. Danach sollten die ProbandInnen die Anwendung im Ein- und Mehrbenutzermodus auf Funktionalität testen, indem sie eine feste Liste an Aufgaben erledigen, z.B. eine Datei öffnen und darin navigieren oder eine Audioverbindung zum Prüfer herstellen (siehe Appendix). Nach Bearbeitung aller Aufgaben wurde das Telefonat beendet. Für jede Aufgabe wurde im dichotomen Antwortformat erfasst, ob sie durchgeführt werden konnte; etwaige Fehler konnten als Freitext hinzugefügt werden. Darauf folgte die Erhebung der Zufriedenheit mit spezifischen Eigenschaften der Anwendung (Audioqualität, Geschwindigkeit, Bedienbarkeit, Datenschutz) und der Gesamtzufriedenheit. Abschließend wurde die Gebrauchstauglichkeit mit einer deutschen Version des System Usability Scale (SUS) [31] erhoben. Alle Fragen wurden über eine standardisierte fünffach gestufte Likert-Skala erfasst. Zur statistischen Auswertung der Ergebnisse wurde das Programm SPSS Statistics 20 verwendet. Aufgrund der Stichprobengröße wurden dazu nonparametrische Verfahren verwendet.

3 Ergebnisse

Es konnte eine Lösung entwickelt werden, die ForscherInnen und KlinikerInnen entlang des biosignalbasierten Forschungsprozesses unterstützt. Beginnend bei der datenschutzsensiblen Erfassung der Daten, deren standortunabhängiger Visualisierung und Auswertung samt Möglichkeit zur Live-Konsultation, über die Unterstützung in der Entwicklung der Analyseverfahren und der automatisierten Analyse der Daten samt Zurückspielen der Daten in die Visualisierung, bis hin zur vereinfachten Auswertung der Ergebnisse dank der strukturierten Speicherung und durchgehender Assoziation zusammenhängender Daten und Dateien. Aufbauend auf dem in Abschnitt 2.2 vorgestellten Systementwurf mussten dafür mehrere Komponenten entwickelt und integriert werden.

3.1 XNAT-Plugin

Es wurde ein XNAT-Plugin entwickelt, um Polysomnographien in XNAT zu unterstützen [23]. Das Plugin führt den Datentyp `psgScanData` ein, der vom bereits vorhandenen Typ `imageScanData` abgeleitet ist. Der neue Datentyp enthält die in der EDF-Spezifikation aufgeführten statischen und dynamischen Header-Felder. Diese Felder sind direkt in der nativen XNAT-Oberfläche sichtbar (siehe Abbildung 6 in [23]). Zudem wurde vom vorhandenen Typ `reconstructedImageData` der Typ `psgEDFData` abgeleitet. Dieser ist von seiner Struktur her

identisch zu `psgScanData` und wird genutzt, um für jede hochgeladene PSG automatisch eine Kopie mit normalisierten Kanalbezeichnungen anzulegen.

3.2 Normalisierung von EDF-Kanalbezeichnungen

Bezüglich des Problems divergierender Kanalbezeichnungen (siehe Abschnitt 2.3.2) wurde manuell eine Liste häufig verwendeter Varianten von Kanalnamen und deren Norm-Äquivalenten zusammengestellt. Sie umfasst derzeit 350 Varianten von 88 Standardsignalen, darunter 47 EEG-Kanäle, 11 EMG-Kanäle und 9 EOG-Positionen. Basierend auf dieser Liste stehen ein Matlab-Script zur Konvertierung von EDF-Dateien³ und eine JavaScript-Lösung (siehe Abschnitt 3.6) zur Verfügung. Mit diesen lassen sich die Kanalbezeichnungen normalisieren, wodurch Aufzeichnungen unterschiedlicher Geräte beispielsweise leichter in automatisierten Analysen ausgewertet werden können.

3.3 Code-Upload

Um die Integration neuer Verfahren in die Plattform zu erleichtern und EntwicklerInnen eine Möglichkeit zu geben, ihre Verfahren schnell zu validieren, wurde ein Code Management Service (CMS) entwickelt. Dieser stellt ein Web-Interface bereit, über das Analyseverfahren in die Plattform geladen und mit dort gespeicherten Signaldaten getestet werden können. Ergebnisse der Ausführung werden direkt im Web-Interface angezeigt. Das CMS wurde ursprünglich entwickelt, um Matlab-Scripte auf DICOM-Daten zu testen (siehe Abbildung 4.4. in [32]), kann aber mit beliebigen Daten verwendet werden. Damit sich auch Ergebnisse reproduzieren lassen, die mit einer älteren Versionen eines Verfahrens erstellt wurden, steht der Code aller hochgeladenen Verfahren unter automatischer Versionskontrolle durch Git.

3.4 Daten-Upload

Um den in Abschnitt 2.3.1 beschriebenen Daten-Upload-Prozess zu vereinfachen, Fehlerquellen zu minimieren und allgemein die Akzeptanz der Plattform zu erhöhen, wurde ein dediziertes Web-Interface zum Hochladen von Signaldaten entwickelt. Damit sind im besten Fall nur zwei Schritte nötig, um Daten in ein vordefiniertes Projekt zu laden: Dateien werden mit der Maus auf das Interface gezogen und der Upload-Button gedrückt (siehe Abbildung 5 in [23]). Dabei werden Einzeldateien unterstützt, aber auch Zip-Archive und Verzeichnisse, die automatisch nach passenden Dateien durchsucht werden. Patientennamen werden lokal durch zufällige Pseudonyme ersetzt. Diese werden lokal im Browser gespeichert, lassen sich als Sicherungsdatei exportieren und werden für identische Patientennamen wiederverwendet, um Daten konsistent zu ersetzen. Die Authentifizierung gegenüber XNAT geschieht durch

³Dagmar Krefting, <https://github.com/sommonetz/psgScan2edfData>

hinterlegte Default-Zugangsdaten und erlaubt das Hochladen in ein Default-Projekt. Es sind aber auch individuelle Logins möglich, um Daten in geschützte Projekt zu laden. Optional können der Name des/der Hochladenden und eine Erstdiagnose hinzugefügt werden.

3.4.1 Performance-Tests

In Nutzertests an verschiedenen Standorten in Deutschland konnten alle AnwenderInnen erfolgreich Dateien hochladen. Es zeigte sich allerdings, dass dies ab einer gewissen Dateigröße überlinear länger dauerte. Um den Grund herauszufinden, wurden Performance-Tests durchgeführt. Dazu wurden relevante Messpunkte entlang des Upload-Prozesses definiert und der Daten-Uploader so umgeschrieben, dass diese gemessen und automatisch an einen eigens dafür entwickelten Telemetriedienst übermittelt werden. Als Testdaten dienten vier PSG-Dateien (53 MB, 106 MB, 228 MB, 396 MB) und eine synthetische 1 kB große Datei, um ein besseres Bild dateigrößenunabhängiger Effekte zu bekommen. Als Testsystem wurden drei virtuelle Maschinen mit unterschiedlichen Mengen an RAM aufgesetzt (512 MB, 1024 MB, 2048 MB). In jeder VM wurde jede Testdatei zehn Mal hochgeladen. Die Ergebnisse zeigen (siehe Abbildung 8 in [23]), dass die Dauer des Einlesens einer Datei und des anschließenden Hochladens mit zunehmender RAM-Menge abnimmt. Da dieselbe Internetverbindung genutzt wurde, sollte zumindest das Hochladen gleich schnell bleiben. Es wird vermutet, dass dadurch, dass jeweils die komplette Datei in den RAM geladen werden muss, um die Patientendaten zu pseudonymisieren, das Betriebssystem, gerade auf Systemen mit wenig RAM, für größere Dateien Teile davon auf das Dateisystem auslagern (swappen) muss, was danach zu längeren Lesezeiten führt.

3.5 Analyse-Management

Nach dem Upload einer PSG werden XNAT-Pipelines angestoßen, um eine normalisierte Version der PSG zu erzeugen und um Analysen zu starten. Diese Analysen sollen in der OpenStack-Cloud ausgeführt werden, was aber nicht direkt aus XNAT heraus möglich ist. Um beide Systeme zu verbinden, wurde ein Job Management Service (JMS) entwickelt. Aus XNAT-Pipelines heraus können diesem über eine REST-API Aufgaben samt Parametern übergeben werden. Das JMS startet daraufhin eine Ausführungsumgebung in Form einer virtuellen Maschinen oder eines Docker-Containers in der OpenStack-Cloud und führt darin über SSH ein Shell-Script aus, das wiederum beispielsweise das jeweilige Analyseverfahren und die zu analysierenden Daten aus XNAT lädt, das Verfahren ausführt und die daraus hervorgegangenen Ergebnisse in XNAT hinzufügt. Die Ausführungsumgebung wird danach vernichtet, was sicherstellt, dass keine sensiblen Daten in der Cloud zurückbleiben.

3.6 Inspektions-, Auswertungs- und Kollaborations-Anwendung

Basierend auf der Anforderungsanalyse und dem Systementwurf wurde die Anwendung zur Inspektion und Auswertung von EDF-Dateien umgesetzt. Der Kern der Anwendung besteht aus zwei Teilen: einem Parser, der EDF-Daten einliest und in eine darstellbare Datenstruktur umwandelt und dem Visualisierer, der die geparsten Daten entgegennimmt und darstellt. Zudem wurden mit der Kollaborations-Erweiterung und der XNAT-Integration weitere Komponenten entwickelt, die an den Kern anknüpfen. Damit die Anwendung auch offline funktioniert, wird sie beim ersten Aufruf komplett geladen und mit Hilfe eines Service Workers⁴ lokal gespeichert. Eine Internetverbindung ist dann nur noch für die Kollaborations-Erweiterung nötig, oder um entfernte Dateien zu laden.

3.6.1 Parser

Der Parser konvertiert EDF-Daten in eine JavaScript-native Datenstruktur. Die über die Datei verstreut liegenden Daten werden dabei für jeden Kanal zusammengeführt (siehe Abbildung 1 in [28]) und die implizite Messzeit eines jeden Datenpunkts explizit hinzugefügt. Dadurch ist es in der Weiterverarbeitung leichter möglich, Daten eines konkreten Zeitraums zu adressieren. Allerdings verbraucht diese Lösung auch deutlich mehr Speicherplatz als die Originaldaten, weswegen es auf normalen Klinik-Rechnern mit relativ wenig Arbeitsspeicher nicht möglich wäre, eine übliche PSG von 300 MB auf einmal zu parsen. Deshalb wurde der Parser dahingehend erweitert, dass er nicht nur komplette Dateien, sondern auch Segmente einer Aufzeichnung parsen kann. Der Parser ist gegenüber die Herkunft der Daten agnostisch und unterstützt jeden Adapter, der eine asynchrone `read`-Funktion implementiert. Aktuell sind Adapter für drei Datenquellen enthalten: lokale Dateien, entfernte Dateien und WebRTC-Verbindungen. Beim Einlesen einer neuen EDF-Datei wird allen Kanalnamen automatisch ihre in Abschnitt 3.2 beschriebene Standardbenennung hinzugefügt. Hinsichtlich der in Abschnitt 2.5.2 genannten Komprimierung ist es möglich, dem Parser einen Frequenzparameter zu übergeben. Dieser beschreibt die maximale Anzahl von Werten, die pro Sekunde erwartet wird. Darauf basierend reduziert der Parser die Messwerte pro Zeitbereich auf den jeweils kleinsten und größten Wert (Min-Max-Paare), wodurch die Ergebnisdaten weniger Speicherplatz beanspruchen.

3.6.2 Visualisierer

Der Visualisierer nimmt die vom Parser aufbereiteten Daten entgegen und stellt die Signale aller Kanäle als Graphen dar. Die Benutzeroberfläche orientiert sich dabei an in Schlaflaboren etablierten Anwendungen und den Empfehlungen der AASM [12]. Sie ermöglicht die

⁴Eine Browsertechnologie, die es ermöglicht, Dateien lokal zu speichern und Serveranfragen lokal zu beantworten, wodurch eine Webanwendung auch ohne Verbindung zu einem Server verwendet werden kann.

Navigation innerhalb einer Aufzeichnung, das Wechseln zwischen gängigen Betrachtungszeitfenstern, unterstützt Scoring, also das Zuordnen von Schlafstadien zu einer Epoche, stellt das sich daraus ergebende Hypnogramm dar und erlaubt, die Metadaten einer Datei einzusehen. Als weiterer Teil der diagnostischen Auswertung können relevante Signalmuster markiert und als Ereignisse gespeichert werden (siehe Abbildung 2 in [28]). Für häufige Ereignisse wie obstruktive Apnoen oder Hypopnoen gibt es Tastaturkürzel. Lokale EDF-Dateien lassen sich direkt per Drag-and-Drop öffnen und betrachten. Die URLs von entfernt liegenden Dateien können als URL-Parameter übergeben werden.

3.6.3 Kollaborations-Erweiterung

Die Kollaborations-Erweiterung ermöglicht Echtzeit-Kommunikation mit einer weiteren Partei. Dazu muss eine Partei (Host) eine EDF-Datei öffnen und durch einen Klick auf den Teilen-Button eine PIN generieren. Diese PIN muss der zweiten Partei (Gast) über ein beliebiges Medium mitgeteilt werden. Nach dem Öffnen der Anwendung kann der Gast durch Eingabe der PIN der Sitzung beitreten. Fortan sehen beide Parteien denselben Ausschnitt der EDF-Datei und können die Anwendung normal bedienen. Ihre Aktionen werden synchronisiert und mit gleicher Priorität behandelt. Um die Kommunikation zu erleichtern, sieht jede Partei die Live-Position des Mauszeigers ihres Gegenübers als roten Punkt (siehe Abbildung 4 in [28]). Zudem können sie aus der Anwendung heraus einen Audiokanal öffnen. Es gibt keine Funktionalität, um widersprüchliches Verhalten aufzulösen, z.B. wenn sich beide Parteien gleichzeitig in verschiedene Richtungen bewegen. Solche Konflikte müssen derzeit auf sozialer Ebene gelöst werden, etwa indem die jeweiligen Absichten über den Audiokanal kommuniziert werden.

Technisch wird die Freigabefunktion als weiterer Adapter des Parsers implementiert (siehe Abschnitt 3.6.1). Wie bei den anderen Adaptern werden auch hier nur die aktuell benötigten Segmente der EDF-Datei übertragen. Statt die bereits geparsten Daten zu serialisieren, werden EDF-Rohdaten ausgetauscht, da sich dieser Weg als schneller erwiesen hat und Komprimierungsprobleme umgangen werden, wenn die Parteien unterschiedliche Bildschirmauflösungen haben. Die empfangenen Daten werden standardmäßig nicht persistiert, der Gast kann aber die gesamte Datei über einen Button herunterladen.

3.6.4 Pseudonymisierung

Um den Datenschutz zu wahren, pseudonymisiert die Anwendung standardmäßig jede Datei vor der Netzwerkübertragung lokal, indem Patientendaten durch zufällige Zeichenketten ersetzt werden. Dieses Verhalten kann durch den/die BenutzerIn deaktiviert werden. Pseudonyme werden wie beim Code Management Service (siehe Abschnitt 3.3) persistent im Browser abgelegt und für identische Benutzerkennungen wiederverwendet. Sie können jederzeit als

Excel-Datei⁵ exportiert werden und lassen sich auch im selben Format wieder importieren, z.B. um pseudonymisierte Dateien zu re-identifizieren.

3.6.5 XNAT-Integration

Die XNAT-Integration ermöglicht Authentifizierung und Datenaustausch mit XNAT. Lokale EDF-Dateien können zu XNAT hochgeladen werden, woraufhin sie automatisch in der Plattform analysiert werden [33]. Die Ergebnisse dieser Analysen, z.B. erkannte Artefakte oder Apnoen, werden direkt in die Anwendung zurückgespielt und dort neben den bereits vorhandenen Ereignissen angezeigt (siehe Abbildung 6 in [28]). Zudem können in XNAT abgelegte Dateien über einen Link direkt in der Anwendung betrachtet werden. Dabei werden wieder nur die jeweils relevanten Teile der Datei geladen und angezeigt.

3.6.6 Performance-Optimierungen

Aufgrund der Testergebnisse aus Abschnitt 2.5.2 wurden verschiedene auf Canvas basierende Bibliotheken evaluiert und sich letztlich für die Dygraphs-Bibliothek entschieden, da diese die beste Performance erzielte.

Bezüglich der auch in Abschnitt 2.5.2 beschriebenen Performance-Metriken wurden einige Optimierungen vorgenommen. So wurde, um einen schnellen FMP zu erzielen, die Fähigkeit des Parsers zum segmentierten Laden genutzt. Initial lädt die Anwendung nur die Header-Daten und die Signaldaten der ersten Epoche, woraufhin die ersten Graphen gezeichnet werden können. Im Anschluss werden die Daten für zwei Betrachtungszeitfenster links und rechts vom aktuellen Ausschnitt geladen. Fortan werden Daten in der aktuellen Bewegungsrichtung der NutzerIn vorgeladen und nicht mehr benötigte Daten verworfen.

Mit diesen Änderungen und der Nutzung von Dygraphs konnte ein FMP von rund 1100 ms erreicht werden. Das umfasst das Laden der Anwendung mit leerem Cache von einem Server (~ 440 ms), das Laden einer auf einem Server liegenden EDF-Datei (~ 450 ms) und das Rendern der ersten Epoche (~ 210 ms).

Dygraphs ist von Hause aus nicht auf das hochfrequente Neuzeichnen der Graphen optimiert. Deswegen wurden einige Änderungen an der Bibliothek vorgenommen, beispielsweise Memoisation⁶ von wiederholt berechneten Werten, Ersetzen der linearen durch eine binäre Suche beim Auflösen von Zeit- zu Datenpunkten, Verlagern der Positionierung von Elementen von der CPU auf die GPU und Pooling von HTML-Elementen, damit diese wiederverwendet

⁵Die Erfahrung zeigt, dass die meisten NutzerInnen mit diesem Format, anders als mit CSV oder JSON, effizient und ohne weitere Hilfe arbeiten können.

⁶Eine Technik um Programme zu beschleunigen, bei der die Ergebnisse aufwendiger Berechnungen zwischengespeichert und fortan für identische Ausgangswerte wiederverwendet werden.

werden können. Dadurch konnte die Anwendung auf einem durchschnittlichen PC bei konstantem Scrolling durch die Daten circa 7 mal pro Sekunde die Graphen aktualisieren, was zu einer Latenz von rund 143 ms führt. Dieser Wert ist akzeptabel ($n < 200$ ms), erreicht aber nicht die ideale Latenz von 16 ms.

3.6.7 Benutzerakzeptanztest

Im Einbenutzermodus funktionierten folgende Aufgaben bei allen ProbandInnen ($n = 16$): Öffnen einer EDF-Datei, Einsehen der Metadaten, Ändern der Epoche per Pfeiltaste und Anlegen, Verschieben und Löschen eines Events. Die Zeitauflösung konnte bei einem Probanden (6,2%) nicht verändert werden. Die Epoche konnte bei zehn ProbandInnen (62,5%) nicht per Maus geändert werden, als Grund wurde durchgehend angegeben, dass die eingesetzte Maus kein horizontales Scrollen unterstütze.

Im Mehrbenutzermodus konnten zwei ProbandInnen (12,5%) keine Zugangsdaten für eine weitere Partei erzeugen, als Grund wurde jeweils die Krankenhausfirewall angegeben. Die ProbandInnen konnten die nachfolgenden Aufgaben nicht bearbeiten und wurden daher bei deren Auswertung ausgeschlossen. Die folgenden Aufgaben funktionierten bei allen verbleibenden ProbandInnen ($n = 14$): Beitritt einer weiteren Partei, Anzeige von EDF- und Metadaten für andere Partei, Pseudonymisierung der Metadaten für andere Partei, Herunterladen der EDF-Datei, Synchronisierung von Mauszeigerposition, Epochenänderung und dem Anlegen, Ändern und Löschen von Events und Herunterladen von Events als Excel-Datei. Das Erstellen einer Audioverbindung funktionierte bei einem Probanden (7,1%) nicht, als Grund wurden Mikrofonprobleme angegeben.

Bezüglich der Benutzerzufriedenheit gaben zwei ProbandInnen keine Einstufung der Audioqualität und des Datenschutzes an. Die Ergebnisse zeigen eine sehr hohe Zufriedenheit mit der Audioqualität ($n = 14$; $MD = 5$; $R = 2$), eine generelle Zufriedenheit mit der Geschwindigkeit ($n = 16$; $MD = 4$; $R = 2$), eine sehr hohe Zufriedenheit mit der Bedienbarkeit ($n = 16$; $MD = 4,5$; $R = 1$) und dem Datenschutz ($n = 14$; $MD = 5$; $R = 2$), sowie eine sehr hohe Gesamtzufriedenheit ($n = 16$; $MD = 5$; $R = 1$).

Die Fragen des System Usability Scale wurden von allen ProbandInnen beantwortet. Die einzelnen Subskalen B1 ($MD = 3$; $R = 3$), B2 ($MD = 4$; $R = 1$), B3 ($MD = 3,5$; $R = 1$), B4 ($MD = 3$; $R = 3$), B5 ($MD = 4$; $R = 1$), B6 ($MD = 4$; $R = 2$), B7 ($MD = 4$; $R = 2$), B8 ($MD = 4$; $R = 1$), B9 ($MD = 4$; $R = 2$) und B10 ($MD = 4$; $R = 2$) spiegeln eine mittlere bis hohe Benutzbarkeit des Programms wieder. Anhand der Subskalenwerte der ProbandInnen wurde ein mittlerer SUS-Score ($MD = 86,3$; $R = 23$; $M = 87,2$; $SD = 7,7$) berechnet. Dieser Wert ($> 80\%$) ist nach Bangor et al. [34] überdurchschnittlich ($> \sim 70\%$) und deutet auf eine gute bis exzellente Gebrauchstauglichkeit hin.

4 Diskussion und Ausblick

Die geplante cloudbasierte virtuelle Forschungsumgebung für Biosignaldaten konnte erfolgreich umgesetzt werden und alle Komponenten sind miteinander integriert, wodurch ein großer Teil des Forschungsprozesses aus wenigen spezialisierten Benutzeroberflächen heraus unterstützt wird.

Durch den Code-Management-Service lassen sich Verfahren einfach in die Plattform integrieren und auf einer großen Menge an Daten testen. Dadurch können EntwicklerInnen Ergebnisse sammeln und die Verfahren anpassen. Diese Möglichkeit zur schnellen Iteration beschleunigt den Entwicklungsprozess und kann die Qualität der Verfahren erhöhen. Allerdings haben sich in der Zwischenzeit Tools wie Docker und Plattformen wie GitHub und GitLab in der Softwareentwicklung etabliert. Der Service sollte daher so umgebaut werden, dass Code nicht mehr manuell hochgeladen, sondern als Docker-Image beschrieben in einem Git-Repository verwaltet wird. Testläufe könnten dann automatisch bei Code-Änderungen initiiert und deren Ergebnisse durch den Service auf Pull-Request-Ebene als Kommentar hinzugefügt werden. Dadurch wäre weniger manuelle Arbeit durch die EntwicklerInnen nötig.

Auch der Job Management Service könnte davon profitieren, nicht mehr nur ein Docker-Image zu verwalten, in dem verschiedene Verfahren ausführbar sein müssen, sondern über GitHub- oder GitLab-Repositories oder eine Docker-Registry für jedes Verfahren angepasste Docker-Images zu laden. Durch diesen Ansatz könnte auch die im CMS integrierte Versionskontrolle der Verfahren öffentlich gemacht werden, wodurch spezifische Versionen eines Verfahrens referenziert werden könnten und sich darauf basierende Auswertungen leichter reproduzieren ließen.

Durch die sogenannte „Reproducibility Crisis“ hat Reproduzierbarkeit und deren Notwendigkeit, um wissenschaftliche Ergebnisse überprüfen zu können, wieder mehr Aufmerksamkeit bekommen [35]. Diesbezüglich konnten durch die Kombination von Code- und Datenmanagementsystem – insbesondere die Versionskontrolle der Verfahren, die Definition der Ausführungsumgebung, der Fokus auf die FAIR Data Principles und die Wahl von EDF als offenes Format für Biosignaldaten – wichtige Aspekte guter wissenschaftlicher Arbeit umgesetzt werden. Forschungsdaten lassen sich dadurch nachhaltig nachnutzen und auf diesen Daten basierende Auswertungen verlässlich reproduzieren. In darauf aufbauenden Arbeiten konnten Möglichkeiten zur Archivierung der vollständigen Ausführungsumgebung aufgezeigt werden [5]. Zukünftig sollte versucht werden, Reproduzierbarkeit nicht nur zu ermöglichen, sondern auch zu erleichtern. Dazu könnte Code – basierend auf den Erkenntnissen aus Projekten wie deep.Teaching [36] oder den SimpleITK Notebooks [37] – in Jupyter Notebooks in die Plattform eingebunden werden, wodurch nicht nur die Ergebnisse sondern auch der Entstehungsprozess biomedizinischer Auswertungsverfahren dokumentiert werden könnte.

Unter den gegebenen finanziellen Mitteln und der zur Verfügung stehenden Entwicklungszeit wurde mit XNAT eine stabile Grundlage für das Datenmanagement gewählt. Der damit einhergegangene Aufwand für Updates und Anpassungen der verbundenen Services ist immer noch deutlich geringer als wahrscheinlich für die Anpassung anderer Systeme oder gar für eine Eigenentwicklung entstanden wäre. Und da es frei verfügbar ist, wodurch keine Lizenzkosten anfallen und eigenständig gehostet werden kann, konnte es bisher auch über Projektlaufzeiten hinaus betrieben werden. Aktuell können Biosignale nur durch ein eigenes Plugin hinzugefügt werden, doch angesichts der jüngst entschiedenen Aufnahme neurophysiologischer Wellenformen in den DICOM-Standard [38] werden z.B. Polysomnographien zukünftig nativ von XNAT unterstützt, was die Wahl des Systems weiter bestärkt. Unabhängig davon hat sich bis heute leider kein CDMS etabliert, das sowohl eCRF unterstützt als auch das Management von Massendaten wie Biosignalen.

Durch die Entscheidung, die Anwendung zur Inspektion und Auswertung von EDF-Daten als SPA umzusetzen, läuft sie ohne Installation und besondere Benutzerrechte auf quasi jedem Rechner. Dadurch sind NutzerInnen unabhängiger in der Wahl ihres Standorts bzw. Arbeitscomputers. Dieser Ansatz wird nach wie vor als vielversprechend erachtet, insbesondere das clientseitige Parsen von Datenformaten wie EDF.

Vor allem aber der damit verbundene neuartige Ansatz zur client-seitigen Pseudonymisierung von Patientendaten hat sich aufgrund der daraus resultierenden Risikoreduktion von Datenschutzproblemen in Drittsystemen als großer Vorteil erwiesen. Die Funktion ist standardmäßig aktiviert, kann also nicht versehentlich vergessen werden und erfordert auch sonst kein Zutun der NutzerInnen; kann aber abgeschaltet werden, um z.B. Mehrfachpseudonymisierung zu verhindern. Wie im „Leitfaden zum Datenschutz in medizinischen Forschungsprojekten“ der TMF dargelegt, schlussfolgert Prof. Alexander Roßnagel in einem Rechtsgutachten, dass pseudonymisierte Daten anonymisierten Daten gleichgestellt sind, wenn NutzerInnen keinen Zugriff auf die Zuordnungsregeln haben [39, S. 43]. Das könnte auf die Anwendung zutreffen, da die Zuordnung von Klarnamen zu Pseudonymen nur im Browser gespeichert und nie übertragen wird. Es bliebe aber ein Restrisiko, da unberechtigter Zugriff auf die Zuordnung und damit eine Re-Identifizierung der Signaldaten nicht vollständig ausgeschlossen werden kann. Die Pseudonymisierung könnte aber auch leicht zur Anonymisierung ausgebaut werden, indem die Zuordnungen einfach nicht gespeichert werden. Dadurch wären die Daten laut TMF-Leitfaden „nicht mehr personenbeziehbar, so dass für sie auch das Datenschutzrecht samt Einwilligungsvorbehalt nicht mehr anzuwenden ist“ [39, S. 42].

Die Studienergebnisse aus den Benutzerakzeptanztest belegen die Funktionsfähigkeit des Programms in Bezug auf die gesetzten Anforderungen. So werden die Inspektions- und Auswertungsmöglichkeiten gut durch die BenutzerInnen angenommen und auch die Performance für alltägliche Arbeit wird als gut erachtet. Gesamt wurde eine gute bis exzellente Gebrauchs-

tauglichkeit bescheinigt. Fehler wie die mangelnde Mausbedienung müssen allerdings verbessert werden und auch für etwaige Störfaktoren wie zu restriktiv konfigurierte Firewalls müssen Lösungen gefunden werden.

Bezüglich der Umsetzung als SPA haben sich aber auch Nachteile aufgezeigt, vor allem hinsichtlich technischer Limitationen in den Möglichkeiten, bestimmte Funktionen in einer Webanwendung umzusetzen, die native Anwendungen nicht haben. Beispielsweise ist der Zugriff auf das Dateisystem begrenzt und sollen Dateien vor dem Hochladen verändert werden (z.B. pseudonymisiert), muss immer noch die komplette Datei eingelesen werden, was auf schwächeren Rechnern bis hin zu einem Absturz der Seite führen kann. Die Streams-API unterstützt zukünftig Fetch Upload Streaming, ein Verfahren, durch das eine Server-Anfrage gestartet werden kann, noch bevor die zu sendenden Daten vollständig vorliegen. Damit ließe sich das Problem lösen, die Funktion wird aber aktuell nur in Chrome implementiert.

Längerfristig sollten Analysemethoden wie z.B. Filter, Frequenzanalyse oder automatische Artefakterkennung implementiert werden, um die Benutzerfreundlichkeit zu erhöhen und generell die Einstiegshürde zur Nutzung dieser Verfahren zu senken. Alphonse et al. [40] haben bereits EEG-Analysemethoden als Weblösung vorgestellt und Stegman et al. konnten mit dem WebBCI-Projekt zeigen – wenngleich auch nur für einfache Verfahren – dass die EEG-Verarbeitung performant in JavaScript umgesetzt werden kann, obwohl die mathematischen Verarbeitungsmöglichkeiten der Sprache im Vergleich mit z.B. MATLAB sehr eingeschränkt sind [41].

Verschiedene Komponenten des Systems wurde auf den Jahreskongressen der GMDS (Deutsche Gesellschaft für Medizinische Informatik) und der DGSM (Deutsche Gesellschaft für Schlafmedizin) vorgestellt. Die DGSM bietet ihren Mitgliedern kostenlos eine Instanz der Inspektions-, Auswertungs- und Kollaborations-Anwendung an, um die Kommunikation zwischen ForscherInnen, KlinikerInnen und gegebenenfalls PatientInnen zu fördern.

Die meist nachgefragte Erweiterung für den EDF-Visualisierer ist eine Videounterstützung⁷, insbesondere für die Telekonsultation, um die Daten und die daraus resultierenden Diagnosen mit PatientInnen besprechen zu können, vor allem in ländlichen bzw. weniger urbanen Gebieten. Allerdings wirft die Videoverarbeitung Datenschutzprobleme auf, da in der Regel die Gesichter der PatientInnen erkennbar sind. Zudem erfordert eine Videounterstützung, um eine akzeptable Leistung zu erreichen, auch höhere Datenübertragungsraten, die in Deutschland nicht flächendeckend verfügbar sind. Außerdem werden Biosignal- und Videodaten typischerweise mit verschiedenen, nicht synchronisierten Geräten aufgezeichnet und beide Quellen zu synchronisieren ist technisch herausfordernd. Es müsste daher untersucht werden, ob und wie sich diese Funktionalität datenschutzkonform und performant im Browser umsetzen ließe.

⁷PatientInnen werden in der Regel während des Schlafs gefilmt, um die Interpretation der aufgezeichneten Biosignale zu erleichtern.

Die Unterstützung von mehr als zwei TeilnehmerInnen im Kollaborationsmodus ist möglich und vorgesehen, aber die Menge der zu synchronisierenden Daten steigt quadratisch mit der Anzahl der TeilnehmerInnen, wodurch Verbindungsprobleme mit jeder zusätzlichen Partei wahrscheinlicher werden. Daher müsste womöglich ein Rollensystem eingeführt werden, das zwischen rein zuschauenden und bearbeitungsberechtigten NutzerInnen unterscheidet, um die Stabilität der Anwendung zu gewährleisten.

Um die in Abschnitt 1.3 erwähnte Interrater-Reliabilität zu verbessern, ist geplant, die Plattform als Grundlage einer webbasierten adaptiven Lernumgebung einzusetzen – ähnlich dem AASM Inter-Scorer Reliability Program –, um medizinisch-technische AssistentInnen für Funktionsdiagnostik (MTAF) aus- und fortzubilden. Diese sollen dabei diagnostisch relevante Signalmuster in Beispieldaten erkennen und bekommen anschließend direkt einen Vergleich mit Konsensauswertungen, die, wie von Penzel et al. vorgeschlagen [42], zwischen ExpertInnen erarbeitet wurden. Das würde die Auszubildenden nicht nur unabhängiger von Zeit, Ort und Dauer ihrer Ausbildung machen; sondern wäre auch finanziell und qualitativ vorteilhaft, da die Notwendigkeit der räumlichen Nähe zwischen AusbilderIn und Auszubildenden personal- und damit kostenintensiv ist und die Qualifizierung aktuell stark von der lokalen Expertise und den lokal vorhandenen Trainingsdaten abhängt.

In Bezug auf Online-Repositories wie sleepdata.org und physionet.org, die freien Zugang zu Biosignal-Aufzeichnungen bieten, könnte die Visualisierungs-Anwendung verwendet werden, um Aufzeichnungen direkt von der jeweiligen Website aus zu betrachten, wodurch sie nicht mehr heruntergeladen werden müssten und für Interessierte leichter zugänglich wären.

Literaturverzeichnis

- [1] Michael Fraser. „Virtual Research Environments: Overview and Activity“. In: *Ariadne* 44 (06/2005). URL: <http://www.ariadne.ac.uk/issue44/fraser/>.
- [2] Anaïs Le Jeannic, Céline Quelen, Corinne Alberti und Isabelle Durand-Zaleski. „Comparison of two data collection processes in clinical studies: electronic and paper case report forms“. In: *BMC Medical Research Methodology* 14.1 (12/2014), S. 7. ISSN: 1471-2288. DOI: 10.1186/1471-2288-14-7.
- [3] Amine Naït-Ali, Hrsg. *Advanced Biosignal Processing*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. ISBN: 978-3-540-89505-3, 978-3-540-89506-0. DOI: 10.1007/978-3-540-89506-0.
- [4] Mark D. Wilkinson u. a. „The FAIR Guiding Principles for scientific data management and stewardship“. In: *Scientific Data* 3.1 (12/2016), S. 160018. ISSN: 2052-4463. DOI: 10.1038/sdata.2016.18.

-
- [5] Christoph Jansen, Maximilian Beier, Michael Witt, Sonja Frey und Dagmar Krefting. „Towards Reproducible Research in a Biomedical Collaboration Platform following the FAIR Guiding Principles“. In: *Companion Proceedings of the 10th International Conference on Utility and Cloud Computing - UCC '17 Companion*. ACM Press, 2017, S. 3–8. ISBN: 978-1-4503-5195-9. DOI: 10.1145/3147234.3148104.
- [6] Thomas Süptitz, Stephan JJ Weis und Torsten Eymann. „Was müssen Virtual Research Environments leisten? - Ein Literaturreview zu den funktionalen und nichtfunktionalen Anforderungen.“ In: *Wirtschaftsinformatik Proceedings 2013*. 01/2013, S. 21. URL: <https://aisel.aisnet.org/wi2013/21>.
- [7] Jörg Moldenhauer, Martin Haimerl, Michael Walz und Gerald Weisser. „Datenschutz und Systemsicherheit für die medizinische Informationsverarbeitung“. In: *Rechner- und sensorgestützte Chirurgie*. Gesellschaft für Informatik e. V., 2001, S. 156–164.
- [8] „Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung) (Text von Bedeutung für den EWR)“. In: *Amtsblatt der Europäischen Union* L 119 (05/2016), S. 1–88. URL: <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=celex%3A32016R0679>.
- [9] Vanessa Ibáñez, Josep Silva und Omar Cauli. „A survey on sleep assessment methods“. In: *PeerJ* (2018). ISSN: 2167-8359. DOI: 10.7717/peerj.4849.
- [10] Richard B Berry, Rita Brooks, Charlene Gamaldo, Susan M Harding, Robin M Lloyd, Stuart F Quan, Matthew T Troester und Bradley V Vaughn. „AASM Scoring Manual Updates for 2017 (Version 2.4)“. In: *Journal of Clinical Sleep Medicine* (2017). ISSN: 1550-9389, 1550-9397. DOI: 10.5664/jcsm.6576.
- [11] A Roebuck, V Monasterio, E Geder, M Osipov, J Behar, A Malhotra, T Penzel und G D Clifford. „A review of signals used in sleep analysis“. In: *Physiological Measurement* (2014). ISSN: 0967-3334, 1361-6579. DOI: 10.1088/0967-3334/35/1/R1.
- [12] Conrad Iber, Sonia Ancoli-Israel, A L Chesson und Stuart F Quan. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. American Academy of Sleep Medicine Westchester, IL, 2007.
- [13] Heidi Danker-Hopfe, Peter Anderer, Josef Zeitlhofer, Marion Boeck, Hans Dorn, Georg Gruber, Esther Heller, Erna Loretz, Doris Moser, Silvia Parapatics, Bernd Saletu, Andrea Schmidt und Georg Dorffner. „Interrater reliability for sleep scoring according to the Rechtschaffen & Kales and the new AASM standard“. In: *Journal of Sleep Research* 18.1 (03/2009), S. 74–84. ISSN: 09621105, 13652869. DOI: 10.1111/j.1365-2869.2008.00700.x.

-
- [14] Richard S. Rosenberg und Steven Van Hout. „The American Academy of Sleep Medicine Inter-scoring Reliability Program: Sleep Stage Scoring“. In: *Journal of Clinical Sleep Medicine* 09.01 (01/2013), S. 81–87. ISSN: 1550-9389, 1550-9397. DOI: 10.5664/jcsm.2350.
- [15] T. Penzel, I. Fietze und C. Veauthier. „Möglichkeiten der automatischen Schlafstadienklassifikation und ihre Grenzen“. In: *Klinische Neurophysiologie* 46.03 (08/2015), S. 128–135. ISSN: 1434-0275, 1439-4081. DOI: 10.1055/s-0035-1559649.
- [16] Thomas Penzel, Niels Wessel, Maik Riedl, Jan W. Kantelhardt, Sven Rostig, Martin Glos, Alexander Suhrbier, Hagen Malberg und Ingo Fietze. „Cardiovascular and respiratory dynamics during normal and pathological sleep“. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 17.1 (03/2007). ISSN: 1054-1500, 1089-7682.
- [17] N. Wessel, M. Riedl, J. Krämer, A. Müller, T. Penzel und J. Kurths. „Synchronisation and coupling analysis: Applied cardiovascular physics in sleep medicine“. In: *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2013, S. 6567–6570.
- [18] Mirjam Körner. „Interprofessional teamwork in medical rehabilitation: a comparison of multidisciplinary and interdisciplinary team approach“. In: *Clinical Rehabilitation* (2010). ISSN: 0269-2155, 1477-0873. DOI: 10.1177/0269215510367538.
- [19] Ross J Anderson. *Security in clinical information systems*. British Medical Association London, 1996.
- [20] Johannes Drepper und Sebastian Semler. *IT-Infrastrukturen in der patientenorientierten Forschung - Aktueller Stand und Handlungsbedarf - 2016*. Technologie- und Methodenplattform für die Vernetzte Medizinische Forschung e.V, 2017. ISBN: 978-3-89838-723-1.
- [21] Daniel Marcus, Timothy R Olsen, Mohana Ramaratnam und Randy L Buckner. „The Extensible Neuroimaging Archive Toolkit - An informatics platform for managing, exploring, and sharing neuroimaging data“. In: *Neuroinformatics* (2007). ISSN: 1539-2791, 1559-0089. DOI: 10.1385/NI:5:1:11.
- [22] Roy Thomas Fielding und Richard N. Taylor. „Architectural Styles and the Design of Network-Based Software Architectures“. Diss. 2000. ISBN: 0599871180. DOI: 10.5555/932295.
- [23] Maximilian Beier, Christoph Jansen, Geert Mayer, Thomas Penzel, Andrea Rodenbeck, René Siewert, Michael Witt, Jie Wu und Dagmar Krefting. „Multicenter Data Sharing for Collaboration in Sleep Medicine“. In: *Future Generation Computer Systems* (2017). ISSN: 0167739X. DOI: 10.1016/j.future.2016.03.025.

- [24] Bob Kemp, Alpo Värri, Agostinho C Rosa, Kim D Nielsen und John Gade. „A simple format for exchange of digitized polygraphic recordings“. In: *Electroencephalography and Clinical Neurophysiology* (1992). ISSN: 00134694. DOI: 10.1016/0013-4694(92)90009-7.
- [25] Bob Kemp und Jesus Olivan. „European data format ‘plus’ (EDF+), an EDF alike standard format for the exchange of physiological data“. In: *Clinical Neurophysiology* (2003). ISSN: 13882457. DOI: 10.1016/S1388-2457(03)00123-8.
- [26] Guo-Qiang Zhang, Licong Cui, Remo Mueller, Shiqiang Tao, Matthew Kim, Michael Rueschman, Sara Mariani, Daniel Mobley und Susan Redline. „The National Sleep Research Resource: towards a sleep data commons“. In: *Journal of the American Medical Informatics Association* (2018). ISSN: 1067-5027, 1527-974X. DOI: 10.1093/jamia/ocy064.
- [27] Bob Kemp, Alpo Värri, Thomas Penzel und Jesus Olivan. *Standard texts and polarity rules*. 2003. URL: <https://edfplus.info/specs/edftexts.html>.
- [28] Maximilian Beier, Thomas Penzel und Dagmar Krefting. „A Performant Web-Based Visualization, Assessment, and Collaboration Tool for Multidimensional Biosignals“. In: *Frontiers in Neuroinformatics* 13 (09/2019). ISSN: 1662-5196. DOI: 10.3389/fninf.2019.00065.
- [29] Daniel E Kee, Liz Salowitz und Remco Chang. „Comparing Interactive Web-Based Visualization Rendering Techniques“. In: *Tufts University, Medford, MA* (2012). URL: <https://www.cs.tufts.edu/~remco/publications/2012/InfoVis2012-Poster-RenderingTechniques.pdf>.
- [30] Tom Horak, Ulrike Kister und Raimund Dachsel. „Comparing Interactive Web-Based Visualization Rendering Techniques“. In: *Interactive Media Lab, TU Dresden, Germany* (2018). URL: <https://imld.de/cnt/uploads/Horak-2018-Graph-Performance.pdf>.
- [31] Matthias Rauer. *Quantitative Usability-Analysen mit der System Usability Scale (SUS)*. 04/2011. URL: <https://blog.seibert-media.net/blog/2011/04/11/usability-analysen-system-usability-scale-sus/> (besucht am 20.01.2021).
- [32] Christoph Jansen, Maximilian Beier, Michael Witt, Jie Wu und Dagmar Krefting. „Extending XNAT towards a Cloud-based Quality Assessment Platform for Retinal Optical Coherence Tomographies“. In: *Scalable Computing: Practice and Experience* 16.1 (02/2015). ISSN: 1895-1767. DOI: 10.12694/scpe.v16i1.1062.
- [33] Michael Witt, Christoph Jansen, Stefanie Breuer, Maximilian Beier und Dagmar Krefting. „Artefakterkennung über eine cloud-basierte Plattform“. In: *Somnologie* (2017). ISSN: 1432-9123, 1439-054X. DOI: 10.1007/s11818-017-0138-0.

- [34] Aaron Bangor, Philip T. Kortum und James T. Miller. „An Empirical Evaluation of the System Usability Scale“. In: *International Journal of Human-Computer Interaction* 24.6 (07/2008), S. 574–594. ISSN: 1044-7318, 1532-7590.
- [35] Open Science Collaboration. „Estimating the reproducibility of psychological science“. In: *Science* 349.6251 (08/2015), aac4716–aac4716. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aac4716.
- [36] Christian Herta, Benjamin Voigt, Patrick Baumann, Klaus Strohmenger, Christoph Jansen, Oliver Fischer, Gefei Zhang und Peter Hufnagel. „Deep Teaching: Materials for Teaching Machine and Deep Learning“. In: *5th International Conference on Higher Education Advances (HEAD'19)*. Universitat Politècnica València, 2019. ISBN: 978-84-9048-661-0. DOI: 10.4995/HEAD19.2019.9177.
- [37] Ziv Yaniv, Bradley C. Lowekamp, Hans J. Johnson und Richard Beare. „SimpleITK Image-Analysis Notebooks: a Collaborative Environment for Education and Reproducible Research“. In: *Journal of Digital Imaging* 31.3 (06/2018), S. 290–303. ISSN: 0897-1889, 1618-727X. DOI: 10.1007/s10278-017-0037-8.
- [38] *Digital Imaging and Communications in Medicine (DICOM) Supplement 217: Neurophysiology Waveforms*. 07/2020. URL: ftp://medical.nema.org/medical/dicom/final/sup217_ft2_NeurophysiologyWaveforms.pdf.
- [39] Klaus Pommerening, Johannes Drepper, Krister Helbing und Thomas Ganslandt. *Leitfaden zum Datenschutz in medizinischen Forschungsprojekten: Generische Lösungen der TMF 2.0*. Medizinisch Wissenschaftliche Verlagsgesellschaft, 10/2014. ISBN: 978-3-95466-295-1. DOI: 10.32745/9783954662951.
- [40] Joshy Alphonse und Shyam Diwakar. „Deploying a Web-based Electroencephalography Data Analysis Virtual Laboratory“. In: *Procedia Computer Science* 171 (2020), S. 2420–2425. ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2020.04.261>.
- [41] Pierce Stegman, Chris Crawford und Jeff Gray. „WebBCI: An Electroencephalography Toolkit Built on Modern Web Technologies“. In: *Augmented Cognition: Intelligent Technologies*. Hrsg. von Dylan D. Schmorow und Cali M. Fidopiastis. Bd. 10915. Cham: Springer International Publishing, 2018, S. 212–221. ISBN: 978-3-319-91469-5, 978-3-319-91470-1. DOI: 10.1007/978-3-319-91470-1_18.
- [42] Thomas Penzel, Xiaozhe Zhang und Ingo Fietze. „Inter-scorer Reliability between Sleep Centers Can Teach Us What to Improve in the Scoring Rules“. In: *Journal of Clinical Sleep Medicine* 09.01 (01/2013), S. 89–91. ISSN: 1550-9389, 1550-9397. DOI: 10.5664/jcsm.2352.

Eidesstattliche Versicherung

„Ich, Maximilian Beier, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Möglichkeiten zur standortunabhängigen und reproduzierbaren Biosignalanalyse in der multizentrischen klinischen Forschung“ / „Possibilities for location-independent and reproducible biosignal analysis in multicenter clinical research“, selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren/innen beruhen, sind als solche in korrekter Zitierung kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) werden von mir verantwortet.

Ich versichere ferner, dass ich die in Zusammenarbeit mit anderen Personen generierten Daten, Datenauswertungen und Schlussfolgerungen korrekt gekennzeichnet und meinen eigenen Beitrag sowie die Beiträge anderer Personen korrekt kenntlich gemacht habe (siehe Anteilserklärung). Texte oder Textteile, die gemeinsam mit anderen erstellt oder verwendet wurden, habe ich korrekt kenntlich gemacht.

Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Erstbetreuer/in, angegeben sind. Für sämtliche im Rahmen der Dissertation entstandenen Publikationen wurden die Richtlinien des ICMJE (International Committee of Medical Journal Editors; www.icmje.org) zur Autorenschaft eingehalten. Ich erkläre ferner, dass ich mich zur Einhaltung der Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis verpflichte.

Weiterhin versichere ich, dass ich diese Dissertation weder in gleicher noch in ähnlicher Form bereits an einer anderen Fakultät eingereicht habe.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§§156, 161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

Datum

Unterschrift

Anteilsklärung

Publikation 1 C. Jansen, M. Beier, M. Witt, J. Wu und D. Krefting, „Extending XNAT towards a Cloud-based Quality Assessment Platform for Retinal Optical Coherence Tomographies“, Scalable Computing: Practice and Experience, 2015

Ich habe die Qualitätssicherungsplattform anteilig konzipiert und umgesetzt, das Matlab-SSH-Docker-Image erstellt und veröffentlicht, den Code Management Service und den OCT-Migrations-Dienst konzipiert und implementiert und die Performance-Tests konzipiert, durchgeführt und ausgewertet. Daraus sind die Abschnitte zu Code Management Service, Code Upload, Usability und Performance sowie die Tabellen 3.1 und 4.2 und die Abbildungen 3.1, 4.4, 4.6 und 4.9 entstanden. Darüber hinaus habe ich am restlichen Manuskript und der Umsetzung des Reviewer-Feedbacks mitgearbeitet.

Publikation 2 M. Beier, C. Jansen, G. Mayer, T. Penzel, A. Rodenbeck, R. Siewert, M. Witt, J. Wu und D. Krefting, „Multicenter data sharing for collaboration in sleep medicine“, Future Generation Computer Systems, 2017

Ich habe die Kollaborationsplattform anteilig konzipiert und umgesetzt, den PSG-Uploader und den EDF-Visualisierer konzipiert, implementiert und veröffentlicht, die Performance-Tests konzipiert, durchgeführt und ausgewertet und Literaturrecherche betrieben. Daraus sind die Abschnitte 1, 2, 3.1, 3.4, 3.5, 4 (außer 4.6), 5 und anteilig Appendix A sowie die Tabellen 1, 2 und anteilig A.3 und A.4 und die Abbildungen 1, 4, 5, 6, 7, 8 und 9 entstanden. Zudem habe ich das Reviewer-Feedback eingearbeitet.

Publikation 3 M. Beier, T. Penzel und D. Krefting, „A Performant Web-Based Visualization, Assessment, and Collaboration Tool for Multidimensional Biosignals“, Frontiers in Neuroinformatics, 2019

Ich habe die Software konzipiert, implementiert und veröffentlicht, die Performance-Tests konzipiert, implementiert, durchgeführt und ausgewertet, die Nutzerinterviews durchgeführt und Literaturrecherche betrieben. Daraus sind alle Abschnitte und Abbildungen entstanden. Zudem habe ich das Feedback der Koautoren und der Reviewer eingearbeitet.

Unterschrift, Datum und Stempel des betreuenden Hochschullehrers

Unterschrift des Doktoranden

Ausgewählte Publikationen

Publikation „Extending XNAT towards a Cloud-based Quality Assessment Platform for Retinal Optical Coherence Tomographies“

Scalable Computing: Practice and Experience, Ausgabe 16, 2015

C. Jansen, M. Beier, M. Witt, J. Wu und D. Krefting

<https://doi.org/10.12694/scpe.v16i1.1062>

Publikation „Multicenter data sharing for collaboration in sleep medicine“

Future Generation Computer Systems, Ausgabe 67, 2017

M. Beier, C. Jansen, G. Mayer, T. Penzel, A. Rodenbeck, R. Siewert, M. Witt, J. Wu und
D. Krefting

<https://doi.org/10.1016/j.future.2016.03.025>

Publikation „A Performant Web-Based Visualization, Assessment, and Collaboration Tool for Multidimensional Biosignals“

Frontiers in Neuroinformatics, Ausgabe 13, 2019

M. Beier, T. Penzel und D. Krefting

<https://doi.org/10.3389/fninf.2019.00065>

Lebenslauf

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Publikationsliste

Ausgewählte Publikationen

- C. Jansen, **M. Beier**, M. Witt, J. Wu und D. Krefting, „Extending XNAT towards a Cloud-based Quality Assessment Platform for Retinal Optical Coherence Tomographies“, Scalable Computing: Practice and Experience, Ausgabe 16, 2015
DOI: 10.12694/scpe.v16i1.1062
- **M. Beier**, C. Jansen, G. Mayer, T. Penzel, A. Rodenbeck, R. Siewert, M. Witt, J. Wu und D. Krefting, „Multicenter data sharing for collaboration in sleep medicine“, Future Generation Computer Systems, Ausgabe 67, 2017
DOI: 10.1016/j.future.2016.03.025
Journal Impact Factor 2017: 4.639
- **M. Beier**, T. Penzel und D. Krefting, „A Performant Web-Based Visualization, Assessment, and Collaboration Tool for Multidimensional Biosignals“, Frontiers in Neuroinformatics, Ausgabe 13, 2019
DOI: 10.3389/fninf.2019.00065
Journal Impact Factor 2019: 2.649

Zusätzliche Publikationen

- C. Jansen, **M. Beier**, M. Witt, S. Frey und D. Krefting, „Towards Reproducible Research in a Biomedical Collaboration Platform following the FAIR Guiding Principles“, Companion Proceedings of the 10th International Conference on Utility and Cloud Computing, 2017
DOI: 10.1145/3147234.3148104
- M. Witt, C. Jansen, S. Breuer, **M. Beier** und D. Krefting, „Artefakterkennung über eine cloud-basierte Plattform“, Somnologie, Ausgabe 21, 2017
DOI: 10.1007/s11818-017-0138-0

Konferenzen

Ergebnisse der ausgewählten Publikationen wurden auf der CCGrid 2014 und 2015, der GMDS 2014, 2015 und 2018 und der DGSM 2017, 2018 und 2019 präsentiert.

Danksagung

Prof. Dr. Thomas Penzel danke ich für die Möglichkeit, an der Charité promovieren zu können und die hervorragende fachliche Unterstützung.

Prof. Dr. Dagmar Krefting möchte ich besonders danken – all dies war nur dank ihres unermüdlichen Engagements, Vertrauens, ihrer Motivation und steten Unterstützung möglich. Eine bessere Betreuung hätte ich mir nicht wünschen können.

Ich danke meinen Kollegen, vor allem Christoph Jansen, Michael Witt und Maryna Khvastova, für ihre Hilfe und die großartige Zusammenarbeit.

Danke auch an die DFG, das IFAF, die Stiftung Charité und den Open-Access-Publikationsfonds der Charité für die finanzielle Förderung.

Zu guter Letzt danke ich meiner Familie, insbesondere meiner Ehefrau Marie, für den Rückhalt und die Geduld.

Nutzerakzeptanztest des EDF-Visualisierers

Teilnehmerdaten

Testnummer: Testnummer Datum: Datum Uhrzeit: Uhrzeit

Name: Vorname Nachname Fachrichtung: z.B. Internist / Psychologin

Institut: Institut / Klinik / Unternehmen

Fragen zum Einbenutzermodus

E1. Eine EDF-Datei konnte geöffnet werden ja nein

etwaige Notizen

E2. Die Daten aller Kanäle der ersten Epoche wurden visualisiert ja nein

etwaige Notizen

E3. Die Metadaten der Datei konnten eingesehen werden ja nein

etwaige Notizen

E4. Die Epoche konnte per Pfeiltasten geändert werden ja nein

etwaige Notizen

E5. Die Epoche konnte per Maus geändert werden ja nein

etwaige Notizen

E6. Die Zeitauflösung konnte geändert werden ja nein

etwaige Notizen

E7. Es konnte ein Event angelegt werden ja nein

etwaige Notizen

E8. Das Event konnte verschoben werden ja nein

etwaige Notizen

2 | Nutzerakzeptanztest

E9. Das Event konnte gelöscht werden ja nein

etwaige Notizen

Fragen zum Mehrbenutzermodus

M1. Es konnten Zugangsdaten für eine weitere Partei erzeugt werden ja nein

etwaige Notizen

M2. Eine weitere Partei konnte der Sitzung beitreten ja nein

etwaige Notizen

M3. Es konnte eine Audioverbindung hergestellt werden ja nein

etwaige Notizen

M4. Die EDF-Daten sind für die andere Partei sichtbar ja nein

etwaige Notizen

M5. Die Metadaten der Datei konnten durch die andere Partei eingesehen werden ja nein

etwaige Notizen

M6. Die Metadaten sind für die andere Partei pseudonymisiert ja nein

etwaige Notizen

M7. Die EDF-Datei konnte durch die anderen Partei heruntergeladen werden ja nein

etwaige Notizen

M8. Die Mauszeigerpositionen wurden synchronisiert ja nein

etwaige Notizen

M9. Das Ändern einer Epoche wurde synchronisiert ja nein

etwaige Notizen

3 | Nutzerakzeptanztest

M10. Das Anlegen eines Events wurde synchronisiert ja nein

etwaige Notizen

M11. Das Verschieben eines Events wurde synchronisiert ja nein

etwaige Notizen

M12. Die Events konnten als Excel-Datei heruntergeladen werden ja nein

etwaige Notizen

M13. Die Excel-Datei enthielt alle Events ja nein

etwaige Notizen

M14. Das Löschen eines Events wurde synchronisiert ja nein

etwaige Notizen

Fragen zur Zufriedenheit

Z1. Wie zufrieden sind Sie mit der Audioqualität der Verbindung? Bitte wählen

etwaige Notizen

Z2. Wie zufrieden sind Sie mit der Geschwindigkeit der Anwendung? Bitte wählen

etwaige Notizen

Z3. Wie zufrieden sind Sie mit der Bedienbarkeit der Anwendung? Bitte wählen

etwaige Notizen

Z4. Wie zufrieden sind Sie mit dem Datenschutz der Anwendung? Bitte wählen

etwaige Notizen

Z5. Wie zufrieden sind Sie insgesamt mit der Anwendung? Bitte wählen

etwaige Notizen

4 | Nutzerakzeptanztest

Fragen zur Benutzbarkeit (System Usability Scale)

B1. Ich kann mir sehr gut vorstellen, das System regelmäßig zu nutzen. Bitte wählen

etwaige Notizen

B2. Ich empfinde das System als unnötig komplex. Bitte wählen

etwaige Notizen

B3. Ich empfinde das System als einfach zu nutzen. Bitte wählen

etwaige Notizen

B4. Ich denke, dass ich technischen Support brauchen würde, um das System zu nutzen. Bitte wählen

etwaige Notizen

B5. Ich finde, die verschiedenen Funktionen des Systems sind gut integriert. Bitte wählen

etwaige Notizen

B6. Ich finde, dass es im System zu viele Inkonsistenzen gibt. Bitte wählen

etwaige Notizen

B7. Ich kann mir vorstellen, dass die meisten Leute das System schnell zu beherrschen lernen. Bitte wählen

etwaige Notizen

B8. Ich empfinde die Bedienung als sehr umständlich. Bitte wählen

etwaige Notizen

B9. Ich habe mich bei der Nutzung des Systems sehr sicher gefühlt. Bitte wählen

etwaige Notizen

B10. Ich musste eine Menge Dinge lernen, bevor ich mit dem System arbeiten konnte. Bitte wählen

etwaige Notizen

5 | Nutzerakzeptanztest

Fehlerprotokoll

F1. Traten Fehler auf und wenn ja, welche?

ja nein

Kommagetrennte Aufzählung aller Fehler

Zusätzliche Anmerkungen

z.B. Kritik, Anregungen