

Teil A

1. Inhaltliche Kurzbeschreibung des Vorhabens

Das Projekt AIStudyBuddy nutzt moderne KI-Technologien, um die individuelle Planung und Reflexion von Studienverläufe zu unterstützen. Im Fokus stehen zwei Zielgruppen: Studierende erhalten mit dem StudyBuddy ein Tool zur informierten, evidenzbasierten Planung des Studiums. Es bietet grafische Repräsentationen des Studienfortschritts und gibt handlungsleitendes Feedback. Dieses stützt sich auf regelbasierte Studienverlaufspläne sowie durch KI-Technologie ermittelte Verlaufsprofile, die zu erfolgreichen Studienabschlüssen führen. Generische Studienpläne werden so durch ein Werkzeug zur individuellen Studienplanung ergänzt, die kontinuierlich angepasst, begründet, reflektiert wird. Studiengangsdesigner erhalten mit BuddyAnalytics ein interaktives Tool, das Planungsentscheidungen wie die kompetenzorientierte Curriculumsentwicklung und Studienberatung unterstützt. Das Projekt kombiniert KI-Paradigmen der datengestützten (Process Mining) und regelbasierten KI (Answer Set Programming, ASP). Mit Process Mining wird das Studienverhalten anhand der Daten aus Campus-, Lernmanagement- und Prüfungssystemen analysiert. Es stellt reale Studienverläufe den intendierten gegenüber. Mithilfe von Answer Set Programming werden transparente Begründungen für Feedback generiert, das für Nichtdomänenexperten verständlich ist. Alle Komponenten sind Teil einer Referenzarchitektur, die Prinzipien wie Ethics-by-Design und Privacy-Preservation folgt. Sie wird von den Verbundpartnern kollaborativ entwickelt, pilotiert, evaluiert und iterativ optimiert. Strategisches Ziel ist der standardisierte Austausch von Studiendaten als erster Schritt zum hochschulübergreifenden Studienmonitoring. Das Projekt widmet Determinanten von Akzeptanz bei der Einführung KI-basierter Unterstützungstechnologien besondere Aufmerksamkeit. Im Ergebnis soll die Kombination aus evidenzbasiertem Studienmonitoring, interaktiver Studienverlaufsplanung und datengestütztem Curriculumsdesign zu erfolgreicherem Studienverläufen und Absolvent:innen führen.

Schlüsselwörter: Studienmonitoring | Studienverlaufsplanung | Process Mining | Regelbasierte KI | Studierendendiversität | Technikfolgenabschätzung | Wirksamkeit

2. Darstellung der Ausgangslage

Lehren, Lernen und Prüfen finden im Kontext rasanten akademischen und gesellschaftlichen Wandels statt. Kennzeichnend hierfür ist die Digitalisierung sowohl der Kompetenzvermittlung mitsamt ihrer Prüfung als auch der Verwaltung von Studienprozessen. Studierende erzeugen dabei immer mehr Daten in den digitalen Systemen ihrer Hochschulen, darunter Moduleinschreibungen, Prüfungsanmeldungen und Leistungsartefakte in Campus-, Lernmanagement- oder Prüfungssystemen (Ifenthaler, 2020). Solche Daten werden im Kontext von Learning Analytics u.a. mit Methoden des Maschinellen Lernens und Data Mining analysiert (Chatti et al., 2012). Zu den Zielen gehören die Klassifikation von Studierenden, Inhalten und Interaktionen (Scheffel et al., 2012), die Analyse von Lernstrategien (Matcha et al., 2019) oder sozialer Lernnetzwerke (Fournier, Kop & Sitlia, 2011) sowie die Lernendenmodellierung (Abyaa et al., 2019). In Deutschland existieren Projekte wie LISA (Fortenbacher, Pinkwart & Yun 2017), STELA (Ifenthaler, 2017), FragSte (Berens et al., 2019) oder LerSys (HTWK Leipzig, 2018). Dabei sind europaweit klare Trends und Herausforderungen bei der Einführung von Learning Analytics an Hochschulen erkennbar (Tsai et al., 2020).

Hochschulen beginnen aktuell damit, das Potenzial KI-basierter Verfahren für Lehr-/Lernprozesse nicht nur in Forschungsprojekten, sondern auch im realen Einsatz zu erproben (Lukarov, 2018; Berens et al., 2019). Im Fokus steht dabei meist die Betrachtung von Lern- und Prüfungsprozessen (Zawacki-Richter et al., 2020). Weniger Beachtung findet die Analyse übergreifender Studienverlaufsdaten, für die Daten aus verschiedenen Hochschulsystemen zusammengeführt werden (Baumann et al., 2015, 2020). Individuelle Studienberatungen, Studienverlaufsempfehlungen oder die

Gestaltung von Studiengängen erfolgen nur eingeschränkt evidenzbasiert, da belastbare Daten zum Studienmonitoring und Entscheidungsregeln fehlen.

Gleichzeitig erleben Hochschulen eine soziodemografische Diversifizierung. Studierende unterscheiden sich immer stärker in ihren Eingangsprofilen, Bildungsbiografien und Lebenssituationen. Ebenso ermöglichen Hochschulen eine zunehmende Vielfalt im Studienangebot und den Ausgestaltungsmöglichkeiten in Studiengängen. Wahlbereiche, interdisziplinäre Angebote, zusätzliche Leistungen, kooperative Studiengänge und die Anerkennung extern erbrachter Leistungen führen zu immer weniger idealtypischen Studienverläufen. Folglich werden die Wege zum Studienabschluss vielfältiger und komplexer. Generische Studienverlaufspläne können individuelle Planungsbedarfe oftmals nicht mehr zufriedenstellend abbilden. Auch innerhalb der Studierendenschaft sinkt die Bedeutung traditionell vorhandenen impliziten Wissens über mögliche Studienverläufe. Individuelle Erfahrungswerte sind oft nicht mehr zwischen Studierenden übertragbar und Hochschulen beobachten eine zunehmende Disaggregation der Anfängerkohorten im Laufe der Zeit. Im Ergebnis haben Studierende nur noch eingeschränkte Handlungsoptionen in den drei Kernbereichen der Studienverlaufsgestaltung: (i) Planung des eigenen Studienverlaufs, (ii) Erkennung ungünstiger Entwicklungen und (iii) zielorientierte Anpassung ihres Studienplans.

Das Projekt verfolgt das Ziel, den Fortschritt moderner KI-Technologien nutzen, um Chancengerechtigkeit umzusetzen und erfolgreiche diverse Studienverläufe zu unterstützen. Studienplanung soll dezentralisiert, jederzeit zugänglich und änderbar sein. Studierende erhalten einen Raum, individuelle Wege durch ihr Studium zu finden und auch Irrwege zu erkennen und zu korrigieren. Hierzu bestehen an allen beteiligten Hochschulen bereits Vorarbeiten. So wird an der Bergischen Universität Wuppertal (BUW) das KI-basierte Frühwarnsystem FragSte für drohenden Studienabbruch für zwei Hochschulen sowie das Leistungscockpit ASSIST entwickelt, welche Feedback zur Studiensituation geben. An der Ruhr-Universität Bochum (RUB) wurden in den vergangenen Jahren unterschiedliche Tools und Plattformen wie der Studiport und die so genannten RUB-Checks entwickelt, die evidenzbasiert Orientierung und Unterstützung im frühen Studienverlauf bieten. An der RWTH Aachen (RWTH) wird an einer zentralen Learning Analytics Infrastruktur gearbeitet, welche über Schnittstellen zu verschiedenen Hochschulsystemen Daten im Datenstandard *Experience API* (xAPI) in einem *Learning Record Store* speichert. Die Infrastruktur erlaubt das Zusammenführen verschiedener Studierendendaten und daraus resultierend plattformübergreifende Analysen und Maßnahmen zur Kohortenverfolgung. RUB und RWTH sondieren und erproben im Projekt KI:edu.nrw u. a. die systematische Nutzung von Daten verschiedener Systeme für gezielte didaktische Interventionen in Studienprozessen. Der Projektverbund aus RWTH, RUB und BUW bietet eine einzigartige Möglichkeit, die Effekte und Potenziale der Studierendendiversität in der Ausdifferenzierung von Studienverläufen an drei Hochschulen mit unterschiedlichen Schwerpunkten zu betrachten. Die Kombination aus einer hohen Zahl an Studienverläufen und der Berücksichtigung eines breiten Spektrums an Merkmalen, Fächern, Settings und Technologien schafft eine hervorragende Datengrundlage für reliable, valide, generalisierbare und transferierbare Erkenntnisse. Eine der Stärken des interdisziplinären Projektverbunds liegt im kontinuierlichen, fest in der Projektstruktur verankerten Transfer von Methoden und Technologien zwischen den Verbundpartnern mit ihren komplementären Besonderheiten der Studierendenschaften, Studienangebote, Lehr-Lern-Arrangements und Infrastrukturen. Initial werden an den drei Hochschulen neun Fakultäten als Pilotpartner eingebunden; im nachfolgenden Regelbetrieb werden sukzessiv weitere Fakultäten erschlossen. So können standort-, organisations- und fachspezifische Bedarfe schnell identifiziert und berücksichtigt werden und die Übertragbarkeit zwischen Hochschulen kann pilotiert werden.

3. Ableitung und Beschreibung der Projektidee und der Projektziele

Das Projekt fokussiert auf zwei Zielgruppen: (i) **Studierende** und (ii) **Studiengangsgestalter** (z.B. Curricula-Planer:innen, Kommissionen für Lehre, Gremien, Dekanate). Darüber hinaus wer-

den strategische Stakeholder wie Hochschuladministrationen und bildungspolitische Akteure adressiert. Zu diesem Zweck wird eine KI-basierte Systemarchitektur mit zwei verschiedenen Zugängen konzipiert und implementiert: (i) **StudyBuddy** für Studierende und (ii) **BuddyAnalytics** für Studiengangsdesigner. Die Basis bildet eine Modellierung der Studien- und Prüfungsordnungen, um die Regeln und Bedingungen zum Studienabschluss in Form von Verlaufsgraphen zu formalisieren. Zusätzlich sollen Prozess- und Verlaufsdaten aus Hochschulsystemen wie Campusmanagementsysteme (CaMS), Lernmanagementsysteme (LMS), Prüfungssysteme (PS) und Nutzungsdaten aus dem zu entwickelnden StudyBuddy selbst verwendet werden. So entsteht eine selbstreferenzielle und kontinuierlich aktualisierende Datenbasis zur KI-basierte Entscheidungsunterstützung.

Studierende, die als kompetente Gestalter:innen ihres Studienverlaufs aufgefasst werden, sollen durch einen intelligenten **StudyBuddy** in der Planung und Reflexion des eigenen Studienverlaufs unterstützt werden. Dazu werden innovative und bewährte KI-Technologien der zwei KI-Paradigmen *Process Mining* (datengestützte KI) und *Answer Set Programming* (ASP; regelbasierte KI) kombiniert (vgl. Abb. 1 in Kap. 6). Durch grafische Repräsentation von zu belegenden Modulen abhängig vom aktuellen Studienfortschritt erhalten Studierende konkrete handlungsleitende Informationen und Empfehlungen über verfügbare Optionen, wie die Auswahl belegbarer Module, deren Abhängigkeiten (sog. "Constraints" im Sinne der KI) und formalen Voraussetzungen. Sie erhalten Beispiele dafür, wie sich Studierende mit ähnlichen Verlaufsprofilen und Studiensituationen verhalten, um zum erfolgreichen Studienabschluss zu gelangen. Gegenüber der klassischen Planung in CaMS wird also nicht nur die aktuelle Situation und die Veranstaltungsanmeldungen des kommenden Semesters in den Blick genommen, sondern individuelle, vollständige Studienverläufe können kontinuierlich angepasst, begründet und reflektiert werden.

Für **Studiengangsdesigner** (sowie Hochschulgremien und Fachstudienberatungen) entsteht mit den **BuddyAnalytics** ein Werkzeug, das die eingeschränkte Evidenz darüber, wie Studierende durchs Studium gehen und welche Kombinationen und Reihenfolgen von Modulen sinnvoll sind, durch empirisch gewonnene Erkenntnisse unter Einsatz von Process Mining ersetzt. Durch eine Analyse der Daten verschiedener Hochschulsysteme und interaktive Visualisierung können Indikatoren für Hochschulbildung generiert werden. Mit diesen können Erkenntnissen für zukünftige Planungsentscheidungen, wie Curriculumsentwicklung oder Studienberatung, gewonnen werden.

Des Weiteren sollen die Ergebnisse des Process Mining in das regelbasierte KI-System einfließen. Damit basieren Empfehlungen des StudyBuddy sowohl auf formalen wie auch Erfahrungsdaten, also „weichen“ Regeln. Die Übersetzung dieser aufgedeckten Regeln in ein regelbasiertes Modell ermöglicht deren Darstellung in einer für Nichtdomänenexperten verständlichen Form und unterstützt die Weiterentwicklung von Studiengängen.

Studierende können dadurch den Möglichkeitenraum explorieren, indem sie z.B. den Umfang der zu planenden Credits spezifizieren (z.B. wegen Nebentätigkeit), ihre Planungen auf Kommiliton:innen (*Critical Peers*) abstimmen oder Auslandssemester einbeziehen. Durch die Kombination beider KI-Technologien führt StudyBuddy unterschiedliche Informationen zusammen: (1) in der Prüfungsordnung und dem Modulhandbuch definierte Regeln, welche Veranstaltungen gewählt werden können bzw. müssen, um dem Abschluss näher zu kommen, (2) das in der Universität oft nur informelle und anekdotische Wissen, welche Module und welche Kombination und Reihenfolge von Modulen sinnvoll und erfolgversprechend sind sowie (3) Verweise auf situativ passende Unterstützungsangebote. Dieses Zusammenspiel ermöglicht eine hochqualitative individuelle Planungsunterstützung, die ohne gleichzeitige Berücksichtigung beider KI-Technologien auf diese Weise automatisiert nicht möglich wäre.

Für Studierende entsteht mit dem StudyBuddy somit ein Werkzeug zur Planung des Studiums über mehrere Semester in die Zukunft. Durch die Modellierung der formalen und informellen Regeln des Studiengangs ist zu jedem Planungsschritt unmittelbares Feedback und ein Überblick über die Konsequenzen der aktuellen Planung möglich. So können bspw. auch die für Menschen schwer zu

überblickenden Folgen der Planung für die nächsten Semester sofort erkannt und rückgemeldet werden. Das könnte eine erfahrungsgemäß ungünstige Reihenfolge zweier Module in einem Jahr, die aber durch die aktuelle Planung und den Wunsch nach Regelstudienzeit erzwungen wird, sein. Die Studierenden werden durch StudyBuddy nicht in ihrer Wahlfreiheit eingeschränkt, sondern im Gegenteil zu informierten und evidenzbasierten Entscheidungen befähigt. Um langfristig die Veränderung in Studien- und Prüfungsordnungen im StudyBuddy zu integrieren wird ein Prozess mit Werkzeugunterstützung für die Studiengangsdesigner etabliert.

Studiengangsdesigner gewinnen mit BuddyAnalytics die Möglichkeit zur evidenzbasierten Curriculumsentwicklung. Die Erschließung der Studierendendaten verschiedener Hochschulsysteme und der Nutzungsdaten des StudyBuddy ermöglicht die Betrachtung von realen Studienverläufen und dem Planungsverhalten der Studierenden. Dadurch können Probleme und Diskrepanzen intendierter Studienverläufe aufgedeckt und Interventionen angeregt werden.

Für die Etablierung von StudyBuddy und BuddyAnalytics kommt der Akzeptanz entscheidende Bedeutung zu. Das Projekt widmet daher den Determinanten von **Akzeptanz** (Usability, Ethik, didaktisches Design) bei der Einführung KI-basierter Unterstützungstechnologien besondere Aufmerksamkeit. Es wird im Projekt geprüft, mit welchen technischen und didaktischen Mitteln gerechtfertigtes Vertrauen und Zustimmung für das Erheben und Auswerten personalisierter bzw. pseudonymisierter Daten gesichert werden kann. Dies soll durch eine Referenzarchitektur sichergestellt werden, die Prinzipien wie *Ethics-by-Design* und *Privacy-preserving Data Publishing* berücksichtigt. Darin macht eine Consent Management App transparent, a) welche Verfahren zu welchem Zweck pseudonymisierte Daten nutzen und b) wem diese in welcher Form präsentiert werden. Die Gelingensbedingungen der Einführung der genannten Werkzeuge werden durch Einbindung von Forschung zu ethischen Risiken (Technikfolgenabschätzung) und didaktischen Designs sichergestellt, so dass Akzeptanz, Nutzen und Wirksamkeit in unterschiedlichen Studiensituationen wichtige Gestaltungskriterien bilden.

Darüber hinaus soll das Projekt den **Austausch von Studierendendaten zwischen den Hochschulen** zur Nutzung des Process Mining in der Referenzimplementierung erproben und erste Schritte zum hochschulübergreifenden Studienmonitoring durch eine Clearingstelle sondieren.

4. Methodisches Vorgehen und Zusammenarbeit

Startpunkt des Vorhabens sind die funktionalen Vorarbeiten an den drei Hochschulen. Die Umsetzung des Vorhabens erfolgt inkrementell durch aufeinanderfolgende PDCA-Zyklen. Arbeitspakete werden geplant, umgesetzt, evaluiert und bilden den folgenden Ausgangspunkt. Für die Zusammenarbeit zwischen den Hochschulen und die Bereitstellung von Analyseservices sowie das Lernen aus großen Datenbeständen aller beteiligten Hochschulen erfolgt zunächst die gemeinsame Definition eines Referenzmodells von Studienfortschrittsdaten. Das Referenzmodell wird schrittweise weiterentwickelt, zunächst für exemplarische Fakultäten bzw. Modellstudiengänge an allen Standorten, darunter Informatik (RWTH, BUW), Wirtschaftswissenschaften (RWTH, RUB, BUW), Maschinenbau (RWTH, RUB, BUW) und Germanistik (BUW). Die nutzbaren Datenquellen werden nach und nach erschlossen. Für jede neue Datenquelle werden zuerst die datenschutzrechtlichen Aspekte geklärt, dann die Verfügbarkeit in einem Data Warehouse sichergestellt sowie *privacy-preserving* Schnittstellen implementiert. Der Einbindung in das Referenzmodell geht stets die Abschätzung der Technikfolgen und didaktischen Mehrwerte voraus.

KI-basierte Empfehlungen im StudyBuddy werden zunächst auf einer hohen Aggregatsebene (Studienfortschritt allgemein) gegeben und mit jeder neuen Datenquelle präzisiert. Dabei wird in allen Entwicklungsschritten jeweils nutzerzentriert entwickelt, was insbesondere die formative Evaluation der Wirksamkeit sowie eine kontinuierliche Reflexion über Akzeptanz, Vertrauenswürdigkeit und ethische Fragestellungen umfasst. In ausgewählten Studiengängen sollen experimentelle Settings realisiert werden, bei denen unter Beachtung der Variationen zwischen den Hochschulen und Studiengängen bestimmte Studierendengruppen andere Systemfunktionalitäten erhalten. Zugleich

wird mit jeder weiteren Datenquelle auch der Umfang der Analysen und die bereitgestellte Funktionalität der BuddyAnalytics erweitert. Auch diese Entwicklung erfolgt in Abstimmung mit den verantwortlichen Studiengangsdesignern formativ evaluierend.

Die drei beteiligten Hochschulen bilden eine Steuerungsgruppe, die aus den drei Prorektoren für Studium und Lehre (Prof. Freitag, Prof. Frommer, Prof. Krieg) besteht. Sie lenkt den Entwicklungsprozess, beschließt die strategische Gesamtausrichtung und nimmt fertige Entwicklungsstände ab. Die Gesamtprojektleitung übernimmt die RWTH Aachen; Prof. Schroeder verantwortet die wissenschaftliche Projektleitung, Dr. Malte Persike die operative Projektleitung. In regelmäßigen (Teil-)Projekttreffen werden Arbeitspakete sowie die damit einhergehende Priorisierung abgestimmt und dokumentiert. Unterstützende Einrichtungen, bestehend aus Fachkräften der Verbundpartner, werden funktional eng angebunden. Die enge Zusammenarbeit in den Arbeitspaketen wird durch interdisziplinäre AP-Leitungen gefördert.

Grundlage der Kooperation ist ein projektorientiertes, horizontal-kollaboratives und vernetztes Arbeiten im Projektverbund. Die Arbeitspakete werden in einem gemeinsamen Backlog angelegt, um eine transparente Übersicht über Arbeitsstand und kommende Aufgaben zu schaffen, redundanzfrei und effektiv zu planen sowie die erreichten Zwischenstände zu dokumentieren. Die Arbeitspakete werden so zugeschnitten, dass sie übersichtliche, abgeschlossene, funktionsfähige Ergebnisse erzielen. Das Backlog wird ergänzt durch eine projektweite Wissensmanagement-Plattform, über die die gesamte Projektkommunikation abläuft.

Um den datenschutzrechtlichen Anforderungen sowohl während der Projektlaufzeit, wie auch mit Blick auf die Überführung in einen Regelbetrieb gerecht werden zu können, werden die jeweiligen Datenschutzbeauftragten und die datenverarbeitenden Stellen an den drei Standorten im Vorfeld der Nutzung von Studierendendaten in den relevanten Unterarbeitspaketen eingebunden (insbesondere UAPs 4.2 und 4.3). Mit Blick auf die forschungsethischen Anforderungen, etwa zur informierten Einwilligung zur Datennutzung in der Entwicklungsphase oder zur Teilnahme an den Befragungen), können die nötigen Voten der Ethikkommissionen erst dann beantragt werden, wenn die konkreten Rahmenbedingungen und Studiendesigns bis hin zu konkreten Fragebögen feststehen. Entsprechende Voten werden dann jeweils bei der Ethikkommission beantragt, die für die konkrete forschende Einrichtung zuständig ist, an der die Projektaktivität durchgeführt wird (z.B. zentrale Kommission der Universität oder Ethikkommission der relevanten Fakultät). Ethische Aspekte wie sie etwa in den Leitgedanken der Datenethikkommission des Bundes formuliert werden sind mit Blick auf eine Überführung in den Regelbetrieb zudem zentraler Bestandteil der im Projekt geplanten konstruktiven Folgenforschung.

5. Beitrag zur didaktisch hochwertigen, kompetenzorientierten Lehre

Das Projekt fokussiert Interventionen auf makrodidaktischer Ebene. StudyBuddy und BuddyAnalytics sowie damit verbundene Planungs- und Feedback-Vorgänge verfolgen nicht unmittelbar das Ziel, Lehrveranstaltungen einzelner Lehrender zu verbessern oder die Kompetenzvermittlung in konkreten didaktischen Lehrsituationen anzupassen. Das Projekt liefert stattdessen in zweierlei Weise einen essentiellen Beitrag zu didaktisch hochwertiger und kompetenzorientierter Studienprozesse:

- **Kompetenzerwerb:** Der StudyBuddy leistet einen Beitrag zur Stärkung insbesondere personeller Kompetenzen wie beispielsweise Selbstreflexion und Selbstorganisation. Dabei findet auch die Verknüpfung mit anderen didaktischen Interventionen und Beratungsangeboten Beachtung. Rückmeldungen und Empfehlungen des StudyBuddy ermöglichen, dass Studierende Kontakt zu Fachstudienberatungen oder ihren Lehrenden aufnehmen, um bestimmte Fragen oder ihre Studienplanung zu reflektieren und zu bearbeiten.
- **Curriculumsentwicklung:** Die BuddyAnalytics spiegeln aggregierte Nutzungsdaten an Studiengangdesignern zurück. Sie können aus ihnen u.a. typische und untypische sowie regelmäßig erfolgreiche oder nicht erfolgreiche Studienverläufe erkennen. Die Studiengangdesigner können auf dieser Grundlage Rückschlüsse für die makrodidaktische Gestaltung von Lernsettings

ziehen, d.h. sie können die Organisation sinnvoll aufeinander abgestimmter Studienoptionen im Sinne einer evidenzbasierten Entwicklung von Studiengängen realisieren.

Insgesamt werden Kompetenzorientierung und didaktische Ausrichtung so zu integralen Gestaltungsmerkmalen. StudyBuddy und BuddyAnalytics werden zu Impulsgebern für die Stärkung von Kompetenzen sowie die didaktische Reflexion und Entwicklung von Curricula. Die Zusammensetzung der Projektpartner und die Vielfalt unter den betrachteten Studiengängen sichert hier eine interdisziplinäre Herangehensweise, die Aspekte der technischen, didaktischen und ethischen Konzeption und Umsetzung sowie der Überprüfung von Akzeptanz und Wirksamkeit umfasst.

6. Einsatz von KI zur Verbesserung der Hochschulbildung

Um Studienmonitoring, Verlaufsplanung und Curriculumsdesign mit hochqualitativer, individualisierter Planungsunterstützung zu leisten, müssen sowohl die formalen in Prüfungsordnung und Studienverlaufsplänen festgehaltenen als auch die informellen Regeln des Studiums, wie günstige Kombinationen und Reihenfolgen, gleichzeitig betrachtet werden. Zur Erfüllung all dieser Anforderungen bietet sich ein hybrider Ansatz aus *Process Mining* (PM) (van der Aalst 2012) und regelbasiertem *Answer Set Programming* (ASP) (Janhunen und Niemelä 2016) an (siehe Abbildung 1). (van der Aalst 2012) und regelbasiertem *Answer Set Programming* (ASP) (Janhunen und Niemelä 2016) an (siehe Abbildung 1).

Mit Process Mining kann das tatsächliche Studienverhalten von Studierenden aus den in CaMS, LMS und PS erfassten Prozess- und Verlaufsdaten entdeckt und analysiert werden. Das tatsächliche Studienverhalten kann sich stark von idealtypischen Studienverlaufsplänen unterscheiden, z. B. da Studierende Module aufgrund geschobener Prüfungen in einer anderen Reihenfolge belegen oder ein Nicht-Bestehen ihren Fortschritt verzögert. Mithilfe von Process Discovery (van der Aalst 2010) (u.a. induktive Mining-Techniken unter Verwendung partiell geordneter Ereignisdaten (Leemans et al. 2016)) können reale Studienverläufe den intendierten gegenübergestellt und Abweichungen erkannt werden. Process Discovery verwendet dabei einen Datensatz (Ereignislog) und kreiert Modelle mit verschiedenen möglichen vorausgesetzten Eigenschaften. Ein solches Ereignislog enthält verschiedene Ereignisse, welche zu verschiedenen Fällen gehören. Jedes Ereignis besitzt mindestens als Attribut eine Aktivität und ein Attribut, welches eine (partielle) Ordnung herstellt zwischen verschiedenen Ereignissen. Im Falle von Studierenden Daten kann hier als Fall ein Studierender betrachtet werden und ein jedes Ereignis kann im Zusammenhang mit einem Kurs stehen. Die Aktivitäten wären beispielsweise die Anmeldung zur Vorlesung, die Anmeldung zur Klausur und im Bestfall das Bestehen der Klausur. Zusätzliche Attribute sind möglich, wie zum Beispiel die Note nach der Klausur, das Modul oder die Anzahl der Antritte vor einer erneuten Klausur. Auch Fälle können Attribute haben, wie zum Beispiel das Land der Hochschulzugangsberechtigung. Wichtig ist anzumerken, dass diese Belegung als Aktivität oder Fall- oder Ereignisattribut nicht in jedem Fall eindeutig sind und auch andere Möglichkeiten bestehen. Als andere Perspektive könnte man als Aktivität Module wählen und anstelle nur eines ordnenden Attributs könnte man eine Startzeit und Endzeit einführen, die das erstmalige Belegen bis zum endgültigen Bestehen oder endgültigen durchfallen abbildet. Mit dieser Perspektive kann man untersuchen, wie Module parallel weiterverlaufen und wie lange Studierende üblicherweise verschiedene Module belegen, bis sie diese erfolgreich abschließen können.

Nach der Erstellung eines Modells muss geprüft werden inwieweit dieses gut die Daten repräsentiert und brauchbar ist. Konformitätsprüfungen (Xixi Lu et al. 2014) können mithilfe Alignment-basierter Techniken automatisch Warnungen und evidenzbasierte Richtlinien generieren (Park und van der Aalst 2020). Durch eine umfassende Datenbasis können die gelernten Prozessmodelle für Vorhersagen auf Studienplan- und Studierendenebene verwendet werden. Process Mining spezifische Anonymisierungs- und Verschlüsselungstechniken (Rafiei et al. 2020; Mannhardt et al. 2019) werden eingesetzt, um den Schutz der Privatsphäre (Rafiei und van der Aalst 2020) zu gewährleisten. Hierbei ist zu beachten, dass klassische Techniken nicht die Komplexität eines Ereignislogs

beachten und damit nicht ohne Anpassungen umsetzbar sind. Als Beispiel lässt sich nennen, dass die Kombination von bestimmten Klausuren dazu führen kann, dass eine Zuordnung zu einem bestimmten Studierenden oder einer kleinen Gruppe von Studierenden möglich ist. Insbesondere wird dies kritisch, wenn dabei auch noch die Reihenfolge bekannt ist, welche notwendig ist, um Process Mining durchzuführen. Die spezifischen Techniken achten dabei darauf, dass wichtige Algorithmen in einem hohen Grad noch möglich sind und gleichzeitig Anonymität mit hoher Wahrscheinlichkeit garantiert werden kann.

Häufig werden diese Techniken im Bereich der Industrie (van der Aalst et al. 2007) angewendet, aber sie werden auch für das Gesundheitswesen (Pegoraro et al. 2021) benutzt, um Prozesse zu verbessern. Viele etablierte Algorithmen sind dabei noch nicht in der Lage die partielle Ordnung, die die Daten der Studierenden aufweisen, mit zu berücksichtigen. Diese Besonderheit der Daten und auch die mögliche Ungenauigkeit (Pegoraro und van der Aalst 2019) auf Grund von menschlichen Fehlern bei der Eingabe muss bei der Entwicklung des konkreten Algorithmus beachtet werden. Diese Fehler im Kontext der Universität können zum Beispiel falsch zugeordnete Semesterbezeichnungen oder Tippfehler bei der Namenseintragung sein, wie z. B. zusätzliche Leerzeichen oder andere Groß- und Kleinschreibung.

Bisher bekannte Projekte, die sich mit Studierendendaten beschäftigen, benutzen bisher immer nur Daten aus einem LMS oder PS, (Mukala et al. 2015; Juhaňák et al. 2019; Bogarín et al. 2018a, 2018b). Eine Gesamtbetrachtung aller verfügbaren Daten ist bisher nicht gemacht worden kann aber an vielen Stelle interessante Einsichten geben und damit zur Verbesserung der Studienplanung, aber auch des Studiums der Studierenden führen.

Answer Set Programming (ASP) (Brewka et al. 2011; Janhunen und Niemelä 2016; Gebser et al. 2018) ist ein etabliertes Paradigma der wissensbasierten KI-Systeme. Es basiert auf Logikprogrammierung und nicht-monotonem Schließen (Gelfond und Lifschitz 1988, 1991) und ist eng verwandt mit Propositional Satisfiability Checking (SAT) (Biere et al. 2009), Satisfiability Modulo Theories (SMT) (Barrett und Tinelli 2018) und Prolog (Bratko 2011). ASP erlaubt eine rein deklarative Programmierung von komplexen Problemen, wie zum Beispiel die Regeln einer Prüfungsordnung. Das Hauptkonstrukt von ASP sind sogenannte Regeln, Ausdrücke der Form $\text{Head} \leftarrow \text{Body}$, wo Body eine logische Konjunktion von Literalen und Head eine atomare Formel ist. Intuitiv sagt eine Regel

ModulC \leftarrow ModulA_bestanden, not ModulB,

dass das Fach Modul C belegt werden kann, falls das Modul A bereits bestanden ist und nicht bereits das Modul B belegt wurde.

Der regelbasierte Ansatz erlaubt dabei, harte Regeln wie die Regeln einer Prüfungsordnung in ein deklaratives Programm umzusetzen, das auch in der Lage ist, Begründungen für Feedback sowie Empfehlungen zu generieren. Das Lösen eines ASP-Programms erfolgt dabei grundsätzlich in zwei Schritten: Zunächst wird das Eingabeprogramm durch einen *grounder* wie zum Beispiel *gringo* (Gebser et al. 2019) oder *Alpha* (Eiter et al. 2018) übersetzt, um eine einfachere und variablenfreie Repräsentation des Problems zu erhalten, wobei beispielsweise externe Funktionen bereits ausgewertet und das Programm optimiert wird. Im zweiten Schritt wird dann eine zulässige Belegung mit einem *solver* wie *clasp* (Gebser et al. 2007) oder *dlv* (Adrian et al. 2018) bestimmt. Zusätzlich zu harten Regeln erlaubt ASP auch die Formulierung von weichen Kriterien, entweder in der Form einer zu optimierenden Variablen, wie zum Beispiel die erwartete Gesamtdauer des Studiums, oder einer priorisierten Rangliste von Präferenzen (Gebser et al. 2011), wie etwa der Wunsch, das Modul C zu wählen, mit niedrigerer Priorität jedoch nicht mehr als 34 CPs in einem Semester zu belegen. In beiden Fällen erfolgt aus den weichen Kriterien eine Rangfolge der Lösungen, von denen die beste ausgewählt wird.

Allgemein wird ASP in vielen verschiedenen Bereichen eingesetzt (Erdem et al. 2016), etwa zur Erstellung von individualisierten Produkten (Tiihonen et al. 2003), auf Haushaltsrobotern (Erdem et al. 2012; Erdem und Patoglu 2018), und auf dem Space Shuttle (Nogueira et al. 2001). Besonders

zu erwähnen sind auch die bei KBS durchgeführten Steuerung einer Flotte von Robotern in der Produktionslogistik mithilfe von ASP (Schäpers et al. 2018), sowie *teaspoon* (Banbara et al. 2019), eine ASP Anwendung, welche die Raumebelegung und den Stundenplan von Lehrenden und Studierenden berechnet und optimiert. Relevant für dieses Projekt ist auch *zlog* (Adrian et al. 2018), ein Tool für intelligentes Routing in einem Call Center, das Anrufer:innen basierend auf einer Klassifizierung zu geeigneten Mitarbeiter:innen durchstellt. Solch eine Klassifizierung könnte in unserem Fall auch durch Process Mining erfolgen, indem bestimmte Gruppen von Studierenden unterschieden werden, woraus dann entsprechende gruppenspezifische Regeln für den StudyBuddy abgeleitet werden.

Regeln aus einem ASP können genutzt werden, um mithilfe eines ASP Solvers eine gegebene Planung im StudyBuddy zu bewerten und Empfehlungen zu geben. Da im Gegensatz zum maschinellen Lernen die Semantik der Regeln, mit denen das System arbeitet, transparent ist, kann das KI-System im StudyBuddy detaillierte Begründungen für Feedback und Empfehlungen generieren.

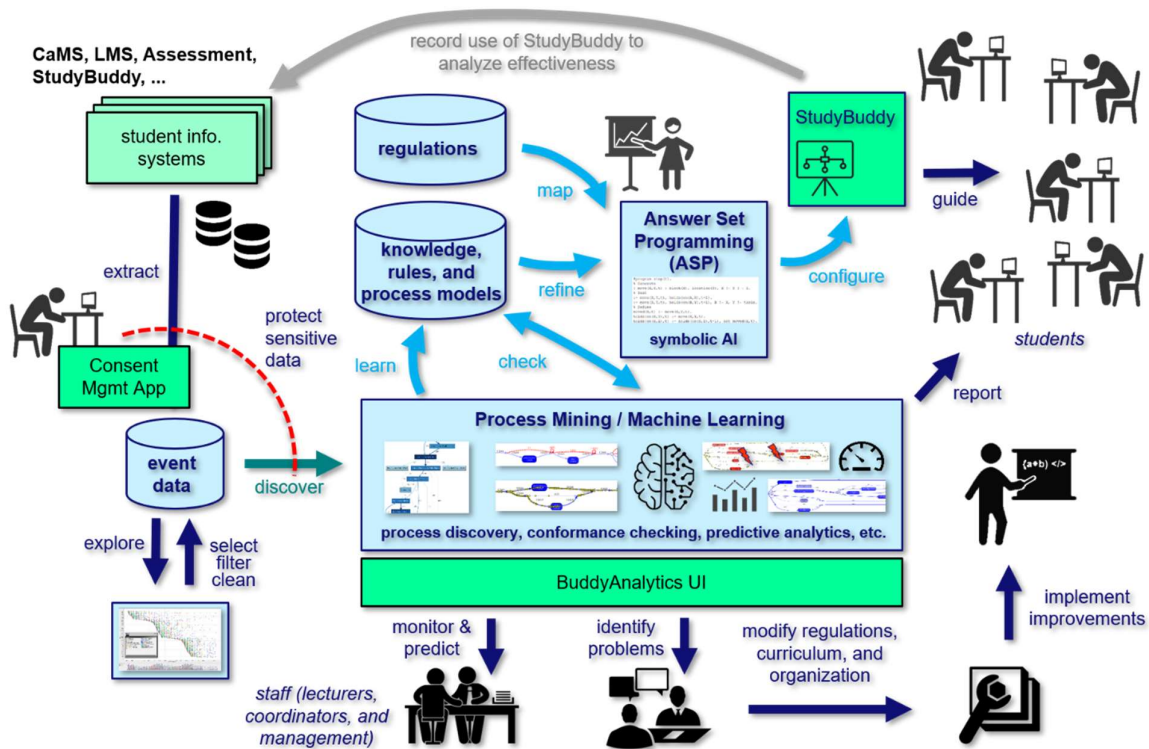


Abbildung 1. Überblick über die Erfassung, Aufbereitung und Verarbeitung von Studierendendaten in der KI-basierten Architektur.

Die KI-basierte Architektur, wie in Abbildung 1 zu sehen, kombiniert Methoden aus dem Process Mining mit Methoden des Answer Set Programmings. Dabei verschiedene Kombinationen dieser Methoden möglich, die in diesem Projekt untersucht und umgesetzt werden sollen.

Zunächst werden aus unterschiedlichen Hochschulsystem Daten extrahiert und gespeichert. Hierbei gilt, dass Studierende zu jedem Zeitpunkt in einer Consent Management App die Kontrolle darüber haben, ob und wie ihre personalisierten oder pseudonymisierten Daten verwendet werden. Durch datenschutzrechtliche und ethische Absicherung bei der Datengewinnung sowie explizite Transparenz im Einsatz der KI-Technologien sollen Vertrauen und Akzeptanz gesichert werden.

Diese unbearbeiteten Daten können durch verschiedene Techniken, z. B. Dotted Charts (Song und van der Aalst 2007), sondiert werden, um dann Teildaten herauszufiltern, auszuwählen oder zu bereinigen. Dieser vorbereitete Datensatz ist die Basis für Process Discovery Algorithmen, die Mo-

delle erstellen, welche die in der Wirklichkeit vorkommende Studienverläufe abbilden. Auf Basis dieser Modelle können ASP-Regeln gelernt werden, die das aus der Prüfungsordnung entstandene Programm um Praxiserfahrungen ergänzt. Dies könnte beispielsweise dazu genutzt werden, Studierende zu erkennen, bei denen ein Studienabbruch droht, um diese dann entsprechend informieren zu können. Außerdem können die Informationen aus dem Process Discovery genutzt werden, um auf einer Peer Group basierend Vorschläge zu machen, welche Module von Interesse sein könnten.

Alternativ kann ASP in Process Mining integriert werden, zum Beispiel um bestimmte Abhängigkeiten zwischen Modulen, die sich aus der Prüfungsordnung ergeben, einfach im Process-Mining-Modell darstellen und insbesondere für Konformitätsüberprüfungen verwenden zu können. Dies könnte zum einen der höheren Ausdruckstärke, zum anderen der einfacheren Lesbarkeit des Modells dienen. So können Studierende informiert werden, wenn ihr geplanter oder bereits getätigter Studienverlauf vom vorhergesehenen Studienverlauf abweicht. Außerdem würde ein solches inklusives Model Studiengangsdesignern erlauben, Anpassungen im Studienverlaufsplan vorzunehmen. Solche Anpassungen können zum Beispiel eine veränderte Reihenfolge von Fächern oder das Einführen eines neuen Pflichtmoduls sein. Auch kann dies die Basis sein, um Studierende zu führen und ihnen direktes Feedback zu geben. Hier ist zu denken an die Empfehlung von Modulen auf Basis des bisherigen Studienverlaufes. So kann eine mögliche Empfehlung aufgrund des Process Minings bestimmt und mit ASP auf Konformität überprüft werden. Außerdem können Studierende informiert werden, falls eine geplante Kombination von Wahlmodulen nicht der Prüfungsordnung entsprechend regelgetreu ist. Die tatsächliche Umsetzung des inklusiven Modells und wie dieses genutzt wird, um Vorhersagen zu machen ist Teil des Projektvorhabens. Die Ergebnisse der Benutzung des StudyBuddys und BuddyAnalytics können messbar gemacht werden durch das Aufzeichnen der Benutzung des StudyBuddys und der Markierung von Veränderungen im Studienverlaufsplan.

7. Einbettung geplanter Maßnahmen

Die **RWTH** zielt mit dem prämierten Konzept „Studierende im Fokus“ darauf ab, den individuellen Studienerfolg bei gleichbleibendem Leistungsanspruch und angemessener Studiendauer zu erhöhen und Abbruchquoten zu senken. Voraussetzung ist die Gewährleistung eines reibungslosen, optimierten Studien- und Prüfungsverlaufs, vollständige Information, Beratung und Betreuung der Studierende, ein umfassendes Angebot sowie intensive und individuelle Förderung vor und während des Studiums. Das Projektvorhaben trägt wesentlich dazu bei, diese Voraussetzung sicherzustellen.

Die **RUB** stellt in ihrem Zukunftskonzept Lehre das forschende Lernen in den Fokus des Studiums. Studierende sollen früh aktiv am wissenschaftlichen Denken und Arbeiten beteiligt werden. Eine wesentliche Aufgabe der Curriculumentwicklung besteht darin, Studierenden mehr Selbstständigkeit zu ermöglichen. Das Vorhaben fügt sich in das Zukunftskonzept Lehre nahtlos ein, da es einerseits Studierenden hilft, bessere Studienentscheidungen zu treffen, andererseits Lehrende bei der Studiengangkonzeption unterstützt.

Für die **BUW** besteht die wichtigste Aufgabe akademischer Lehre darin, Menschen zum eigenverantwortlichen Umgang mit Erkenntnis und ihrer Anwendung zu bilden. Um die Effizienz und Effektivität des Studiums zu steigern, wird eine kontinuierliche Verbesserung der Dienstleistungs- und Beratungsorientierung angestrebt. Schon jetzt setzt die Bergischen Universität ein KI-basiertes Frühwarnsystem mit Daten aus dem CaMS, PS und zukünftig LMS ein, das drohende Studienabbrüche früh identifiziert und Studierende informiert. Das Projektvorhaben entwickelt dieses nicht nur weiter, es ermöglicht Studierenden auch, sich kritisch mit KI-Technologien auseinanderzusetzen und ihre gesellschaftliche Relevanz zu bewerten.

Das Vorhaben fügt sich zudem in die Angebote der **Digitalen Hochschule NRW (DH-NRW)** ein, die ein Teil der Digitalisierungsstrategie des Landes ist. Das beantragte Verbundprojekt möchte die

DH-NRW nutzen, um hochschulübergreifende Services für die digitale Weiterentwicklung des Lehrens, Lernen und Studierens zu entwickeln, zu betreiben und für alle Hochschulen auch über NRW hinaus zur Nutzung zur Verfügung stellen. Weitere Maßnahmen und Einbettung und Vernetzung sind in Teil B, Abschnitt 3 beschrieben.

8. Planung zu Verwertung und Transfer der Ergebnisse

Das Projekt wird einen Studienverlaufsplaner in Form eines KI-basierten StudyBuddy und einen Curriculumsmonitor in Form einer KI-basierten BuddyAnalytics bereitstellen, deren Komponenten als Open Source Software publiziert und gepflegt werden. Parallel dazu werden die Referenzarchitektur sowie die Analysewerkzeuge für Studienfortschrittsdaten als Services verfügbar gemacht. Der Entwicklungsstand am Projektende soll einen Regelbetrieb der Software an den Verbundhochschulen ermöglichen und an weiteren Standorten einsetzbar sein. Die Verstetigung und nachhaltige Weiterentwicklung der Projektkomponenten wird durch die institutionelle Verankerung gesichert, z.B. an der RWTH durch den Betrieb am CLS. Zur Sicherung der Nachhaltigkeit gehört auch die bereits im Rahmen der Projektplanung begonnene Kooperation von Hochschulschulleitungen, Hochschulverwaltungen und Wissenschaft zwischen den Projektpartnern, die durch die Projektgovernance intensiviert und formalisiert wird. Die Vernetzung des Projekts, u.a. über die DH.NRW, stellt eine Verbreitung der Projektergebnisse über die beteiligten Hochschulen hinaus sicher.

Die Projektpartner verfolgen einen zweigliedrigen Veröffentlichungsplan. Zum einen werden im Projektverlauf kontinuierlich wissenschaftliche Publikationen in Fachzeitschriften (bspw. Computers in Education, Research in Higher Education, Computers & Education), wie auch Vorträge auf wissenschaftlichen Tagungen (bspw. EDM, LAK) stattfinden. Diese werden maßgeblich auch aus den Promotionsarbeiten bei den Projektpartnern generiert. Zum anderen sollen Veröffentlichungen für eine breite Öffentlichkeit erfolgen (bspw. Blogbeiträge im Hochschulforum Digitalisierung oder auf e-teaching.org). Die Publikationen fokussieren sowohl auf technologische Weiterentwicklungen als auch auf die Begleitforschung zu Diversität und Akzeptanzkriterien sowie auf die wissenschaftliche formative und summative Evaluation. Die Konzepte der evidenz- und constraintbasierten Studienplanung und -reflexion (Studierende) sowie der evidenzbasierten Curriculumsplanung (Studiengangsgestaltung) werden in Form von Open Access Case Studies aufbereitet, die als Handreichungen aus der umfassenden systematischen Dokumentation der Arbeitspakete extrahiert werden.

Strukturelle Effekte des Projekts für die Hochschulen sollen über die Projektlaufzeit hinausreichen. Dies betrifft den Paradigmenwechsel zum optimistischen Umgang mit personalisierten bzw. pseudonymisierten Daten durch Privacy-preserving Data Publishing, Ethics-by-Design und transparentem Consent Management. Das Projekt sondiert zudem die Einrichtung einer Clearing Stelle in Abstimmung mit der DH.NRW, um eine datenschutzkonforme Verarbeitung lehrbezogener Daten zu fördern. Zu den Projektmaßnahmen gehört auch die Ausrichtung einer Konferenz im Hybridformat während der zweiten Projekthälfte, um die Vernetzung von Wissenschaft, Serviceeinrichtungen und Hochschulverwaltungen zu den Themen des evidenzbasierten Curriculumsdesigns, des KI-basierten Studienmonitorings und der datengestützten Studienverlaufsplanung zu fördern.

9. Nutzung von Synergien und des strukturellen Mehrwerts der Kooperation

Eine der besonderen Stärken des Projektverbunds liegt im kontinuierlichen, fest in der Projektstruktur verankerten Transfer von Ideen und Technologien zwischen den Verbundpartnern mit ihren komplementären Besonderheiten der Studierendenschaften, Studienangebote, Lehr-Lern-Arrangements und Campusmanagement-Infrastrukturen. Die Verbundpartner bringen umfangreiche Studierendenzahlen mit diversen Studierendekulturen in das Projekt ein. Dabei bringen die RWTH als Technische Universität, die RUB als Volluniversität und die BUW mit ihrem stark interprofessionell angelegten Studienangebot ihre sowohl komplementären wie auch konvergenten Fächerspektren ein, um die Vielfalt der Studienverläufe umfassend adressieren zu können. Die unterschiedlichen Hochschulsysteme z.B. im Bereich des Campusmanagements und des Lernmanagements werden

von den Verbundpartnern als Chance verstanden. Um das Potential der technischen und organisatorischen Diversität zu nutzen, werden als Kernmerkmal der Entwicklung und Implementierung in den verschiedenen Teilprojekten klare Schnittstellendefinitionen in explizit definierten Arbeitspaketen eingesetzt. Dies sichert gemeinsam mit der eng verzahnten Entwicklung die Interoperabilität und vermeidet typische Fehlerquellen bei Verbundprojekten. In Verbindung mit der Entwicklung des Referenzmodells ist somit die Transferierbarkeit der Projektergebnisse über den Projektverbund hinaus ein immanentes Merkmal der Zusammenarbeit.

Strukturelle Effekte des Projekts für die Hochschulen sollen über die Projektlaufzeit hinausreichen. Dies betrifft den Paradigmenwechsel zum optimistischen Umgang mit personalisierten bzw. pseudonymisierten Daten durch Privacy-preserving Data Publishing, Ethics-by-Design und transparentes Consent Management. Dazu gehört auch der interdisziplinäre Ansatz im Projekt der integriert technischen, didaktischen und ethischen Konzeption und Umsetzung sowie der wissenschaftlichen Überprüfung von Akzeptanz und Wirksamkeit. Auch die Kooperation von Hochschulschulleitungen, Hochschulverwaltungen und Wissenschaft zwischen den Verbundpartnern zeichnen das Projekt aus. Auf Landesebene wird durch die Sondierung der Einrichtung einer Clearing Stelle das Fundament für die Ermöglichung eines sicheren und datenschutzkonformen Datenaustauschs zwischen Hochschulen gelegt.

Die Vernetzung des Projekts u.a. über die DH-NRW, ORCA.nrw und den KI-Campus stellt eine Verbreitung der Projektergebnisse sicher. Zu den Projektmaßnahmen soll zudem die Ausrichtung einer Tagung gehören, in der Wissenschaft und Hochschulverwaltungen zum Thema des KI-basierten Studienmonitorings zusammenkommen.