



ENTWICKLUNG DURCH BILDUNG

Arbeits- und Forschungsberichte aus dem Projekt
EB – Bildung als Exponent individueller und regionaler Entwicklung
Nr. 7

Der Regionalmonitor Hochschulbildung

Datenkonstruktion und Gestaltung

David Van de Water und Bastian Steinmüller

unter Mitwirkung von Dominik Böhmer

2017

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



Impressum:

E^B – Bildung als Exponent individueller und regionaler Entwicklung
– Evidenzbasierte Bedarfserschließung und vernetzte Kompetenzentwicklung
Förderkennzeichen: 16OH21008

Herausgeber:

Hochschule Kaiserslautern
Prof. Dr.-Ing. Hans-Joachim Schmidt
Schoenstraße 11
67657 Kaiserslautern

Technische Universität Kaiserslautern
Jun.-Prof. Dr. Matthias Rohs
Erwin-Schrödinger-Straße
67663 Kaiserslautern

Hochschule Ludwigshafen
Dr. Doris Arnold
Ernst-Boehe-Str. 4
67059 Ludwigshafen am Rhein

2017

ISSN 2364-8996

Lizenz

Arbeits- und Forschungsberichte aus dem Projekt E^B sind unter einer Creative-Commons-Lizenz veröffentlicht:
Namensnennung - Nicht kommerziell - Keine Bearbeitungen 4.0 International Lizenz. <http://creativecommons.org/licenses/by-nc-nd/4.0/>





Van de Water, D; Steinmüller, B.: Hochschulbildung

Inhaltsverzeichnis

Abbildungsverzeichnis.....	3
1 Einleitung	4
2 Nutzergruppen, Anforderungen und Herausforderungen	4
3 Datenkonstruktion	6
3.1 Beschreibung einer Bildungsregion.....	7
3.2 Datenquellen und relevante Merkmale.....	9
4 Informationstechnische Grundlagen	12
4.1 Allgemeine Systemstruktur	12
4.2 Software-Architektur	14
4.3 Grundlegende Systembausteine.....	15
5 Chancen und Grenzen	17
Literaturverzeichnis	18



Van de Water, D; Steinmüller, B.: Hochschulbildung

Abbildungsverzeichnis

Abbildung 1 Karte der Bildungsregion im Projekt E ^B	8
Abbildung 2 Systemstruktur und verwendete Programmiersprachen	13
Abbildung 3 Grundlegende Software-Architektur	14



1 Einleitung

Der Regionalmonitor Hochschulbildung (folgend Regionalmonitor genannt) trägt als interaktive webbasierte Datenbank dazu bei, den Übergang von einer angebotsorientierten zu einer bedarfsorientierten Entwicklung von Studienangeboten im Bereich der wissenschaftlichen Weiterbildung praktisch umzusetzen (vgl. Rohs, Vogel & Marks, 2015). Das Ziel liegt in der Bereitstellung empirischer Daten zur demographischen und wirtschaftlichen Entwicklung in der Region E^B (Marks, 2015), um die Gestaltung von Studienangeboten stärker als bisher an derzeitigen und zukünftigen Bedarfen zu orientieren. Die Daten werden in einem webbasierten Regionalmonitor integriert und für Fragen zur Gestaltung von Studienangeboten aufbereitet. Dadurch kann ein wichtiger Beitrag zu einer fundierten und ressourceneffizienten Entwicklung von Studienangeboten sowie zur Verbindung von Hochschul- und Regionalentwicklung geleistet werden.

Der Arbeitsbericht befasst sich eingangs mit den potentiellen Nutzergruppen des Regionalmonitors und geht auf Anforderungen ein, die mit dem Ziel einer nachhaltigen Etablierung des Systems über die Projektlaufzeit hinaus verbunden sind. Sodann wird die dahinter liegende Datenkonstruktion erläutert, was eine Beschreibung der für die Zielsetzung relevanten Merkmale und der verwendeten Datenquellen umfasst. Weiterführend werden aus informationstechnischer Sicht der Grundaufbau sowie wesentliche Funktionsbausteine der Software dargestellt. Abschließend erfolgt eine differenzierte Auseinandersetzung mit den Herausforderungen, die mit der technischen Konstruktion eines solchen Systems einerseits und mit der Nutzbarkeit empirischer Daten für die Angebotsentwicklung andererseits verbunden sind.

2 Nutzergruppen, Anforderungen und Herausforderungen

Potentielle Nutzergruppen des Regionalmonitors im Hochschulbereich sind zunächst in der Studienangebotsentwicklung tätige Personen, die prospektiv die Initiierung neuer Studienangebote vorantreiben und retrospektiv die Gestaltung eines neuen Studienangebots beispielsweise in Bezug auf Studierbarkeit und Absolventenverbleib legitimieren müssen (Technische Universität Kaiserslautern, 2016). Mit der Bereitstellung von fächerorientierten Daten zu Fachkräften und Zielgruppen in der Region soll die zu erwartende Relevanz von neuen Studienangeboten genauer als bisher eruiert und die Entscheidung für oder gegen deren Implementierung fundierter getroffen werden können. Ferner kann ein solches System die Leitung einer

Van de Water, D; Steinmüller, B.: Hochschulbildung

Hochschule oder eines Anbieters wissenschaftlicher Weiterbildung dabei unterstützen, die Positionierung der Hochschule bzw. des Weiterbildungsanbieters in der jeweiligen Region zu stärken. So kann etwa durch die Berücksichtigung prognostizierter Fachkräftebedarfe eine stärkere Bindung an regionale Unternehmen unterstützt werden. Weiterhin können Professor_innen, die als Experten ihres Fachs Ideengeber für neue Studienangebote darstellen, sowie Hochschulstrateg_innen in Fachgebieten, Fachbereichen etc. den Regionalmonitor nutzen. Derartige Synergien aus der Verbindung der intrinsischen Motivation, das eigenes Fachgebiet durch ein neues Studienangebot zu stärken, mit den erkennbaren Bedarfen in der Region, sind nicht zuletzt mit Blick auf die beruflichen Chancen der zukünftigen Absolvent_Innen vielversprechend.

Die Entwicklung des Regionalmonitors ist mit Anforderungen verknüpft, die sowohl die Datenkonstruktion als auch die technische Integration der Daten betreffen. Eine Spannung existiert zwischen der Datenqualität und dem Datenschutz. Das Bestreben, die soziodemographische Entwicklung einer eng abgegrenzten Region mit empirischen Daten zu erfassen, wird bisweilen durch zu geringe Fallzahlen limitiert. Gerade im vorliegenden Fall zeichnet sich die betreffende Region durch eine Vielzahl kleinerer Gemeinden aus. In vorliegenden Umfragen wird der Bereich Westpfalz nicht mit ausreichenden Fallzahlen abgedeckt, um ‚repräsentative‘ Aussagen zu treffen. Je nach Auswahl beziehungsweise Kreuzung von Merkmalen, selbst nach Aggregation einzelner Datensätze, fallen die Stichprobengrößen zu klein aus. Sehr geringe Fallzahlen können dem datenschutzrechtlichen Grundsatz widersprechen, Rückschlüsse von veröffentlichten Daten auf einzelne Personen auszuschließen.¹ Eine damit einhergehende Einschränkung ist, dass Einzeldatensätze (Rohdaten) nicht in das System importiert werden können. Denn der direkte Zugang zu formal anonymisierten Originaldaten (z.B. Mikrozensus) ist seitens der statistischen Ämter gar nicht zulässig (Statistische Ämter des Bundes und der Länder, 2016b).

Eine Lösungsmöglichkeit ist die so genannte „Kontrollierte Datenfernverarbeitung“, bei der Datennutzer auf Basis eines Dummydatensatzes eine Auswertungssyntax deklarieren, diese an die Statistischen Ämter senden und nach einer Anonymitätsprüfung die *aggregierten* Daten

¹ So wird im Umgang mit amtlichen Statistiken gefordert: „Sind die Einzelangaben dem Befragten oder Betroffenen nicht zuzuordnen, aus den Daten also keine Rückschlüsse auf die Auskunftgebende Stelle oder Person mehr möglich (absolute Anonymität), können diese Einzelangaben auch außerhalb der amtlichen Statistik genutzt werden (§16 Abs. 4 BStatG)“ (Statistische Ämter des Bundes und der Länder, 2016a).

Van de Water, D; Steinmüller, B.: Hochschulbildung

erhalten. Folglich ist die Art der Daten, welche durch das System zur Verfügung gestellt werden, bereits in der Entwicklungsphase strikt festgelegt, sodass die Nutzerinnen und Nutzer weniger Raum für die Auswertung nach eigenen, spezifischen Fragestellungen erhalten.

Regionen sind einem ständigen Wandel unterzogen, sodass die demografische und wirtschaftliche Struktur über die Zeit variiert. Daraus resultiert die Anforderung, die integrierten Daten stetig zu aktualisieren. Einerseits setzt dies eine skalierbare Modellierung der Software- bzw. Datenbankstruktur voraus, damit der Import neuer Datendateien gewährleistet ist. Andererseits – und inhaltlich von großer Relevanz – ist die Auswahl der Datenquellen davon betroffen. So ist mit dem Rückgriff auf amtlich erhobene Daten, wie etwa dem Mikrozensus, zwar eine nachhaltige Verfügbarkeit gesichert, da diese Erhebungen teils gesetzlich verankert sind und dauerhaft durchgeführt werden. Jedoch enthalten diese nicht immer alle benötigten Merkmale. Diese Überlegungen tangieren auch die weitere Anforderung der nachhaltigen Etablierung des Systems über die Projektlaufzeit hinaus, die ebenfalls nur mit dem Rückgriff auf Daten, deren Erhebung bereits institutionell verankert ist, eingehalten werden kann. Obwohl aufgrund dieser Festlegung der Kern des Regionalmonitors auf amtlich erhobenen Daten basiert, ist die Integration spezieller Erhebungen, die etwa inhaltlich nur einen speziellen Bereich abdecken oder nur einmalig erhoben werden, nicht ausgeschlossen. Denkbar ist hier das Einbinden von PDF-Dateien, die bereits aufbereitete statistische Informationen etwa in Form von Grafiken oder auch Texten beinhalten und damit etwaige Unschärfen in den amtlichen Daten auflösen können.

Eine weitere Anforderung ist schließlich die Übertragbarkeit des Systems auf andere Regionen. Auch dies legitimiert den Einsatz amtlicher Daten, da diese in der Regel deutschlandweit erhoben werden, somit weitere Regionen umfassen, und zusätzlich interregionale Vergleiche vereinfachen.

3 Datenkonstruktion

Die Datenkonstruktion für den Regionalmonitor umfasst die Auswahl, Akquise und Aufbereitung von empirischen Daten über die Region E^B. Dabei kommen in besonderem Maße forschungstheoretische Prämissen zum Tragen, wie sie nachfolgend bei der Eingrenzung des Regionsbegriffs dargestellt werden. So erfolgt die Auswahl relevanter Merkmale insbesondere vor dem Hintergrund von Einflussfaktoren, die sich gerade aufgrund ihrer regionalen Bedeutung auf die Entwicklung des Weiterbildungsstandortes auswirken.

3.1 Beschreibung einer Bildungsregion

Eine Bildungsregion² wird auf verschiedenen Ebenen beschrieben. Ein erster Überblick wird über die Darstellung allgemeiner demographischer Daten wie Altersstruktur, Wanderungssaldi oder Geschlechterverhältnis geschaffen. Daten über den Arbeitsmarkt liefern Informationen zum Fachkräfteangebot sowie zur Fachkräftenachfrage mit dem Ziel der Abschätzung von zukünftigen Fachkräftebedarfen. Daten über die anvisierten Zielgruppen geben Auskunft über potentielle Teilnehmerinnen und Teilnehmer für geplante Studienangebote. Kriterien wie Qualifikation, Staatsangehörigkeit oder Arbeitszeit erleichtern es, sowohl die potenzielle Nachfrage nach einem Studiengang einzuschätzen, als auch eine passgenaue didaktische Gestaltung zu ermöglichen. Ferner werden Daten über den aktuellen Bildungsmarkt bereitgehalten. Die Kenntnis über eine mögliche Sättigung des Bildungsmarkts in einem bestimmten Themenfeld kann für die Einführung eines neuen Studienangebotes entscheidend sein.

Die Daten stammen zum Teil aus sehr unterschiedlichen Quellen und weisen hinsichtlich ihrer geographischen Konsistenz wie auch ihrer inhaltlichen Breite eine erhebliche Heterogenität auf. Dies ist unvermeidlich, möchte man eine Region in einer validen Bandbreite mit bildungsrelevanten Daten beschreiben. Um dennoch eine klare Handhabung zu gewährleisten, bedarf es zweierlei: der (1) sinnvollen Integration der Daten in (2) eine thematisch breitgefächerte und intuitiv zugängliche Anwendung, welche die Daten regionalbasiert aufbereitet (es wird in Kapitel 4 näher beschrieben). Dreh- und Angelpunkt der Datenakquise, -integration und -aufbereitung muss daher die Region selbst sein.

Der dabei zugrundeliegende Regionsbegriff beruht auf Weiterbildung als zentralem analytischen Bezugswert. Er fokussiert im Speziellen auf den Bildungsinteressierten selbst und beschreibt das Einzugsgebiet um den Weiterbildungsstandort Kaiserslautern, bildet also strukturell wie räumlich eine Bildungsregion ab (Marks, 2015). Es lässt sich ableiten, dass *Region* hier unmittelbar *Bildung* impliziert. Dabei bestimmt die Ausrichtung der *Bildungsregion* auf ihre Akteure zugleich das Leitmotiv ihrer inhaltlichen Basis: die Orientierung an Zielgruppen von Bildung (ebd.). Auf Basis der Studierendenstatistik der TU Kaiserslautern zu Wohnorten der

² Der Begriff „Bildungsregion“ wird in der Literatur, anders als hier, vor allem mit Fokus auf Schulbildung verwendet. Siehe dazu beispielsweise Stern et al. (2008) oder Emmerich (2010).

Van de Water, D; Steinmüller, B.: Hochschulbildung

Studierenden, Daten zur Weiterbildungsteilnahme nicht traditioneller Zielgruppen in der Westpfalz aus dem Adult Education Survey (Bundesministerium für Bildung und Forschung, 2015) sowie Angaben, die vom Statistischen Bundesamt zur beruflichen Mobilität in der Region erhoben werden, wurden Prämissen bestimmt, welche die hier verwendete Bildungsregion determinieren.

Sie rekurren in ihrem Kern auf Weiterbildung als zentralem regionalem Bezugswert, verbinden die Regionsdefinition also mit dem regionalen Zugehörigkeitsgefühl der Bildungsinteressierten. Geographisch begrenzt wird diese lose Regionsvorstellung durch ein angenommenes Maximum der zumutbaren zeitlichen Belastung zum Erreichen der Hochschulen als Weiterbildungsstätten durch die Bildungsinteressierten: Sie müssen innerhalb von 60 Minuten (einfacher Weg) erreicht werden können (ÖPNV). „Die Reichweite, die sich durch die Fahrtzeit ergibt, wird zuletzt durch die administrativen Grenzen der Landkreise bzw. kreisfreien Städte ergänzt. Daher wird nicht die Erreichbarkeit binnen 60 Minuten im Einzelnen geprüft, sondern erfolgt in einer durchschnittlichen Abschätzung für den gesamten Kreis“ (Marks, 2015, S. 13). (vgl. Abbildung 1).

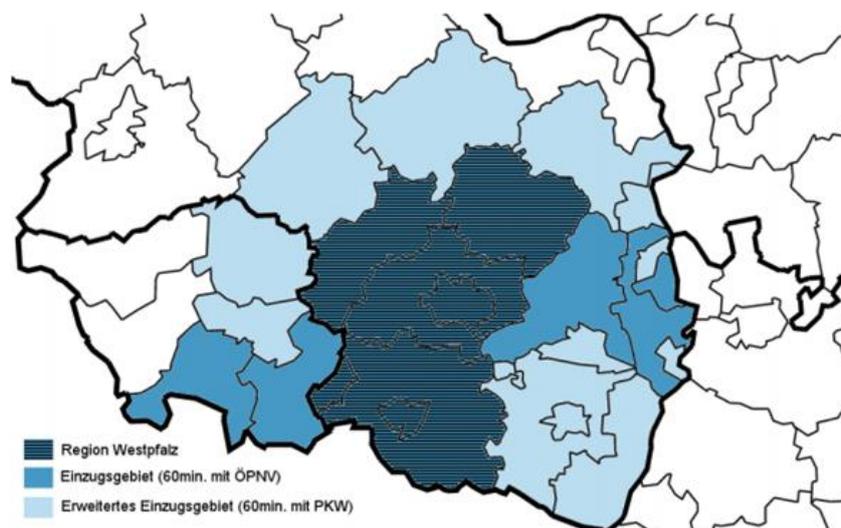


Abbildung 1 Karte der Bildungsregion im Projekt E^B (Marks, 2015, S. 14)

Für weiterführende Informationen zur Bildungsregion, etwa ihre räumliche Ausdehnung, beinhaltete Landkreise oder demographische Kennzahlen, siehe Marks (2015).

Van de Water, D; Steinmüller, B.: Hochschulbildung

Diesem Verständnis von Region grundlegend verhaftet, fällt die Auswahl adäquater Datenquellen voraussetzungsvoll aus und ihre Akquise und Integration ist mit einigen Hürden verbunden. Das nachfolgende Kapitel führt das je Datenquelle aus.

3.2 Datenquellen und relevante Merkmale

Regionalkompass

Es ist die Grundlage des Projektes, die oben beschriebene Region möglichst facettenreich mit empirischen Daten zu beschreiben. Die verwendeten Daten müssen dabei immer einerseits dem Anspruch genügen, die strukturellen Besonderheiten der Region sichtbar zu machen und andererseits vor dem Hintergrund von Weiterbildung als thematischem Fokus des Projekts ausgewählt werden.

Die E^B-Region ist in weiten Teilen mit der Westpfalz kongruent, reicht jedoch stellenweise über deren Grenzen hinaus (vgl. Abbildung 1). Ihr wirtschaftliches Zentrum und Sitz der wichtigsten Bildungseinrichtungen, Kaiserslautern, ist Mittelpunkt eines Gebiets, das sich durch eine ländliche Prägung auszeichnet und als strukturschwach gilt (z. B. Prognos AG, 2016). Um diesen Charakteristika Rechnung zu tragen, wurden Daten nötig, die sozialstrukturelle Bevölkerungsmerkmale möglichst breit abbilden und in genügend großer Fallzahl erhoben wurden, um auch für bevölkerungsarme Gegenden noch Aussagekraft zu besitzen. Diese Daten tragen dazu bei, ein grundlegendes Verständnis der Region zu gewinnen: differenziert nach sozialen und ökonomischen Kennzahlen und mit regionaler Tiefe, d.h. bis auf Ebene der Landkreise hinunter. In diesem Sinne bereiten sie den Boden dafür, die ergänzenden Daten der Bereiche Arbeitsmarkt, Bildungsmarkt und Zielgruppen synergetisch und zielführend zu interpretieren.

Die regionale Eingrenzung, insbesondere die selbst gewählten Grenzen der E^B-Region, erschweren die Datenakquise erheblich. Gegenüber üblichen regionalstatistischen Einheiten, etwa Bundesländern oder sonstigen Verwaltungsbezirken, die in nahezu allen überregionalen Umfragen Berücksichtigung finden, wurde die verwendete Region bisher nicht statistisch abgebildet. Insofern mussten Daten aus kleineren statistischen Einheiten aggregiert werden, um ein regionales Abbild zu schaffen, was wiederum zum Teil datenschutzrechtlichen Einschränkungen unterliegt und darüber hinaus nicht für alle gewünschten Items realisierbar ist.

Van de Water, D; Steinmüller, B.: Hochschulbildung

Arbeitsmarkt

Die Daten für die Darstellung der Beschaffenheit des regionalen Arbeitsmarktes werden vom unabhängigen Wirtschaftsforschungsinstitut (WifOR, 2016) bereitgestellt, das als Kooperationspartner des E^B-Teilprojekts der Universität Kaiserslautern fungiert. Sie umfassen Informationen zum Fachkräfteangebot, zur Fachkräftenachfrage, Engpässen und Trends und sind zu Berufsgruppen und Jahren – derzeit für den Zeitraum von 2007 bis 2030 – aggregiert. Indem einerseits die Fachkräftebedarfe der vergangenen Jahre bis heute und andererseits die prognostizierten Bedarfe für die kommenden Jahre nachvollziehbar sind, erhalten die Nutzerinnen und Nutzer einen empirisch fundierten Einblick in die Arbeitsmarktsituation. In Abgrenzung zum Fachkräftemonitor Rheinland-Pfalz (IHK-Arbeitsgemeinschaft Rheinland-Pfalz, 2016) ist die Präsentation der Daten in der Benutzungsoberfläche aus der Sicht von Hochschulangehörigen gestaltet. So wurden die bereitgestellten Berufsgruppen den an Hochschulen weit verbreiteten Fachbereichen zugeordnet. Eine Herausforderung stellt die Zusammensetzung der Berufsgruppen dar, die von WifOR gruppiert werden. Sie subsumieren mitunter ein breites Spektrum an Berufen, was eine mehrfache Zuordnung mancher Berufsgruppen zu verschiedenen Fachbereichen erfordert. Beispielsweise ist die Berufsgruppe 'Werbung, kaufm. und redaktionelle Medienberufe, geistes- und wirtschaftswissenschaftliche Berufe' den Fachbereichen 'Kunst, Musik, Design', 'Sprach- und Kulturwissenschaften' sowie 'Wirtschaftswissenschaften, Rechtswissenschaften' zugewiesen.

Zielgruppen

Um sich potentiellen neuen Zielgruppen wissenschaftlicher Weiterbildung auf Basis soziodemographischer Daten zu nähern, bietet der Mikrozensus die reichhaltigste und umfassendste Datenquelle. Er wird mit einer sehr hohen Stichprobe erhoben und periodisch als Panel erhoben. Der Mikrozensus „[...] ist eine repräsentative Haushaltsbefragung der amtlichen Statistik in Deutschland. Rund 830 000 Personen in etwa 370 000 privaten Haushalten und Gemeinschaftsunterkünften werden stellvertretend für die gesamte Bevölkerung zu ihren Lebensbedingungen befragt. Dies sind 1% der Bevölkerung, die nach einem festgelegten statistischen Zufallsverfahren ausgewählt werden“ (Statistisches Bundesamt, 2016). Die seit 1957 jährlich durchgeführte Haushaltsbefragung stellt die einzige gesetzlich verpflichtende statistische Erhebung dar und enthält wichtige Daten zur Bevölkerungsstruktur sowie zur wirtschaftlichen



Van de Water, D; Steinmüller, B.: Hochschulbildung

und sozialen Lage der deutschen Bevölkerung. Für das Projekt E^B stellt diese eine der zentralen Datenquellen zur Beschreibung potentieller Zielgruppen der evidenzbasierten Angebotsentwicklung dar und liefert, nach Reanalyse, interessante demographische Kennzahlen der Bevölkerung. Dazu zählen insbesondere Altersstruktur, Geschlechterverhältnis, Staatsangehörigkeit, Familienstand, Kinderanzahl im Haushalt, sowie Variablen zu Bildung, Beruf und Arbeitszeiten (u.a. höchster beruflicher Abschluss, Nettoeinkommen, Verteilung der Arbeitszeit der Erwerbspersonen, Samstags-/Sonntagsarbeit etc.).

„Weiterbildung ist gekennzeichnet durch Heterogenität: Verschiedene Personengruppen nehmen in unterschiedlichen Situationen oder Kontexten verschiedene Weiterbildungsaktivitäten wahr – sei es aufgrund verschiedener Zugangsmöglichkeiten oder Teilnahmemotivationen“ (BMBF, 2015, S. 18). Dieses differenzierte Verständnis der Lage und der Möglichkeiten von Weiterbildungsrezipient_innen spiegelt sich in der Projektlogik von E^B wider. Insofern bietet der Adult Education Survey, aus dessen selbstbeschreibender Einleitung dieses Zitat stammt, eine weitreichende Datenquelle, um die sozialstrukturellen Daten der Region Merkmalen von Weiterbildungsaktivitäten gegenüberzustellen. In ihm werden nicht Weiterbildungsbeteiligungen erhoben, sondern auch Informationen über deren Beschreibungen, wodurch Analysen nicht nur auf Personenebene sondern auch auf Ebene der Weiterbildungsaktivitäten möglich werden (BMBF, 2015). Die Betrachtung erfolgt unterteilt nach betrieblicher, individuell-berufsbezogener und nicht-berufsbezogener Weiterbildung, dabei nach zeitlichen und finanziellen Rahmenbedingungen unterscheidend und hinsichtlich der Teilnahmemotive kontrastierend.

Beide Datenquellen ergänzen sich dahingehend, als sie miteinander kombiniert Rückschlüsse auf Bedürfnisse und Anforderungen neuer Zielgruppen von Weiterbildung zulassen. Indem sie auf die E^B-Bildungsregion rekurrieren, tragen sie dazu bei, regionale Besonderheiten aufzudecken bzw. regionalen Förderbedarf sichtbar zu machen.

Bildungsmarkt

Die Darstellung des wissenschaftlichen Bildungsmarktes wird mit Daten des Hochschulkompasses realisiert (Hochschulrektorenkonferenz, 2016). Dieser bietet Auskunft über in Deutschland bestehende Studienangebote mit verschiedenen Abschlüssen vom Erststudium bis zur Promotion und beinhaltet Vollzeitangebote genauso wie alternative Studienformen. Der technische Datenzugriff erfolgt über eine REST-Schnittstelle (Representational State Transfer, zur

Van de Water, D; Steinmüller, B.: Hochschulbildung

Technik vgl. Fielding, 2000), die eine automatisierte und individualisierte Datenabfrage ermöglicht. Der Zugang zum Hochschulkompass wird vom Projektpartner Virtueller Campus Rheinland-Pfalz (VCRP, 2016) bereitgestellt. Dies impliziert eine Besonderheit im Datenzugriff, weil die Daten lediglich durch spezifizierte Suchanfragen an eine entfernte Datenbank zugänglich sind.

4 Informationstechnische Grundlagen

4.1 Allgemeine Systemstruktur

Der Regionalmonitor ist ein webbasiertes System und ist strukturell gemäß dem klassischen Client-Server-Modell entworfen (Tanenbaum, 2003, S. 18). Auf der Seite des Clients (Frontend) wird den Nutzerinnen und Nutzern über einen Browser eine grafische Benutzungsoberfläche (Graphical User Interface, GUI) bereitgestellt, welche über interaktive Elemente eine individuelle Datenauswahl und Datenvisualisierung ermöglicht. Auf der Seite des Servers (Backend) werden alle systemrelevanten und statistischen Daten in einer Datenbank verwaltet, wobei die Schnittstelle zwischen Benutzereingaben und Datenselektion durch ein serverseitiges Skript gesteuert wird. Der Einsatz des Systems in verschiedenen Systemumgebungen wird einerseits durch den Gebrauch nicht-proprietärer Software und andererseits durch die Berücksichtigung verschiedener Plattformen während der Programmierphase gewährleistet.³

Der Einsatz moderner Webtechnologien (Pomaska, 2012) ermöglicht eine hohe Usability. So werden etwa dynamische Inhalte nicht durch ein Refresh der gesamten Webseite geladen, sondern lediglich die neu benötigten Daten mittels *Asynchronous Java Script und XML* (kurz: AJAX) beim Server angefordert und nach Erhalt durch Manipulation des Document Object Models (kurz: DOM) in die bestehende Seitenstruktur integriert. Dadurch wird die Reaktionszeit auf benutzergesteuerte Ereignisse deutlich reduziert.

³ So ist die Schnittstelle zur Datenbank software-technisch modular aufgebaut, sodass prinzipiell verschiedene relationale Datenbanken unterstützt werden; des Weiteren differiert das Verhalten der Skriptsprache PHP in manchen Bereichen auf verschiedenen Betriebssystemen wie etwa bei der Angabe von Dateipfaden.

Van de Water, D; Steinmüller, B.: Hochschulbildung

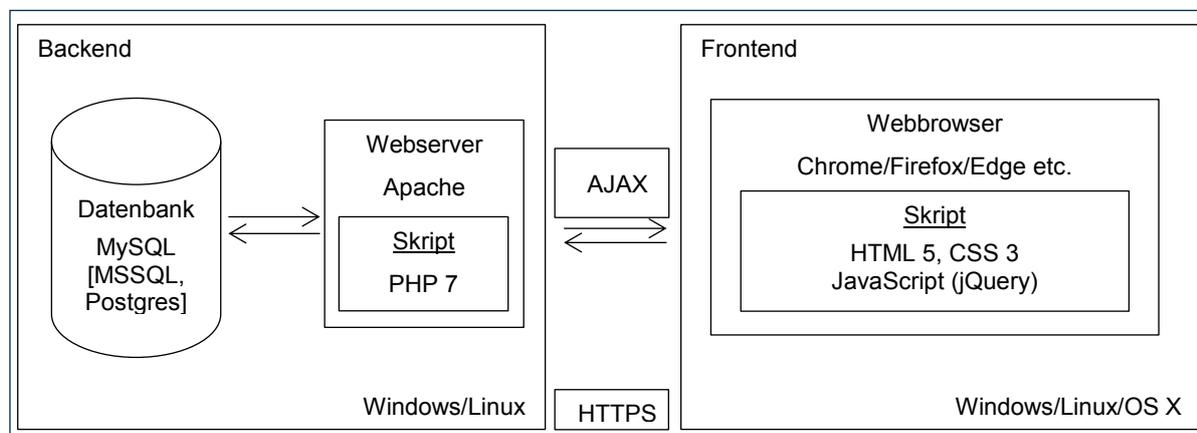


Abbildung 2 Systemstruktur und verwendete Programmiersprachen (eigene Darstellung)

In Abbildung 2 Systemstruktur und verwendete Programmiersprachen (eigene Darstellung) ist die grundlegende Systemstruktur mit den unterstützten Systemen und verwendeten Programmier- bzw. Auszeichnungssprachen dargestellt. Serverseitig werden Systemumgebungen mit Linux- und Windowsbetriebssystemen unterstützt.

Die Datenverwaltung wird durch die Verwendung einer relationalen Datenbank realisiert (Studer, 2016), die in Bezug auf Änderungen und Ergänzungen der Datenstruktur eine hohe Flexibilität bietet (Steiner, 2014, S. 12). Auf diese Weise wird den Anforderungen der Nachhaltigkeit und Flexibilität Rechnung getragen, so dass neue Datenquellen hinzugefügt und Anpassungen, die bei der Übertragung des Systems auf eine andere Region erforderlich werden, möglich sind.⁴ Eine höhere Kompatibilität wird durch den Verzicht auf datenbankspezifische Komponenten wie `constraints` erreicht. Dies gilt beispielsweise für die auf Microsoft SQL-Servern (kurz: MSSQL) erlaubten CHECK-Einschränkungen (Microsoft Corporation, 2016), durch die der gültige Wertebereich einer Tabellenspalte definiert und die Datenintegrität erhöht wird. Da diese etwa von MySQL-Servern nicht unterstützt werden, wird auf die datenbankseitige Überprüfung der Integritätsbedingungen verzichtet, wodurch das System auf beiden Datenbanksystemen lauffähig bleibt.

⁴ Zu beachten ist, dass die Flexibilität der Datenbank entscheidend von der Struktur des Datenschemas abhängt. So erlauben „flache Tabellen“ etwa eine einfache Verknüpfung von Tabellen untereinander (Staud 2005, 45).

Van de Water, D; Steinmüller, B.: Hochschulbildung

Die serverseitigen Skripte sind in PHP programmiert. Unter den serverseitigen Programmiersprachen für Webseiten ist sie mit 82,1%⁵ die meist genutzte Sprache (W3Techs, 2016a). Generell lässt die starke Verbreitung einer Software eine große Community und eine stetige Weiterentwicklung insbesondere mit Blick auf Sicherheitslücken vermuten. Ferner bietet PHP eine große Flexibilität hinsichtlich der Unterstützung von Datenbanken und Webservern (A-chour et al., 2016). Das Entwicklungs- und Livesystem wird auf einem Apache-Webserver betrieben (The Apache Software Foundation, 2016), welcher mit 52,0% Marktanteil ebenfalls am häufigsten eingesetzt wird (W3Techs, 2016b).

Clientseitig wird ein Browser vorausgesetzt, der die modernen Webtechnologien wie HTML5 und CSS 3 ausreichend unterstützt, was auf verschiedenen Webseiten überprüfbar ist (Leenheer, 2016). Weitestgehend sind interaktive Elemente mit jQuery umgesetzt (The jQuery Foundation, 2016) – ein JavaScript-Framework, welches die Browserkompatibilität für die gängigsten Browser bereitstellt. Die Kommunikation zwischen Client und Server geschieht verschlüsselt über eine HTTPS-Verbindung.

4.2 Software-Architektur

Das Websystem ist modular in Anlehnung an das 3-Schichten-Modell aufgebaut (Vogel et al., 2009, S. 67). In Abbildung 3 sind die einzelnen Komponenten überblicksartig dargestellt.

⁵ Bei dieser Statistik ist zu berücksichtigen, dass einzelne Webseiten mehrere Skriptsprachen verwenden können.

Van de Water, D; Steinmüller, B.: Hochschulbildung

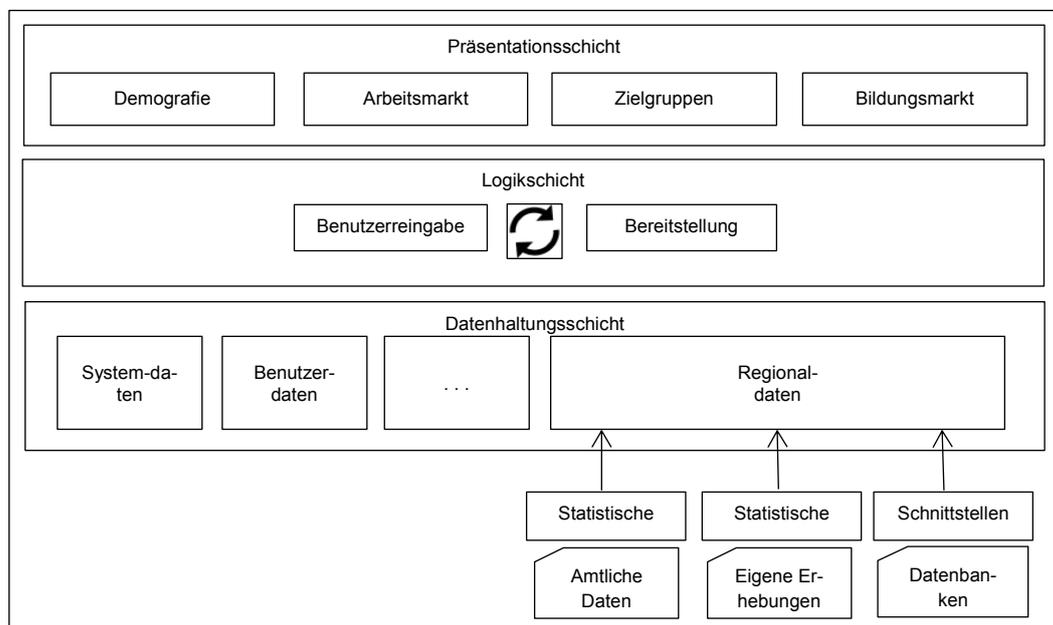


Abbildung 3 Grundlegende Software-Architektur (eigene Darstellung)

In der Datenerhaltungsschicht sind jegliche Regionaldaten und die meisten Systemdaten organisiert. Lediglich für die Initialisierung benötigte Informationen, wie etwa die Zugangsdaten zur Datenbank, werden separat in einer Konfigurationsdatei gespeichert. Das Datenbankschema wird während des Installationsvorgangs umgesetzt, indem alle benötigten Tabellen mit Defaultwerten und den zu dieser Zeit verfügbaren Regionaldaten angelegt werden. Aus datenschutzrechtlichen Gründen werden keine Einzeldatensätze gespeichert, was eine Aggregation der Daten im Vorfeld erfordert.

So werden Rohdaten zunächst statistisch etwa in Form von Häufigkeiten aufbereitet, in CSV-Dateien konvertiert und dann in die Datenbank überführt. Die Logikschicht stellt die eigentliche Programmfunktionalität zur Verfügung, wobei sie die Verbindungskomponente zwischen den Eingaben in der Benutzungsschnittstelle und den in der Datenbank enthaltenen Regionaldaten darstellt. Nach Annahme und Validierung der Benutzereingaben werden gemäß der angesprochenen Programmfunktion die entsprechenden Daten aus der Datenbank extrahiert, für die grafische Anzeige aufbereitet und an die anfragende Stelle zurückgegeben. Die Präsentationsschicht dient als Schnittstelle zur Interaktion zwischen dem Programm und den Nutzern. Durch die grafische Gestaltung können die einzelnen Bedienelemente weitgehend mit der Maus angesprochen werden. Eine Ausnahme bildet lediglich die Eingabe von Logindaten. So

Van de Water, D; Steinmüller, B.: Hochschulbildung

können die Regionaldaten von Anwendern explorativ erkundet und nach ihren Bedürfnissen in der Darstellung angepasst werden.

4.3 Grundlegende Systembausteine

Die Benutzerführung, insbesondere die Zugriffssteuerung auf einzelne Programmodule, wird durch ein Benutzer- und Rollensystem unterstützt. Standardmäßig werden bei der Installation drei Rollenprofile angelegt. Ein *Administrator* kann auf alle Grund- und Verwaltungsfunktionen des Programms zugreifen. Benutzerinnen und Benutzer mit der Rolle *Authenticated* haben Zugriff auf ausgewählte Grundfunktionen wie etwa Daten zum Arbeitsmarkt. Die Rolle mit der niedrigsten Berechtigungsstufe ist mit *Anonym* bezeichnet und erlaubt lediglich die Ansicht grundlegender Elemente der Benutzungsoberfläche wie das Loginformular und das Hauptmenü. Eine Verwendung des Programms ohne Benutzeranmeldung ist nicht möglich.

Zum Zeitpunkt des Aufrufs der Startseite, an dem sich noch keine Person eingeloggt hat, wird automatisch ein Benutzer mit der Rolle *Anonym* eingeloggt, sodass zu jeder Zeit der Aufruf einzelner Programmfunktionen mit der Überprüfung der jeweils benötigten Berechtigungsstufe verbunden ist. Über eine separate Eingabemaske sind die einzelnen Zuweisungskombinationen von Rolle und Programmfunktionen einsehbar. Zudem können die Berechtigungen bestehender Rollen konfiguriert und neue Rollen definiert werden. Bei der Installation werden zwei Standardbenutzer – jeweils ein Benutzer mit der Rolle *Administrator* und der Rolle *Anonym* – angelegt, die für den Betrieb zwingend erforderlich sind. Diese beiden Standardbenutzer und ihre Rollen können daher nicht gelöscht werden. Zu beachten ist, dass die Existenz eines Benutzers ohne Rolle nicht möglich ist, wodurch das Löschen einer Rolle das Löschen aller Benutzer mit dieser Rolle nach sich zieht. Informationen zu aktuellen Benutzersessions werden verschlüsselt in der Datenbank gespeichert.

Die Verwaltung der Regionaldaten ist über ein eigenes Dialogfenster möglich. In zwei unabhängigen Schritten können CSV-Dateien zunächst hochgeladen und dann in die Datenbank eingelesen werden. So ist die Möglichkeit gegeben je Datenquelle mehrere Dateien zu verwalten und die aktuell im System verwendete Datei auszuwählen. Ferner können neue externe Datendateien zunächst auf Ihre Kompatibilität hin getestet und im Fehlerfall auf eine ältere Version zurückgegriffen werden. Bereits beim Hochladen wird die strukturelle Validität jeder Datei überprüft. Beim Einlesen einer Datendatei kann die Zuordnung von CSV-Spalten und

Van de Water, D; Steinmüller, B.: Hochschulbildung

Datenbankfeldern notwendig sein. Dies tritt auf, wenn sich die Struktur der Datendateien, welche von externen Institutionen generiert werden, geändert hat. Dazu zählen etwa die Reihenfolge und die Bezeichnung von Spalten in CSV-Dateien.

In der Systemverwaltung können allgemeine Systemeinstellungen eingesehen und konfiguriert werden. Dazu zählen die Gestaltung der grafischen Benutzeroberfläche (Theme), die zu verwendende Sprache oder je nach Datenquelle die Region, für die Regionaldaten anzuzeigen sind. Weiterhin können Backups von editierbaren Seitenelementen erstellt, importiert und exportiert werden. Dazu gehören beispielsweise Erklärungsseiten, die Impressumsseite und die Fußzeile. Beim Backup werden die dazugehörigen HTML-Texte und medialen Elemente wie Bilder und Videos in einer ZIP-Datei zusammengefasst und in einem Backup-Ordner abgelegt. Einerseits wird so das Erstellen von Sicherheitskopien möglich, welche bei Datenverlust (etwa durch versehentliches Löschen) zur Wiederherstellung genutzt werden können. Andererseits können so Inhalte zwischen verschiedenen Instanzen des Programms – wie einem Test- und einem Produktivsystem – ausgetauscht werden. Die Integrität der Backupdateien wird über einen Hashcode überprüfbar, welcher standardmäßig mit dem SHA-512-Algorithmus aus der SHA-2-Familie erzeugt wird (Eastlake & Hansen, 2006).

Das Programm unterstützt Mehrsprachigkeit, so dass durch entsprechende Bedienelemente die gewünschte Sprache aus dem Spektrum der installierten Sprachpakete auswählbar ist. Bei der Installation wird die Standardsprache festgelegt, die beim Aufruf der Webseite Anwendung findet. Ferner verfügt das System über ein integriertes Autorensystem. Nach dem Einloggen und bei entsprechender Berechtigung des Benutzers, können die Inhaltsseiten für jede Sprache separat erstellt beziehungsweise editiert werden. Dabei bleibt die Struktur der Benutzeroberfläche vollständig erhalten, da je Seite lediglich ein zusätzlicher Bearbeitungsbutton eingeblendet wird. Für die Editierung erscheint ein separates Dialogfenster, der einen WYSIWYG-Editor (What You See Is What You Get) bereitstellt, welcher technisch mit dem CKEditor-Plugin realisiert ist (CKSource, 2016). Die einzelnen Inhaltsseiten können ohne Programmierkenntnisse mit grundlegenden Funktionen der Textformatierung editiert werden, wobei gleichzeitig eine HTML-Code-Ansicht für Fortgeschrittene zur Verfügung steht.

5 Chancen und Grenzen

Die Digitalisierung ermöglicht die Integration verschiedener Datenquellen. So können die verschiedenen Perspektiven auf die Entwicklung von Studienangeboten miteinander in Beziehung gesetzt werden, wie es bisher in dieser Form nicht möglich ist: Daten aus unterschiedlichen Erhebungen, und das heißt letztlich Antworten auf unterschiedliche Fragestellungen, können im Regionalmonitor Hochschulbildung neue Bedeutung gewinnen, da sie einander ergänzen. Die jeweiligen Daten unterscheiden sich allerdings in ihrer inhaltlichen Aussagekraft und decken jeweils unterschiedliche Bereiche ab, die für den Regionalmonitor Hochschulbildung von Interesse sind. Denn die genutzten Daten entstammen verschiedenen Erhebungen, die erstens unterschiedliche Merkmale erfassen und zweitens unterschiedliche Cluster der Merkmale bilden. Daher entsteht eine verschwommene Granularität, beispielsweise wenn mehrere Berufe aufgrund zu geringer Fallzahlen zu Berufsgruppen zusammengefasst werden. Solche Zusammenstellungen sind immer in gewissem Umfang willkürlich, und diese Berufsgruppen decken sich – um im Beispiel zu bleiben – nicht unbedingt mit den wissenschaftlichen Fachgebieten. So entstehen durch die Integration ursprünglich unabhängig erhobener Daten in ein Tool gleichzeitig neue Perspektiven wie auch Unschärfen.

Ist diese Datenintegration in automatisierter Form möglich, etwa durch Datenbankschnittstellen wie REST, so kann eine permanente und nachhaltige Bereitstellung der Daten sichergestellt werden. Das hieße, der Regionalmonitor Hochschulbildung kann mit verhältnismäßig geringem Aufwand viele Jahre aktuell bleiben. Voraussetzung dafür ist allerdings die Zugänglichkeit zu regelmäßigen, standardisierten Erhebungen.⁶ So schafft die Konstruktion des Monitors einerseits eine gewisse Nachhaltigkeit, macht sich aber gleichzeitig von der Zulieferung anderswo (regelmäßig) erhobener Daten abhängig. Insofern lässt sich die Aussagekraft des Regionalmonitors Hochschulbildung nur dann aufrechterhalten, wenn die vorgesehenen spezifischen Zusatzerhebungen periodisch erneuert werden.

Für die Nutzer_innen bietet der Monitor die Möglichkeit der individualisierten und zielorientierten Betrachtung der Daten, welche insbesondere durch die Einbindung interaktiver grafischer Elemente erleichtert wird. Die Vielzahl und Unterschiedlichkeit der einzelnen Datenquellen

⁶ In der Hauptsache entstehen Herausforderungen bei der Zugänglichkeit: Institutionalisierte Erhebungen bieten zwar dauerhaft Daten, sind in ihrer Aussagekraft aber in Bezug auf die Region limitiert. Ergänzende Zusatzerhebungen, wie sie im Rahmen von E^B durchgeführt werden, schließen diese Lücke, werden aber nicht zwangsläufig regelmäßig und dauerhaft durchgeführt.

Van de Water, D; Steinmüller, B.: Hochschulbildung

verlangt ihnen gleichzeitig einen kritischen und reflektierten Blick auf die Ergebnisse ab – etwa hinsichtlich der Repräsentativität und der regionalen Tiefe.

So ergeben sich in Bezug auf die Entscheidung für oder gegen ein neues Studienangebot nur eingeschränkt belastbare Indikatoren dafür, ob ein Bedarf in einem bestimmten Bereich existiert und ob genügend Personen bereit wären an einem entsprechenden Angebot teilzunehmen. Der Regionalmonitor Hochschulbildung hat zwangsläufig nur eine eingeschränkte, transparent beschriebene Aussagekraft, aber er bietet evidenzbasierte Indizien für aktuelle und zukünftige Bedarfe und kann daher als unterstützende Entscheidungsgrundlage heran gezogen werden.

Literaturverzeichnis

Achour, Mehdi et al. (2016). *PHP-Handbuch*. Zugriff am 02.08.2016. Verfügbar unter <http://php.net/manual/de/>.

Bundesministerium für Bildung und Forschung (Hg.) (2015). *Weiterbildungsverhalten in Deutschland 2014. Ergebnisse des Adult Education Surveys. AES Bericht*. Zugriff am 09.11.2016. Verfügbar unter https://www.bmbf.de/pub/Weiterbildungsverhalten_in_Deutschland_2014.pdf.

CKSource (2016). *CKEditor. The best web text editor for everyone*. Zugriff am 03.08.2016. Verfügbar unter <http://ckeditor.com/>.

Eastlake, D., & Hansen, T. (2006). *US Secure Hash Algorithms (SHA and HMAC-SHA)*. Request for Comments (RFC) 4634. Informational. Zugriff am 03.08.2016. Verfügbar unter <https://tools.ietf.org/html/rfc4634> (obsoleted by RFC 6234).

Emmerich, M (2010). Regionalisierung und Schulentwicklung: Bildungsregionen als Modernisierungsansätze im Bildungssektor. In: Altrichter, H., Maag Merki, K. *Handbuch Neue Steuerung im Schulsystem*. Wiesbaden: VS-Verlag.

Fielding, R. Th. (2000). *Architectural Styles and the Design of Network-based Software Architectures* (Dissertation). Zugriff am 05.08.2016. Verfügbar unter https://www.ics.uci.edu/~fielding/pubs/dissertation/fielding_dissertation.pdf.

Hochschulrektorenkonferenz (2016). *Hochschulkompass. Studieren und promovieren in Deutschland*. Zugriff am 05.08.2016. Verfügbar unter <http://www.hochschulkompass.de/>.

IHK-Arbeitsgemeinschaft Rheinland-Pfalz (2016). *IHK Fachkräftemonitor Rheinland-Pfalz. Das Instrument zur Analyse von Fachkräfteangebot und -nachfrage in Rheinland-Pfalz, sowie in seinen Regionen und Branchen bis zum Jahr 2030*. Zugriff am 03.08.2016. Verfügbar unter <http://www.fachkraeftemonitor-rlp.de/>.

Leenheer, N. (2016). *HTML 5 Test. How well does your browser support HTML5?* Zugriff am 02.08.2016. Verfügbar unter <http://html5test.com/>.



Van de Water, D; Steinmüller, B.: Hochschulbildung

Marks, S. (2015). *Region als Bezugsraum für Hochschulentwicklung. Regionsdefinition für das Projekt E^B*. In Wolf, K., Rohs, M. & Arnold, D. (Hrsg.), *Arbeits- und Forschungsberichte aus dem Projekt E^B – Bildung als Exponent individueller und regionaler Entwicklung*, Nr. 1. Kaiserslautern: Technische Universität Kaiserslautern.

Microsoft Corporation (2016). *CHECK-Einschränkungen*. Zugriff am 02.08.2016. Verfügbar unter [https://technet.microsoft.com/de-de/library/ms188258\(v=sql.105\).aspx](https://technet.microsoft.com/de-de/library/ms188258(v=sql.105).aspx).

Pomaska, G. (2012). *Webseiten-Programmierung: Sprachen, Werkzeuge, Entwicklung*. Heidelberg: Springer.

Prognos AG (2016). *Prognos Zukunftsatlas 2016. Das Ranking für Deutschlands Regionen*. Zugriff am 12.12.2016. Verfügbar unter https://www.prognos.com/fileadmin/images/publikationen/Zukunftsatlas2016/Prognos_Zukunftsatlas_2016_Produktblatt.pdf

Rohs, M., Vogel, C., & Marks, S. (2015). *From supply-driven to demand-oriented academic education: Evidence-based development of study courses to match regional skill shortage with new student groups*. 12th PASCAL International Observatory Conference in Catania, Sizilien. Unveröffentlichtes Manuskript.

Statistische Ämter des Bundes und der Länder (2016a). *Datenzugang. Anonymität von Mikrodaten*. Zugriff am 04.08.2016. Verfügbar unter <http://www.forschungsdatenzentrum.de/anonymisierung.asp>.

Statistische Ämter des Bundes und der Länder (2016b). *Datenzugang*. Zugriff am 04.08.2016. Verfügbar unter <http://www.forschungsdatenzentrum.de/datenzugang.asp>.

Statistisches Bundesamt (2016). *Destatis*. Zugriff am 09.11.2016. Verfügbar unter <https://www.destatis.de>

Staud, J. L. (2005). *Datenmodellierung und Datenbankentwurf ein Vergleich aktueller Methoden*. Berlin: Springer.

Steiner, R. (2014). *Grundkurs Relationale Datenbanken: Einführung in die Praxis der Datenbankentwicklung für Ausbildung, Studium und IT-Beruf*. Wiesbaden: Springer Vieweg.

Stern, C., Ebel, Chr., Schönstein, V. (2008). *Bildungsregionen gemeinsam gestalten. Erfahrungen, Erfolge, Chancen*. Gütersloh: Verlag Bertelsmann Stiftung.

Studer, T. (2016). *Relationale Datenbanken: Von den theoretischen Grundlagen zu Anwendungen mit PostgreSQL*. Berlin: Springer Vieweg.

Tanenbaum, A. S. (2003). *Computernetzwerke* (4. Aufl.). München: Pearson Studium.

The Apache Software Foundation (2016). *Apache HTTP Server Project*. Zugriff am 02.08.2016. Verfügbar unter <https://httpd.apache.org/>.

The jQuery Foundation (2016). *jQuery. Write less, do more*. Zugriff am 02.08.2016. Verfügbar unter <https://jquery.com/>.

Technische Universität Kaiserslautern (2016). *Zusammenfassung der wesentlichen formalen Kriterien für die Gestaltung modularisierter Studiengänge und deren Akkreditierung*. Zugriff am 04.08.2016. Verfügbar unter <http://www.uni->



Van de Water, D; Steinmüller, B.: Hochschulbildung

kl.de/fileadmin/refls/tupublic/Handreichungen/
Handreichung_Formalkriterien_modularisierte_STG.pdf.

VCRP (2016). *Virtueller Campus Rheinland-Pfalz*. Zugriff am 09.11.2016. Verfügbar unter <https://www.vcrp.de/>.

Vogel, O., Arnold, I., Chughtai, A., Ihler, E., Kehrer, T., Mehlig, U., & Zdun, U. (2009). *Software-Architektur: Grundlagen - Konzepte - Praxis* (2. Aufl.). Heidelberg: Spektrum Akademischer Verlag.

W3Techs (2016a). *Usage of server-side programming languages for websites*. Zugriff am 02.08.2016. Verfügbar unter https://w3techs.com/technologies/overview/programming_language/all.

W3Techs (2016b). *Usage of web servers for websites*. Zugriff am 02.08.2016. Verfügbar unter https://w3techs.com/technologies/overview/web_server/all.

WifOR (2016). *Willkommen bei WifOR*. Zugriff am 05.08.2016. Verfügbar unter <http://www.wifor.de>.



ENTWICKLUNG DURCH BILDUNG

ISSN 2364-8996