

Datenbewusstsein

Wo, wie und wozu werden Daten gesammelt und verarbeitet? – Datenbewusstsein durch die Exploration von Standortdaten aus dem Mobilfunknetz

In diesem Dokument befinden sich ergänzende Informationen zu dem Unterrichtsmodul mit dem Titel *„Wo, wie und wozu werden Daten gesammelt und verarbeitet? – Datenbewusstsein durch die Exploration von Standortdaten aus dem Mobilfunknetz“* im Kontext eines Unterrichts zum Kontext von Datenbewusstsein¹:

1 Motivation

Im Alltag interagieren wir, v.a. auch Schülerinnen und Schüler, tagtäglich mit datengetriebenen Informatiksystemen. Ein datengetriebenes Informatiksystem erhebt in der Regel während einer Interaktion mit diesem, also dem Nutzer, persönliche Daten und verarbeitet diese für verschiedene Zwecke, wie etwa für das personalisierte Anzeigen von Produkten oder für standortbezogene Dienste (z.B. Empfehlungen). Die Daten werden entweder explizit durch die Nutzerin bzw. den Nutzer bereitgestellt (z.B. durch Eingabe von Profilinformationen auf einer Social Media Plattform) oder implizit bei der Interaktion erhoben, etwa durch Beobachtung der Nutzerinnen und Nutzer (z.B. die Klicks auf Ergebnisse in einer Suchmaschine oder das Verhalten beim Durchsehen eines News Feeds auf einer Social Media Plattform) oder durch Verarbeitung gesammelter Daten generiert.

Ein Beispiel für ein datengetriebenes Informatiksystem, das wir tagtäglich nutzen, ist das Mobilfunknetz, sei es zum Telefonieren, SMS-Schreiben oder den Internetzugang über das mobile Netz. Kaum einer ist sich dabei bewusst, dass dabei verschiedene Daten, z.B. Standortdaten, erhoben werden, wie diese implizite Erhebung von Daten abläuft und wozu dies etwa notwendig ist. Wo, wie und wozu werden diese Daten verarbeitet? In diesem Unterrichtsmodul werden diese Fragen aufgegriffen und geklärt.

In diesen ergänzenden Informationen werden verschiedene Hintergrundinformationen gegeben, die die Lehrkraft beim Unterrichten dieses Moduls hilfreich sind. Neben dem Aufbau und der Funktionsweise des Mobilfunknetzes, gehen wir hier auf das Puzzle aus dem Modul ein und erklären die interaktive Web-Anwendung mit anreichernden Beispielen.

Inhaltsverzeichnis

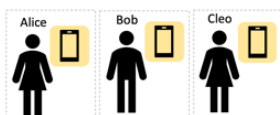
1	Motivation	1
2	Beschreibung von Materialien und weiterführenden Hintergrundinformationen zum Mobilfunknetz ..	2
2.1	Hinweise zum Puzzle über den Aufbau und die Funktionsweise des Mobilfunknetzes	2
	Elemente des Mobilfunknetzes:	2
2.2	Hinweise zu den explorierten Daten von Malte Spitz	3
2.3	Hinweise zur interaktiven Web-Anwendung	3
2.4	Beispiel für Interpretationen der visualisierten Standortdaten	5
2.5	Weiterführende Informationen zum Mobilfunknetz	5
2.6	Beispiele für sekundäre Zwecke der Verwendung von Standortdaten aus dem Mobilfunknetz	7
3	Glossar relevanter Begriffe:	8

2 Beschreibung von Materialien und weiterführenden Hintergrundinformationen zum Mobilfunknetz

2.1 Hinweise zum Puzzle über den Aufbau und die Funktionsweise des Mobilfunknetzes

Puzzleteile:

Personenplättchen:



Die drei Personenplättchen stellen symbolisch die drei Personen mit ihren Handys bzw. mobilen Stationen dar. Diese spielen in den Simulationen der Funktionsweise des Mobilfunknetzes in der zweiten Aufgabe des Arbeitsblattes 2 eine große Rolle.

Elemente des Mobilfunknetzes:



In dem Puzzle kommen stellvertretend für das reduzierte Mobilfunknetz neben den Personenplättchen vier Puzzleteile vor: 1) Antenne + Basisstation (viermal), 2) Vermittlungsstelle, 3) Datenspeicher, 4) Internet. Die Bedeutung der jeweiligen Elemente beschreiben wir im Abschnitt 2.5 genauer. Nachfolgend wird eine Lösung des Puzzles beschrieben. Die Puzzleteile sind, wie dort dargestellt, auf der Puzzlevorlage (s.u.) abzulegen, um das Mobilfunknetz adäquat darzustellen und die Funktionsweise daran zu simulieren. Die Schnittstelle zwischen Vermittlungsstelle und Datenbank wurde zur Reduktion der Komplexität durch eine direkt verbundene Anordnung ersetzt. Die Vermittlungsstelle und die Datenbank sind also als zwei verknüpfte Elemente zu verstehen. Das Puzzleteil Internet (4) steht in äußerst abstrakter Art und Weise stellvertretend für die Verbindung zwischen Vermittlungsstelle und „dem Internet“. Die Thematisierung des, wenn auch grundlegenden, Aufbaus des Internets in diesem Unterrichtsmodul würde im Sinne der Komplexität zu weit führen. (Denkbar wäre an dieser Stelle jedoch in Abhängigkeit der Lerngruppe ein sehr kurzer Exkurs zum Aufbau des Internets.)

Puzzlevorlage:

Die Puzzleteile und Personenplättchen werden in der ersten Aufgabe des Arbeitsblattes 2 im Sinne einer Rekonstruktion des Aufbaus des Mobilfunknetzes auf einer Puzzlevorlage abgelegt. Die Puzzlevorlage wird die Lernenden gemeinsam mit den Puzzleteilen und dem Arbeitsblatt 2 ausgehändigt. In durchgeführten Erprobungen des Unterrichtsmoduls hat sich gezeigt, dass es ratsam ist, die Materialien des Puzzles bereits zuvor ausgeschnitten in den Unterricht mitzubringen, um die effektive Lernzeit nicht wesentlich zu verkürzen, ein Ausschneiden im Unterricht durch die Schülerinnen und Schüler ist unter Umständen auch möglich.

Lösung des Puzzles:

Eine mögliche Lösung für das Puzzle in *Abbildung 1: Beispiel für eine Lösung des Puzzles (AB2)* dargestellt. Variationen bestehen lediglich in der Wahl der Funkzellen, in denen sich die drei Personen aufhalten. Die Elemente des Mobilfunknetzes (graue Puzzleteile) sollten nicht anders gewählt werden. Die Pfeile werden in der zweiten Aufgabe des Arbeitsblattes 2 auf das Puzzle gelegt, um das Herstellen einer Verbindung im Mobilfunknetz nachzustellen. Dazu werden zwei Szenarien zum Telefonieren (Situation 1: orangene Pfeile, Situation 2: grüne Pfeile) aufgegriffen. Optional kann ein drittes Szenario zum *Aufrufen einer Internetseite*

Datenbewusstsein durch die Exploration von Standortdaten aus dem Mobilfunknetz

(blaue Pfeile) als „Sprinteraufgabe“ bearbeitet werden. Die Pfeilrichtung stellt vereinfacht das Senden von Datenpaketen bei der Anfrage der jeweiligen Mobilfunkverbindung dar.

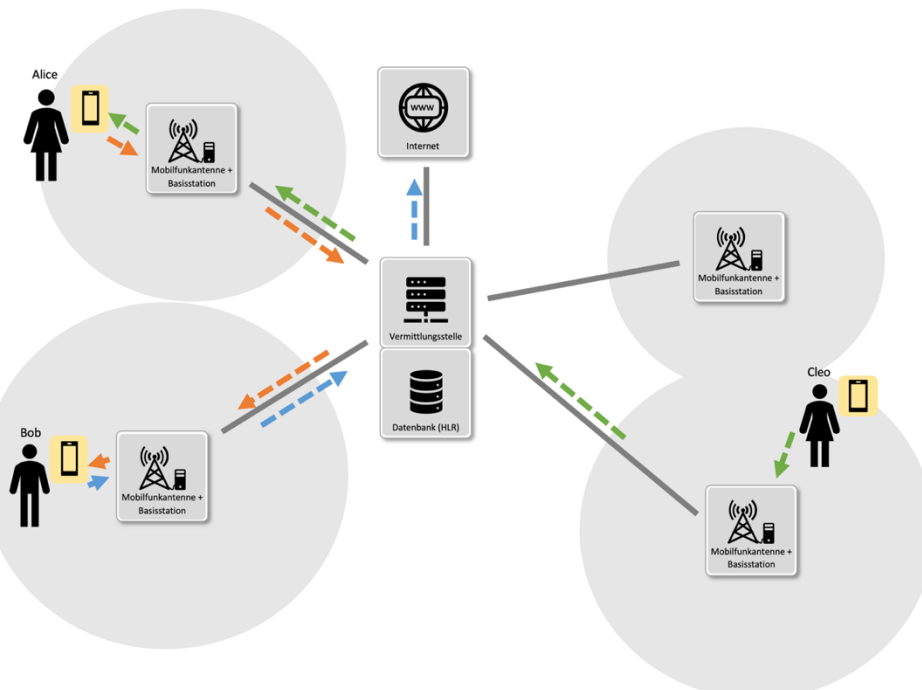


Abbildung 1: Beispiel für eine Lösung des Puzzles (AB2)

2.2 Hinweise zu den explorierten Daten von Malte Spitz

Die Standortdaten wurden während der Nutzung des Mobilfunknetzes von einem Mobilfunkanbieter erhoben. Die Datensätze wurden aus Gründen des Datenschutzes gefiltert, sodass nicht alle erhobenen Daten in der Tabelle aufgefasset werden und die Tabelle somit lediglich als Ausschnitt der erhobenen Daten bezeichnet werden sollte. Die Daten gehören zu einer Einzelperson – genauer zum Politiker Malte Spitz. Veröffentlicht wurden sie gemeinsam mit *Zeit Online* im Kontext des Diskurses um die Vorratsdatenspeicherung. Die inbegriffenen Daten umfassen einen Zeitraum zwischen September 2009 und Februar 2010. Bei der Interpretation der auf der Karte visualisierten Standortdaten sollte gegebenenfalls berücksichtigt werden, dass es sich um inzwischen ältere Daten auf einer aktuellen Karte handelt und somit die weiteren Hinweise auf der Karte (z.B. Cafe, Restaurant, Firmen) nicht unbedingt immer stimmen müssen. Dies hindert den Erkenntnisgewinn der Lernenden aus unserer Sicht jedoch nicht.

Weiterführende Quellen zu den Daten:

<https://www.zeit.de/digital/datenschutz/2011-02/vorratsdaten-malte-spitz>

<https://www.zeit.de/datenschutz/malte-spitz-vorratsdaten>

2.3 Hinweise zur interaktiven Web-Anwendung

Für das Unterrichtsmodul haben wir eine interaktive Webanwendung entwickelt. Diese ist unter folgendem Link erreichbar: go.upb.de/ExplorationStandortdaten

Sie basiert auf einem Jupyter Notebook und ist mit Python entwickelt worden. Von den Lernenden werden hingegen keine Kenntnisse im Umgang mit Jupyter Notebooks sowie keinerlei Programmierkenntnisse vorausgesetzt.

Zunächst müssen die Standortdaten eingelesen werden, was über den Button *Lade die Standortdaten* geht. Prinzipiell ist es auch möglich eigene Standortdaten einzuladen bzw. die Anwendung, um diese zu erweitern, die an dieser Stelle ausgewählt werden könnten.

Nach Laden der Standortdaten wird folgende Bedienoberfläche angezeigt:

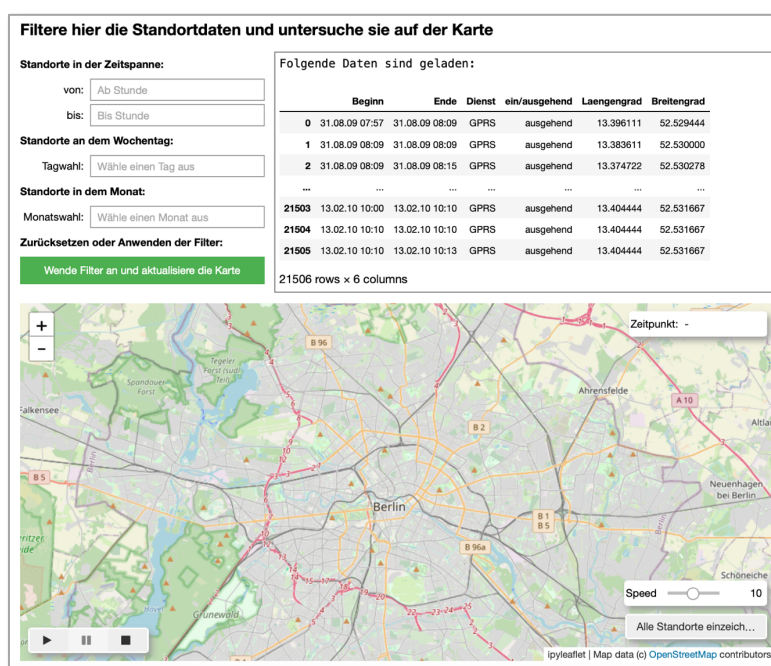


Abbildung 2: Ausschnitt aus der interaktiven Web-Anwendung

Sie besteht im Wesentlichen aus drei Bereichen: 1) Eingabemaske zum Einstellen von Filtern (oben links), 2) Anzeige der verarbeiteten Filter und der Tabelle der aktuellen Standortdaten (oben rechts) und 3) Karte zur Visualisierung der Standortdaten mit mehreren Widgets zur Steuerung der Visualisierung (unten).

In der Eingabemaske (1) können drei Filter separat oder zusammen eingestellt werden. Erstens kann nach einer Zeitspanne gefiltert werden, dann werden alle Standortdaten außerhalb dieser Zeitspanne entfernt. Es können lediglich ganze Stunden eingegeben werden, etwa 16 für 16:00 Uhr. Zweitens kann nach einem Wochentag gefiltert werden, etwa „Montag“, der als Textform in das Textfeld eingegeben bzw. aus dem Dropdown-Menü ausgewählt werden kann. Alle Standortdaten von Tagen außer dem gewählten Wochentag werden entfernt. Drittens kann nach einem Monat gefiltert werden, der ebenfalls als Text eingegeben oder aus dem Dropdown-Menü ausgewählt werden kann. Alle Standortdaten aus anderen Monaten werden entfernt. Ein oder mehrere verwendete Filter werden durch Klicken auf den Button *Wende Filter an und aktualisiere die Karte* auf die aktuellen Standortdaten angewendet. Die Karte wird dann zurückgesetzt und kann zum Visualisieren der aktuellen Standortdaten verwendet werden. Durch Klicken auf den Button *Setze alle Filter zurück* werden die ursprünglichen Standortdaten wiederhergestellt, sodass neue Filter gesetzt werden können.

In der Anzeige (2) werden die zuletzt angewendeten Filter mit jeweils einem Satz beschrieben sowie die aktuellen Standortdaten in tabellarischer Form dargestellt.

Auf der Karte (3) werden die Standortdaten visualisiert. Dynamisch können diese mit dem Play-Widgets unten links nach und nach eingeblendet werden. Dadurch kann der Verlauf der Standortdaten gut nachvollzogen werden. Dieses Einzeichnen der Standorte kann mit dem Pause Button in diesem Widget pausiert und mit dem Stopp Button zurückgesetzt werden. Die Geschwindigkeit des Einzeichnens kann mit dem Schieberegler *Speed* unten rechts angepasst werden. Der Button *Alle Standorte einzeichnen* kann als Ersatz für die dynamische Visualisierung der Standortdaten verwendet werden, um alle Standorte auf einmal einzeichnen zu lassen. Aus Performancegründen ist dieser jedoch für wenige Standortdaten empfohlen und auch reguliert. Oben rechts auf der Karte wird in einem Textfeld der Zeitpunkt des zuletzt eingezeichneten Standortes angegeben. Mit den Buttons Plus und Minus oben links auf der Karte kann die Zoomstufe der Kartenanzeige verändert werden.

Standorte werden als Marker dargestellt. Durch Halten der Maus über einen Marker wird ein Tooltip mit dem Zeitpunkt dieses Standortes eingeblendet. Wenn mehrere Marker nah beieinander liegen, werden diese zu einem Cluster (farbiger Kreis mit einer Zahl) zusammengefasst. Durch Anklicken des Kreises bzw. durch Hineinzoomen in die Karte werden die Cluster wieder (abhängig von der Zoomstufe) getrennt.

Kontaktieren Sie uns gerne bei Fragen, entdeckten Fehlern oder Verbesserungsideen. Wir sind stets daran interessiert, die Web-Anwendung weiterzuentwickeln.

2.4 Beispiel für Interpretationen der visualisierten Standortdaten

In den Explorationsphasen des Unterrichtsmoduls mit der interaktiven Web-Anwendung filtern und interpretieren die Lernenden die visualisierten Standortdaten. Da die Interpretationen subjektiv sind, werden sich diese in der Lerngruppe durchaus widersprechen. So wird für den Beruf typischerweise eine breite Vielfalt angeboten, die durchaus plausibel sind. Für die Interpretationen muss es im Unterrichtsverlauf auch kein richtig oder falsch geben. Wichtig ist, dass für die eigenen Interpretationen mit Bezugnahme auf die Daten argumentiert werden kann.

Beispiel einer Interpretation:

Die meisten Menschen schlafen nachts. In Deutschland würde das dann heißen, dass z.B. zwischen 3:00 bis 4:00 Uhr viele Menschen schlafen. Stellt man nun den Filter für die Zeitspanne auf 3 bis 4 Uhr ein, so stehen insgesamt 294 Standorte zur Verfügung. Nach Einzeichnen auf der Karte können 208 Standorte in Berlin lokalisiert werden. Schaut man diese nun genauer an, so befinden sich 189 von diesen in der Nähe der U-Bahn Haltestelle Rosenthaler Platz (s. Abbildung 3: Auswahl der Standorte für den Filter 3 bis 4 Uhr).



Abbildung 3: Auswahl der Standorte für den Filter 3 bis 4 Uhr

Weiteres Beispiel einer Interpretation:

Die Lernenden können dazu neigen, sehr stark zu interpretieren. So etwa folgendes Beispiel aus einer Erprobung: Ich habe die Standortdaten gefiltert und bekomme heraus, dass er sich oft bei einer bestimmten Firma aufgehalten hat. Die Firma ist in einer bestimmten Branche tätig, was ich mit einer Suchmaschine herausgefunden habe. Deswegen arbeitet er nun also etwa in der Finanzbranche.

Ein Schüler erwiderte darauf in einer Unterrichtserprobung, dass der Aufenthaltsort tagsüber bei einer Firma lediglich die Interpretation ermöglichen würde, dass er vielleicht dort arbeiten würde, aber nicht welchen Beruf er dort nachgeht. Er könnte etwa auch einer Hausmeister-Tätigkeit nachgehen.

Diskussionen dieser Art zeigen ein gutes Verständnis für die Gewinnung einer Information durch Interpretation von (Standort-) Daten und können sehr gewinnbringend für den Unterrichtsverlauf sein. Die Sicherheit der Interpretationen sollte also durchaus beachtet aber auch nicht unterschätzt werden. Die Auswertungsphase dieser Explorationen ist knapp bemessen. Zugunsten von Diskussionen dieser Art kann diese durchaus verlängert werden. Dabei kann dann auch aufgegriffen werden, dass bei solchen Datenauswertungen auch weitere Datenquellen hinzugenommen werden (s. optionale Ergänzung in Phase 2b), um genauere Interpretationen zu ermöglichen.

2.5 Weiterführende Informationen zum Mobilfunknetz

Vereinfachter Aufbau des Mobilfunknetz

Das Mobilfunknetz ist ein datengetriebenes digitales Artefakt. Ein solches wird von einem Netzbetreiber betrieben, von denen in Deutschland Telekom Deutschland, Vodafone und Telefónica Deutschland tätig sind. Seit der Einrichtung des ersten Mobilfunknetzes in den 80er Jahren haben sich die Architektur und die Relevanz des Mobilfunknetzes sowie die Handlungsmuster der Teilnehmer*innen erstaunlich

Datenbewusstsein durch die Exploration von Standortdaten aus dem Mobilfunknetz

weiterentwickelt. So zählten die drei genannten Netzbetreiber Ende 2020 insgesamt über 150 Mio. Teilnehmer*innen¹. Für das Unterrichtsmodul wird der Aufbau und die Funktionsweise des Mobilfunknetzes rekonstruiert, wobei Details ausgelassen werden, die für den Unterrichtsverlauf lediglich eine untergeordnete Rolle spielen. Das Mobilfunknetz besteht aus den folgenden Komponenten:

Mobile Station oder auch Endgerät oder Handy:

Die mobile Station besteht aus einem Handy und einer SIM-Karte. Mit der SIM-Karte wird eine eindeutige weltweite Kennung, die International Mobile Subscriber Identity (IMSI) vergeben, sodass jede Teilnehmerin und jeder Teilnehmer des Mobilfunknetzes damit identifiziert werden kann. Zusätzlich wird jeder SIM-Karte eine eindeutige Telefonnummer zugewiesen, die als eindeutiger Identifizierer der Teilnehmerin bzw. des Teilnehmers im Rahmen des Unterrichtsmoduls verwendet werden kann. Die Identifikation der mobilen Station hängt somit nicht vom eigentlichen Endgerät selbst ab, sondern von der SIM-Karte. Alternativ zu der physischen SIM-Karte wird seit Ende 2015 eSIM angeboten, was ein Verfahren zum Speichern der relevanten Informationen von Teilnehmer*innen auf dem Endgerät ist. Diese Informationen werden auf einem eingebauten nicht auswechselbaren Modul des Endgeräts (auch *Teilnehmer-Identitätsmodul* genannt) gesichert.

Basisstation oder auch Mobilfunkstation:

Die Basisstation, auch als Mobilfunkstation bezeichnet, wird umgangssprachlich oft Sendemast genannt, umfasst jedoch zusätzlich zur Antenne eine Recheneinheit. Eine Antenne besteht aus mehreren Teilen, die als Sektoren bezeichnet werden und in verschiedene Richtungen empfangen und senden können. Eine schematische Darstellung ist in Abbildung 4 dargestellt. Die Antenne sendet und empfängt die elektromagnetischen Wellen zwischen Basisstation und Mobilfunkteilnehmer*in, was zum Telefonieren erforderlich ist. Die empfangenen Signale werden von der Basisstation per Kabel (i.d.R. Glasfaser) zur Vermittlungsstelle weitergeleitet. Eine Basisstation deckt mit ihrem Sendebereich eine Funkzelle ab. Jeder Basisstation ihrer Funkzelle wird eine Cell-ID zugeordnet, die eine Identifikation, u.a. mit dem Standort, zulässt.

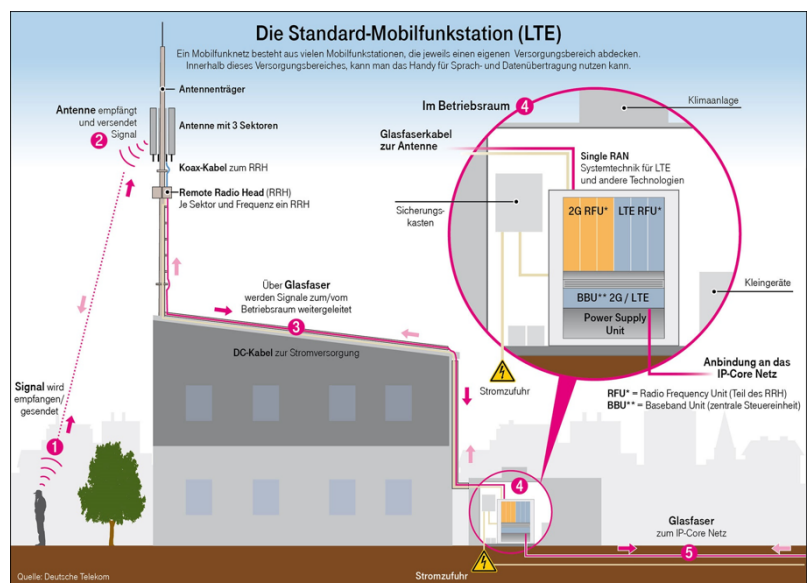


Abbildung 4: Aufbau einer Basis- bzw. Mobilfunkstation von <https://www.telekom.com/de/blog/netz/artikel/heptaband-antenne-5g-530618> (Detailgrad ist nicht für den Unterricht geeignet.)

Funkzelle:

Die Funkzelle stellt den Bereich dar, in dem eine Verbindung zwischen einem Endgerät und der jeweiligen Basisstation hergestellt werden kann. Die Zellgrößen unterscheiden sich im Durchmesser von einigen Metern bis einigen Kilometern. Benachbarte Funkzellen überlappen sich normalerweise, wodurch etwa ein unterbrechungsfreier Übergang von einer Funkzelle zu nächsten ermöglicht werden kann. Verschiedenste

¹ Entnommen von:

https://www.bundesnetzagentur.de/DE/Sachgebiete/Telekommunikation/Unternehmen_Institutionen/Marktbeobachtung/Deutschland/Mobilfunkteilnehmer/Mobilfunkteilnehmer_node.html, zuletzt eingesehen am 14.04.2021

Einflussfaktoren wie das Wetter, Häuser und Bäume oder auch geografische Gegebenheiten beeinflussen die Größe einer Funkzelle. Alle Funkzellen zusammen bilden die Abdeckung des Mobilfunknetzes.

Kontrollstation:

Eine Kontrollstation übernimmt die Verwaltung mehrerer Basisstationen und übernimmt die Datenübertragung zwischen einer Basisstation und einer Vermittlungsstelle. Für das Unterrichtsmodul wird die Kontrollstation nicht explizit ausgewiesen, um die Komplexität des Aufbaus zu reduzieren.

Vermittlungsstelle und zentrale Datenbank:

Eine Vermittlungsstelle wird Mobile-services Switching Centre (MSC) genannt und übernimmt die Vermittlung von Verbindungen zwischen Mobilfunkteilnehmer*innen. Außerdem wird dort eine Verbindung zu anderen Netzen, wie dem Festnetz hergestellt. Weitere Aufgaben einer Vermittlungsstelle sind etwa die Überprüfung der Berechtigung einer mobilen Station für eine Mobilfunkverbindung und die Abrechnung von Gebühren. Eine Vermittlungsstelle besitzt zahlreiche Schnittstellen zu weiteren Bestandteilen des Mobilfunknetzes. Eine davon ist die Verbindung zu der verteilten Datenbank Home Location Register (HLR). In dieser sind Daten über die mobilen Stationen abgespeichert. Dadurch kann eine Vermittlungsstelle u.a. eine mobile Station bzw. den/die jeweiligen Mobilfunkteilnehmer*in im Mobilfunknetz mit der aktuell eingewählten Basisstation lokalisieren. In dem Unterrichtsmodul werden gerade diese aufgezeichneten Standortdaten thematisiert, die für das effiziente Herstellen einer Mobilfunkverbindung notwendig sind. Für das Unterrichtsmodul wird auf eine Vermittlungsstelle mit einer Datenbank reduziert, obwohl ein Mobilfunknetz eigentlich viele Vermittlungsstellen beinhaltet.

Methoden zur Ortung im Mobilfunknetz

Zum Lokalisieren bzw. Orten einer mobilen Station bzw. eines Endgeräts gibt es verschiedene Möglichkeiten, die an dieser Stelle kurz umrissen werden. In dem Unterrichtsmodul werden die Standortdaten der mit der mobilen Station verbundenen Basisstation betrachtet. Diese können aufgrund der Größe der jeweiligen Funkzellen ungenau sein. Mit der Zeit wurden verschiedene Messverfahren entwickelt, die eine genauere Ortung erlauben. So wird etwa bei der Methode Timing Advance die Laufzeit des Signals zwischen der mobilen Station und der Basisstation gemessen, um die Entfernung zwischen diesen beiden Elementen abzuschätzen. Dadurch kann eine höhere Genauigkeit der Ortung erfolgen. Eine weitere Methode ist Enhanced Observed Time Difference (E-OTD), bei der Messung einer Laufzeit zwischen mobiler Station und Basisstation für mehrere benachbarte Basisstationen vorgenommen wird. Die Standortdaten des Unterrichtsmoduls werden oft mit relativ genauen GPS-Daten verglichen. Tatsächlich verwendet die Methode Global Navigation Satellite System (GNSS) eine satellitengestützte Positionsbestimmung seitens des Endgeräts, welches diese Daten an die Basisstation weiterleitet. Mit diesem Verfahren können Genauigkeiten von wenigen Metern Abweichung erzielt werden.

2.6 Beispiele für sekundäre Zwecke der Verwendung von Standortdaten aus dem Mobilfunknetz

Beispiel 1: Analyse der Mobilität während der Coronapandemie

Zur Eindämmung der Covid-19-Pandemie wurden verschiedene Maßnahmen (u.a. „Lockdown“) veranlasst. Die Wirkung der Maßnahmen kann anhand der Mobilität der Bevölkerung untersucht werden. Dazu wurden etwa Mobilfunkdaten des Mobilfunkanbieters Telefónica bereitgestellt, die zuvor anonymisiert und aggregiert wurden. Ein Bericht dazu findet sich bei destatis. Dort sind ebenfalls interessante Visualisierungen eingefügt, welche im Unterricht mit aufgegriffen werden könnten.

<https://www.destatis.de/DE/Service/EXDAT/Datensaetze/mobilitaetsindikatoren-mobilfunkdaten.html>

Beispiel 2: Analyse zur Optimierung des ÖPNV

Mobilfunkanbieter nutzen erhobene Mobilfunkdaten (v.a. auch Standortdaten) für verschiedene Projekte im Kontext der Verkehrsoptimierung oder auch Optimierung des öffentlichen Personennahverkehrs. Telefónica berichtet dazu etwa von mehreren Projekten zum ÖPNV in Leipzig in München, der Verkehrsplanung in Deutschland oder zur Auswertung von Staus auf deutschen Straßen.

<https://www.telefonica.de/analytics/anonymisierte-daten.html>

Ähnlich berichtet auch der Mobilfunkanbieter Deutsche Telekom von einem Projekt mit Verkehrsbetrieben in Nürnberg, wo anonymisierte Standortdaten verwendet werden, um das Verkehrsangebot zu optimieren. Unter dem zweiten Link findet sich auch ein Video, welches dazu die sekundäre Verwendung der Standortdaten beschreibt und eventuell für den Unterricht genutzt werden könnte.

<https://www.telekom.com/de/medien/medieninformationen/detail/data-analytics-handy-schwarm-hilft-strassenbahn-349426>

<https://www.telekom.com/de/medien/mediencenter/medienmappen/medienmappen-2015/data-analytics>

3 Glossar relevanter Begriffe:

Datenbegriff:

Wie im Kernlehrplan NRW für Informatik in Klasse 5 und 6 dokumentiert (s. Inhaltsfeld „Information und Daten“) kann zwischen Information und Daten unterschieden werden: Daten sind (digital) repräsentierte Informationen und können etwa in Informatiksystemen gespeichert und verarbeitet werden. Hinsichtlich des Konzepts Datenbewusstsein ist besonders hervorzuheben, dass Kontexte, Phänomene oder etwa auch Interessen, Emotionen oder Handlungen einer Person anhand ausgewählter Merkmale modelliert werden. Gerade die persönlichen Daten, welche für Datenbewusstsein eine große Rolle spielen, entstammen einem Kontext, wo die jeweilige Person involviert ist oder war. Bei der Thematisierung von Daten sollte der Modellierungsaspekt nicht vernachlässigt werden, da die Kontexte, Phänomene oder Personen nicht vollständig repräsentiert werden. Im Kontext des Datenbewusstseins bedeutet das gerade, dass die erhobenen, persönlichen Daten kein Abbild einer Person darstellen, sondern lediglich modellhaft anhand ausgewählter Merkmale repräsentiert. Dadurch kann auch eine verzerrte Repräsentation einer Person entstehen. Außerdem sollte beachtet werden, dass gewisse Informationen, wie etwa Emotionen oder Interesse, im Sinne der Merkmale für die erhobenen Daten operationalisiert werden, was oft nicht trivial ist (Was heißt es, wenn ein datengetriebenes digitales Artefakt das „Interesse“ der Nutzenden speichert? – Was ist das? Woran wird das fest gemacht?)

Digitale Artefakte und datengetriebene digitale Artefakte:

Im Konzept Datenbewusstsein wurde der Begriff der datengetriebenen digitalen Artefakte (ddA) eingeführt. Dieser beschreibt eine spezielle Art von digitalen Artefakten. Digitale Artefakte sind ein Sammelbegriff für digitale Werkzeuge, Computersysteme aller Art, ihre Bestandteile, ihre Verbindung untereinander. Sie umfassen Sowohl Hardware, Software, Daten und Objekte sowie Algorithmen und Datenstrukturen. Datengetriebene digitale Artefakte sind dann digitale Artefakte, die sich selbst oder ihre Rückmeldung in der Interaktion mit diesem durch die Verarbeitung erhobener Daten verändert. Diese nutzen dann oft zum Beispiel auch Methoden des Maschinellen Lernens.

Architektur und Relevanz (Duale Natur digitaler Artefakte):

Die duale Natur digitaler Artefakte oder auch Dualität beschreibt eine analytische Trennung von Aspekten eines digitalen Artefakts (s.o.). Ein digitales Artefakt kann dieser Auffassung nach aus der Perspektive auf die Architektur und auf die Relevanz beschrieben werdenⁱⁱ. Architektur meint alle technologischen Strukturen und Mechanismen. Relevanz hingegen meint Intentionen, Funktionen, Meinungen, Interpretationen und der Kontext eines digitalen Artefakts.

Explizit und implizit erhobene Daten:

Im Konzept Datenbewusstsein wurden die Begrifflichkeiten der explizit und implizit erhobenen Daten eingeführt. Diese stehen in der Regel in der Verbindung zum Nutzenden - oft stellen sie personenbezogene Daten dar. Die explizit erhobenen Daten sind jene, die der Nutzende mit seiner Handlung intendiert zu erstellen, also direkt und aktiv eingegeben bzw. erzeugt hat. Darüber sind sich Nutzende in der Regel bewusst. Dies sind zum Beispiel bei Social Media Plattformen gepostete Texte und Bilder, bei einer Suchmaschine etwa der Suchbegriff oder beim Telefonieren über das Mobilfunknetz die Telefonnummer desjenigen, den man anrufen möchte. Im Gegensatz dazu, werden implizit erhobene Daten indirekt durch Beobachtung (Tracking) oder Verarbeitung bereits erhobener Daten nebenher zur eigentlichen Handlung des Nutzenden erhoben und generiert. Dieser Datenerhebung sind sich Nutzende oft nicht bewusst. Im Beispiel der Social Media Plattform sind dies etwa Likes und Klicks, bei der Suchmaschine etwa Klicks auf Suchergebnisse oder beim Telefonieren über das Mobilfunknetz etwa Standortdaten der verbundenen Basisstationen.

Primäre und sekundäre Zwecke der Verarbeitung und Verwendung:

Im Konzept Datenbewusstsein wurden die Begrifflichkeiten der primären und sekundären Zwecke der Verarbeitung und Verwendung erhobener Daten eingeführt. Diese beziehen sich auf die Verarbeitung und Verwendung von Daten über einen Nutzenden, die bei der Nutzung von datengetriebenen digitalen Artefakten erhoben werden. Primäre und sekundäre Zwecke beziehen sich auf die Intention, mit der diese zuvor erhobenen Daten verarbeitet und verwendet werden. Primäre Zwecke umfasst, dass die erhobenen Daten dazu verarbeitet und verwendet werden, um das datengetriebenen digitalen Artefakten mit den Features anbieten zu können. Diese beziehen sich auf einer Nutzerperspektive auf die Verarbeitung und Verwendung: Die Daten werden verarbeitet und verwendet, um Nutzenden Features anbieten zu können. Im Beispiel der Suchmaschine ist dies etwa das Anzeigen von Suchergebnissen. Auch inbegriffen wäre, wenn die Suchergebnisse personalisiert geordnet werden. Im Sinne des Features für den Nutzenden würde dies bedeuten, dass der Nutzende gerade die Ergebnisse angezeigt bekommt, die für ihn idealerweise relevant sind. Sekundäre Zwecke bedeutet, dass die Daten verarbeitet und verwendet werden, um andere/weitere Zwecke zu verfolgen – z.B. weitere wirtschaftliche oder wissenschaftliche Zwecke. Diese „Zweitverwertung“ der Daten bezieht sich auf einer Anbieterperspektive auf die Verarbeitung und Verwendung der erhobenen Daten: Wozu kann ein Anbieter eines datengetriebenen digitalen Artefakts die erhobenen Daten noch nutzen? Im Kontext von Streamingdiensten (z.B. Spotify) könnte dies etwa umfassen, dass Nutzungsdaten (z.B. gehörte Musik) zur Analyse der Emotionen der Nutzenden verwendet werden.

Data Moves:

Mit Data Moves werden Datenoperationen beschrieben. Diese umfassen etwa folgendeⁱⁱⁱ:

- Filtern: Bilden einer Teilmenge der Daten
- Gruppieren: Daten in Teilgruppen unterteilen
- Zusammenfassen: Aggregieren von Daten
- Berechnen: Neue Attribute ausgehend von existierenden Daten erstellen (z.B. Ausgehend von zwei Spalten eine dritte Spalte erzeugen)
- Merging/Joining: Datensätze zusammenführen
- Reorganisieren: zum Beispiel ändern der Darstellung der Daten

ⁱ Höper, L. & Schulte, C., (2021). Datenbewusstsein im Kontext digitaler Kompetenzen für einen selbstbestimmten Umgang mit datengetriebenen digitalen Artefakten. In: Gesellschaft für Informatik (Hrsg.), INFORMATIK 2021. Gesellschaft für Informatik, Bonn. (S. 1623-1632).

ⁱⁱ Schulte, C., & Budde, L. (2018). A Framework for Computing Education: Hybrid Interaction System: The need for a bigger picture in computing education. In *Proceedings of the 18th Koli Calling International Conference on Computing Education Research* (pp. 1-10).

ⁱⁱⁱ Erickson, T., Finzer, B., Reichsman, F., & Wilkerson, M. (2018). Data Moves: one key to data science at school level. In *Proceedings of the International Conference on Teaching Statistics (ICOTS-10)* (Vol. 6).