

Data Awareness

Where, how and for what purpose is data collected and processed? - Data awareness through the exploration of location data from the mobile network

Information about the teaching unit:

Title:	Where, how and for what purpose is data collected and processed? - Data awareness through the exploration of location data from the mobile network
Target Group:	Computer science in grades 5 and 6 (all school types)
Time scope:	4 to 6 lessons of 45 minutes each

1 Overview of the teaching unit:

This teaching unit promotes data awareness: The competence to become aware of the role of collected and processed data when using different applications in order to ultimately be able to evaluate their use and identify their own options for action. The aim is to enable learners to interact with data-driven applications in their everyday lives in a self-determined way. The unit consists of three parts and focuses on the collection and processing of location data as an example application when using the mobile network and further in other everyday contexts of the students. In **the first part**, the mobile network is introduced as a context and the structure and functionality of this is developed using the example of making calls with a mobile phone. It is also explained which personal data is collected and what it is primarily used for. For example, this is the location of a base station in the mobile network to which users are connected. This location data is necessary, for example, to ensure the efficient establishment of a mobile connection (primary purpose of the location data collected). In **the second part**, given data from the mobile network is used. This real data of a person contains, among other things, location data (for a more comprehensive explanation of the data, see section **Fehler! Verweisquelle konnte nicht gefunden werden.**). The students analyse this location data using a provided web application with regard to the question of what information can be obtained about a person from location data. They characterise the person they do not know by creating a profile. The reasons why such profiling is prohibited under German law are explained - with strict exceptions requiring authorisation for special purposes (e.g. law enforcement purposes). Students understand these reasons better because they have learnt by example what could be inferred from such data. In **the third part**, the experiences made are transferred to other possible contexts and thus generalised by the students examining other data-driven applications from their everyday lives in which location data is collected, such as in certain apps on their mobile phones (including on the basis of GPS data). As part of an evaluation and assessment of data collection and processing in the various examples, the advantages and disadvantages of collecting and processing personal location data can be discussed in order to provide students with a basis for making informed decisions regarding their interaction with these data-driven applications.

Table of content:

1	Overview of the teaching unit:	1
2	Was meint ‚Datenbewusstsein‘?	2
3	Didaktische Kernidee: Förderung von Datenbewusstsein in diesem Unterrichtsmodul	2
4	Lernvoraussetzungen:	3
5	Ziele:	3
6	Zentrale Leitfragen:	3
7	Zusammenfassender Überblick über das Unterrichtsmodul	Fehler! Textmarke nicht definiert.
8	Überblick über den Unterrichtsverlauf:	4
9	Beschreibungen ausgewählter Materialien und weiterführende Hintergrundinformationen:	10
10	Glossar relevanter Begriffe:	15

2 What does Data Awareness mean?

The data awareness framework for computing education aims to promote awareness and understanding of the collection, processing and use of personal data during the use of data-driven digital artefacts (see glossary in section 10) among students. The collection of personal data during such an interaction occurs through a) active input of information by the user, b) observation and tracking of behaviour and c) processing of previously collected data. This can be divided into explicit data collection, i.e. the active and direct provision of information by the user intended by the action, and implicit data collection, i.e. processes that take place alongside the actual action, such as observation, tracking and generation through data processing (for a detailed description of the terminology, see section 10). Users of data-driven digital artefacts are often aware of the explicitly collected data, but are often unaware of the implicitly collected data. The personal data collected in this way as part of an interaction with a data-driven digital artefact can be processed using various methods, from simple data moves (see glossary in section 10) to machine learning methods. Various purposes are pursued with the processing and utilisation of the data. Collected data can be processed and used for the (‘technical’) operation of functions of the data-driven digital artefact (primary purposes) and/or for additional purposes or, for example, to investigate further developments of the data-driven digital artefact (secondary purposes) (detailed description of the terms in section 10). Primary purposes are to be understood from the perspective of the user and secondary purposes from the perspective of the provider (what else can a provider do with the data?). In terms of the different purposes, a digital doppelgänger (user profile) can be constructed as a model representation of the user. The data awareness framework should raise awareness of data-driven digital artefacts as well as an understanding of the processes of explicit and implicit collection and generation of data and automated data processing for primary or secondary purposes. Students should therefore be enabled to recognise and analyse the collection and processing of personal data in an interaction with a data-driven digital artefact and to make self-determined decisions for interactions of this kind.

3 Educational core idea of fostering data awareness in this teaching unit

The facets of data awareness are implemented in the three parts of the teaching module in order to realise the objectives and thus promote data awareness among learners. The example chosen in the first part describes an interaction system consisting of the user and the mobile phone network as a data-driven digital artefact and the interaction between them. By exploring the structure and functionality of the mobile network in the first part, students can investigate the explicit and implicit collection of personal data in this example using an interactive simulation with a puzzle. The primary purpose of the location data collection is then analysed as an example. The

implicitly collected location data is necessarily processed to establish a mobile phone connection (e.g. when making a phone call) (primary purpose). In the second part, the students explore given location data from the mobile network for an imagined secondary purpose by characterising a person based on their location data (construction of a digital doppelgänger of a user). To do this, they explore the location data in a given web application and create a profile. In this way, the question of what information about a person can be obtained from location data is explored. This exemplary data analysis is highly regulated for a German mobile phone provider, which can then be taken up and reflected upon. Other examples of secondary uses of the data collected in the mobile phone network can also be addressed (see section 9.6). In the third part, the knowledge learnt about data awareness is applied to further examples from their own everyday lives: Interaction with a data-driven digital artefact; explicit and implicit data collection; primary and secondary purposes of use and processing; and construction of a digital doppelgänger. These contexts are then reflected upon and evaluated based on criteria.

4 Learning requirements:

This teaching unit does not require students to have any particular prior knowledge. However, they should have basic experience of using computers and web applications. In addition, a basic understanding of the concept of data (in particular, data vs. information) is desirable, but corresponding explanations could also be integrated into this unit.

5 Goals:

The three parts of this teaching unit essentially focus on the following objectives:

- **Part 1: Structure and functionality of the mobile network**
 - Students recognise the basic structure and functionality of a mobile network by simulating and explaining telephone calls in a mobile network using a puzzle as an example.
 - Students differentiate between the terms explicitly and implicitly collected data and recognise which data is collected explicitly and implicitly when using the mobile network.
 - Students justify the necessity of processing this data (primary purpose) using the implicit collection of location data as an example.
- **Part 2: Exploration of location data**
 - Students explain the procedure for exploring location data using the given location data and can describe essential steps.
 - Students find out personal information about a person they do not know by exploring given location data with a web application (secondary purpose).
- **Part 3: Other contexts with the collection and processing of location data**
 - Students apply their knowledge of data awareness to another example of a data-driven digital artefact from their everyday life by using this example to identify and describe explicit and implicit data collection, its use and processing for primary and exemplary secondary purposes as well as the construction of digital doppelgängers.
 - Students make a reasoned assessment of the collection and processing of location data in the examples discussed, for example by discussing the trade-off between data-saving behaviour and the use of individual or social benefits.

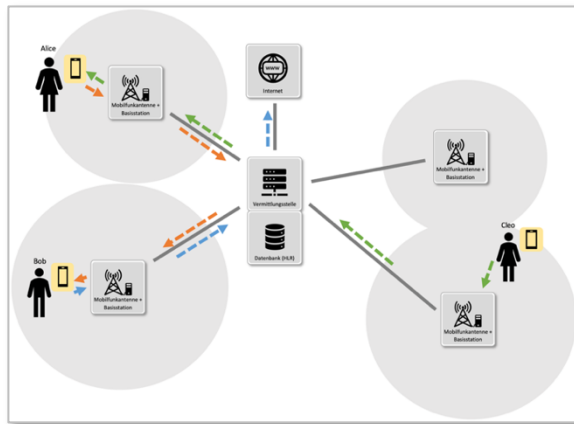
6 Key questions:

- **Part 1: Structure and functionality of the mobile network**
 - How does the mobile network work and what data is collected explicitly and implicitly during use (e.g. when making phone calls)?
 - Why is the implicit collection and processing of location data necessary? (primary purpose)
- **Part 2: Exploration of location data**
 - What knowledge can be gained about a person by analysing their location data? (secondary purpose)
- **Part 3: Other contexts with the collection and processing of location data**
 - In what other contexts is location data collected and what is it used for there?

7 Overview of the lesson plan:

The following table describes the course of the lesson. In the 'Content' column, you will find descriptions of the course of the lesson with relevant didactic comments, which generally focus on promoting data awareness. The corresponding materials are listed in the table for each phase of the teaching module.

Phase	Content	Goals	Materials
Part 1: Structure and functionality of the mobile network			
1a	<p>Introduction to the interaction context and problematisation:</p> <p>As a class discussion, a (thought) experiment on establishing a mobile phone connection when making a call is carried out together: (1) A second mobile phone is called from one mobile phone - what happens? (2) What happens when a mobile phone is placed in a metal box? (3) What happens when both mobile phones are placed together in a metal box? The question of how the mobile phone network is set up and which steps are necessary to establish a mobile phone connection is worked out together.</p> <p><u>Didactic comment:</u></p> <p>The students' own experiences of using the mobile network are taken up (from the user's perspective). In addition, curiosity is to be aroused about the structure and exact functionality of the mobile network (from the provider's perspective). As a result, the question of the structure and functionality of the mobile network is posed. The (thought) experiment helps to raise the question of the 'steps or stations between the two mobile phones'.</p>	<ul style="list-style-type: none"> • Activation of prior knowledge of the mobile phone network (usually from the user's perspective) • Arousing curiosity about the structure and functioning of the mobile phone network 	Supporting slides for the thought experiment (alternatively two mobile phones and a metal box)
1b	<p>Development of the structure and functionality of the mobile network:</p> <p>Together or in individual work, an overview video on the structure and functioning of the mobile phone network is shown (several runs if necessary). The learners note down information on the components of the mobile phone network (worksheet 1), essentially these are: the mobile devices, base stations with mobile phone antennas and corresponding radio cells, switching centres with databases (called data storage here). The terminology of the components of the mobile radio network is then discussed in the plenary session and clarified. The simplified structure of the mobile phone network is then reconstructed using a puzzle, the mode of operation is worked out using two scenarios and initial ideas regarding data collection are developed (worksheet 2).</p>	<ul style="list-style-type: none"> • Understanding the architecture and relevance (see glossary in section 10) of the mobile network • If necessary, familiarisation with the term data using the data collected here as an example; otherwise application of knowledge of the term data (see glossary in section 10) • Introduction of the terms 'explicitly and implicitly collected data' and identification of explicitly and implicitly collected data in this example • Describe the primary purpose of the processing and use of the location data 	Explanation Video (german) , Worksheet 1, Puzzle, Worksheet 2, Worksheet 2 addition (optional), Worksheet 3



Using the two scenarios, the students work out which (personal) data can and must be collected when using the mobile network (using the example of making phone calls) (worksheet 2). These ideas for the data collected are collected in a class discussion. The concept of data is discussed in terms of the extent to which information is understood as data in this example (e.g. the location as a numerical pair of longitude and latitude). The concepts of explicit and implicit data collection are then introduced (worksheet 3). The students categorise the data collected when using the mobile network into these two types (task 1 on worksheet 3). This application of the terminology is then confirmed. The learners then evaluate the use of location data for the primary purpose of establishing a mobile phone connection and develop ideas for secondary purposes (worksheet 3). The concepts of primary purposes (task 2) and secondary purposes (task 3) are introduced in the safeguarding of the processing of worksheet 3 (see explanation of terms in section 10).

Didactic Comment:

The architecture and relevance (see glossary in section 10) of the mobile network is developed in this phase. The given interaction context is then analysed with regard to the explicitly and implicitly collected data (additional ideas for collected data can be evaluated with regard to necessity). Students also justify the processing of location data for the primary purpose of establishing a mobile phone connection. Secondary purposes are initially addressed more superficially in this part, so that only initial ideas for possible secondary purposes are developed. In this part, learners familiarise themselves with the terminology relating to the types of data collection (explicit and implicit data collection) and the types of purposes for using and processing the data (primary and secondary purposes). These will be taken up again in the next phases and serve to explain the role of data in such an interaction system. In the second part, a rather fictitious example of a secondary purpose is discussed in more detail, so that this aspect can serve as a transition to the second part.

Part 2: Exploration of location data

2a	<p>Preparation of the exploration of location data using the interactive web application:</p> <p>Firstly, Malte Spitz's mobile phone data is introduced (see section 9.2), which includes timestamps, services used and location data. No indication should be given as to which person this is. It should only be noted that this data was collected by one person in six months. This introduction can be done using the data table (attached pdf), which should then also be discussed (reference to tasks 4 and 5 on worksheet 2 is possible).</p> <p>The following key question is then addressed in plenary: What could you find out about a person if you had this data?</p> <p>Before the given mobile phone data (especially location data) is explored using a web application, the learners should make assumptions about what could be found out with this data.</p> <p>This is followed by an introduction to the interactive web application in which this location data can be explored (see section 9.3 for a description). The students could first try out the application themselves, for example, before exploring the data in plenary using a common example question. The joint example should be used to convey the procedure for exploration, which the students carry out themselves in the next phase: (1) Develop a question (e.g. Where does the person probably live?), (2) Decide on filters to be set in the web application (e.g. period from 3 to 4 o'clock, period from 3 to 4 o'clock, period from 3 to 4 o'clock, period from 3 to 4 o'clock). (3) visualise and examine the location data on the map, (4) interpret the data and answer the question (e.g. the person lives in Berlin near Zehdenicker Straße) (for a more detailed description of examples, see section 9.4). The procedure for a question can also include 'trying out' different filters.</p> <p><u>Didactic Comment:</u></p> <p>The available mobile phone data can be used as an example to raise awareness of how much data is collected and generated when using the mobile network.</p> <p>With regard to this data, care should always be taken to ensure that it represents the locations of the base stations and not GPS data or similar and therefore has a certain degree of inaccuracy (see description in section 9.5). Exploring the location data to create a model characterisation of the person (cf. the idea of the digital doppelgänger) is a secondary purpose of using and processing the location data. This purpose should not be presented as an actual use and processing, but as a possible one if this data is shared, which is discussed in more detail in phase 2c. In this phase, the exploration of the data is only introduced first, after which the learners explore the data independently with the web application in the next phase.</p>	<ul style="list-style-type: none"> • Improve and demonstrate understanding of the data collected in this interaction context and the amount of data for a person • Practising dealing with the concept of data and the representation of data using the data table and map as examples • Familiarise themselves with the procedure for data exploration using a web application as an example for location data • Understand the characterisation of the person as a secondary purpose of the use and processing of location data (fictitious purpose) 	<p>Table with the location data (pdf), Web-Application (s. section Fehler! Verweisquelle konnte nicht gefunden werden.)</p>
----	---	--	---

2b	<p>Exploration of location data using the interactive web application:</p> <p>In groups, students explore the location data independently using the web application. As a group, they create a joint digital double of the person in the form of a profile. This characterises the person unknown to them, for example with interpretations of their place of residence, place of work or leisure activities (worksheet 4). To this end, the learners can work in groups on different questions in order to be able to make a more comprehensive characterisation together. During the exploration phase in the groups, students can optionally be given the opportunity to draw on further contextual information, such as what is in a particular place or what a particular company does - the students would then explore the context.</p> <p>In the subsequent evaluation of the group results, the various interpretations are discussed and summarised and the students are asked to justify their interpretations. It can also be discussed how confident the students are with a particular interpretation. In addition, the subjectivity of the interpretations can be discussed together by taking up the contextual knowledge with which the students arrived at a particular interpretation - this is particularly useful when dealing with different interpretations. The concept of the digital doppelgänger (introduced on worksheet 4) should be discussed in the backup.</p> <p><u>Didactic comment:</u></p> <p>By filtering as a data move (see glossary in section 10) during exploration, some information about the person can already be found. The learners' interpretations of the characterisation are based on individual contextual knowledge and are therefore subjective. This leads to the emergence of conflicting interpretations, which can and must be discussed profitably. It should also be noted that the characterisations or profiles have a model-like character, which is determined by the selected characteristics in the available data and does not represent an all-encompassing image of the person.</p>	<ul style="list-style-type: none"> • Understanding the subjective nature of information through interpretation (as opposed to data) • Recognise the extent of a person's characterisation based on the exploration of location data • - Get to know the idea of the digital doppelgänger as a model characterisation of a person • Practising the procedure for exploring the data 	Web-Application, worksheet 4
2c	<p>Reflection and evaluation:</p> <p>The collection and processing of personal data when using the mobile network is now reflected on in the classroom discussion and then evaluated by the students. At this point, it should be pointed out that a German mobile network provider is not allowed to use the data collected in this way to characterise people without further action. With the help of the experiences from the exploration, this regulation can also be exemplified by the students.</p> <p><u>Didactic comment:</u></p> <p>On the one hand, care should be taken to ensure that the misconception is not developed that mobile network providers actually analyse an individual's location data in this way. On the other hand, it should be made clear that the location data is necessarily</p>	<ul style="list-style-type: none"> • Reflect on the findings on the collection and processing of data (especially location data) using the example of the mobile phone network and evaluate its 	

	collected here (e.g. for the primary purpose) and that regulation can also be useful.		
Part 3: Other contexts with the collection and processing of location data			
3a	<p>Identify further contexts in the students' everyday lives: Further interaction contexts from the students' everyday lives in which location data is also collected and processed are collected in the plenary session. It should be noted that there is also other location data (mobile network: base station locations; other contexts: e.g. GPS data; see section 9.5). In the next phase, selected examples will be analysed in more detail.</p> <p><u>Didactic comment:</u> This phase is not concerned with the precise reconstruction of various other contexts with regard to data collection and processing; this will take place in the subsequent phase. The initial aim is to identify and describe other contexts in which location data is collected from users. This should enable learners to develop a broader and more differentiated view of the role of location data in their own everyday lives and to recognise data-driven digital artefacts that collect location data.</p>	<ul style="list-style-type: none"> Recognising further interaction contexts from your own everyday life in which location data is collected and processed 	
3b	<p>Application and evaluation: In this phase, students apply the perspectives they have learned about the facets of data awareness to other examples from their everyday lives in which location data plays a role (i.e.: explicit and implicit data collection, primary and secondary purposes, digital doppelgänger). We propose three variants for this, which can be chosen depending on the learning group.</p> <p><u>Variant 1: "Exploration protocol" (relatively open)</u> The students select a data-driven digital artifact in which location data is collected and processed by the users and examine the role of the data with regard to the facets of data awareness. To do this, they create a kind of protocol on the worksheet 5a as an overview of this selected interaction context. The students then present their examples to each other. The students then make a reasoned assessment of these examples with regard to the collection and use of data. Based on the explicit and implicit data collection (or knowledge about possible settings), ideas for possible actions can also be collected (e.g. restricting the collection of data or certain purposes for the use of data, changing one's own behavior with regard to the use of features, ...).</p> <p><u>Variant 2: Group puzzle on further contexts with given information</u> In expert groups, students develop a context for interacting with a data-driven digital artifact that collects location data. Three groups are prepared for this: One group on Google Maps, one on Snapchat and one on WhatsApp. As a group, the students first work on the respective worksheet and deal with the role of</p>	<ul style="list-style-type: none"> Change of perspective from a user to a provider of a data-driven digital artifact Applying the perspectives on the facets of data awareness to further examples from your own everyday life Evaluation of the collection and processing of location data in the mobile network and other contexts Reflecting on the collection and processing of location data in your own everyday life Reflecting on one's own ability to act in the context of interactions with data-driven digital artifacts 	<p><u>Variant 1:</u> Worksheet 5a</p> <p><u>Variant 2:</u> Worksheet 5b</p>

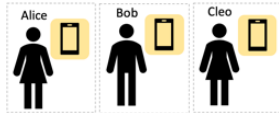
<p>data in the respective context. The composition of the groups is then changed in the sense of a group puzzle so that at least one learner from each expert group is represented in a group. In these mixed groups, the learners work on worksheet AB5b with the note Group jigsaw discussion. This instructs the learners to evaluate their results from the expert groups. They also discuss possible courses of action in the group based on the previous results of the groups.</p> <p>This evaluation and elaboration on the key question is discussed and evaluated in the plenary session. As a safeguard, learners should be encouraged to formulate a personal conclusion on their interaction with these exemplary data-driven digital artefacts or develop an attitude towards them.</p> <p><u>Didactic comment:</u></p> <p>In this application and assessment phase, students should develop a broader view of the collection and processing of location data in their everyday lives. They should also be able to describe the collection of location data in different contexts and different purposes of processing and use.</p> <p>This should be supported by the change of perspective from users to providers of data-driven digital artifacts. In addition, based on this perspective and the facets of data awareness, they develop ideas for possible courses of action in the context of interaction with a data-driven digital artifact. This should lead to students perceiving themselves as capable of acting and not developing a resigned attitude. In both variants, learners should make a personal (reasoned) assessment or attitude towards data-driven digital artefacts in the various interaction contexts and position themselves accordingly.</p>		
---	--	--

8 Descriptions of selected materials and further background information:

8.1 Notes on the puzzle about the structure and functionality of the mobile network

Puzzle pieces:

People:



The three person tiles symbolically represent the three people with their cell phones or mobile stations. These play a major role in the simulations of how the mobile network works in the second task on worksheet 2.

Elements of the mobile network:



In the puzzle, there are four puzzle pieces representing the reduced mobile network in addition to the people tiles: 1) antenna + base station (four times), 2) switching center, 3) data storage, 4) Internet.

The meaning of each element is described in more detail in section 9.5. A solution to the puzzle is described below. As shown there, the puzzle pieces are to be placed on the puzzle template (see below) in order to adequately represent the mobile network and simulate how it works. The interface between the exchange and the database has been replaced by a directly connected arrangement to reduce complexity. The exchange and the database are therefore to be understood as two linked elements. The Internet puzzle piece (4) represents the connection between the switching center and “the Internet” in an extremely abstract way. Discussing the structure of the Internet in this teaching module, albeit a basic one, would go too far in terms of complexity. (Depending on the learning group, however, a very short excursus on the structure of the Internet would be conceivable at this point).

Puzzle template:

In the first task of worksheet 2, the puzzle pieces and people tiles are placed on a puzzle template to reconstruct the structure of the mobile phone network. The puzzle template is handed out to the learners together with the puzzle pieces and worksheet 2. Trials of the teaching module have shown that it is advisable to bring the puzzle materials to class already cut out beforehand so as not to significantly reduce the effective learning time; it may also be possible for pupils to cut them out in class.

Solution to the puzzle:

A possible solution to the puzzle is shown in Figure 1. Variations exist only in the choice of radio cells in which the three people are located. The elements of the mobile network (gray puzzle pieces) should not be chosen differently. In the second task on worksheet 2, the arrows are placed on the puzzle to simulate establishing a connection in the mobile phone network. Two scenarios for making a phone call (situation 1: orange arrows, situation 2: green arrows) are used for this purpose. Optionally, a third scenario for accessing a website (blue arrows) can be worked on as a “sprinter task”. The direction of the arrow simplifies the sending of data packets when the respective mobile connection is requested.

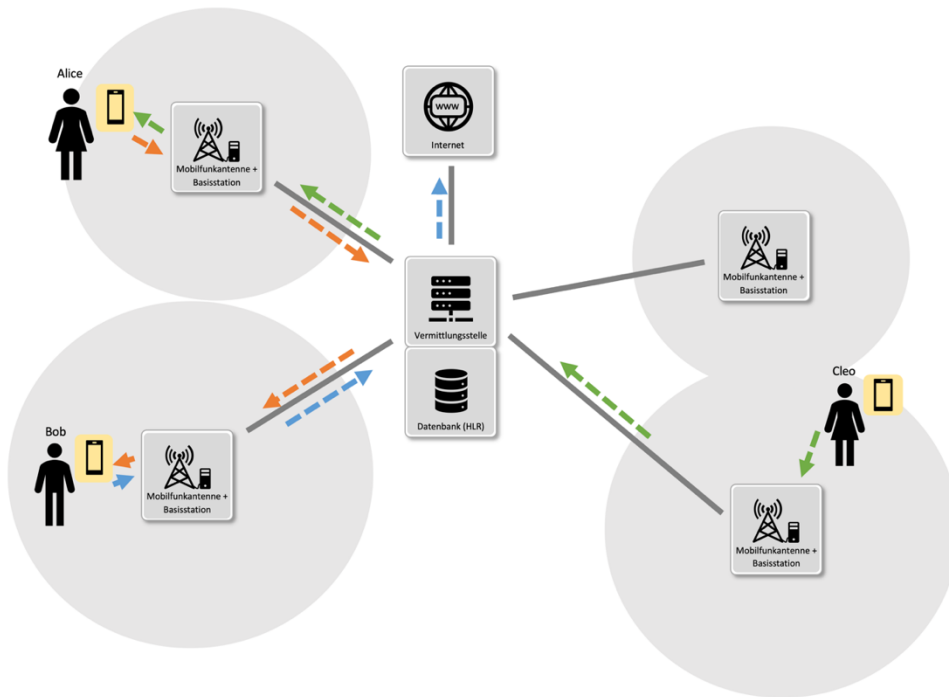


Figure 1: Example of a solution to the puzzle (worksheet 2)

8.2 Notes on the data explored by Malte Spitz

The location data was collected by a mobile network provider during use of the mobile network. The data records were filtered for reasons of data protection so that not all of the data collected is included in the table and the table should therefore only be described as an excerpt of the data collected. The data belongs to a single person - more precisely to the politician Malte Spitz. It was published jointly with Zeit Online in the context of the debate on data retention. The data included covers a period between September 2009 and February 2010. When interpreting the location data visualized on the map, it should be borne in mind that this is now older data on a current map and therefore the other references on the map (e.g. café, restaurant, companies) are not necessarily always correct. In our view, however, this does not hinder the knowledge gained by the learners.

Further sources for the data:

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

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

8.3 Notes on the interactive web application

We have developed an interactive web application for the teaching module. This can be accessed via the following link: <http://go.upb.de/ExplorationLocationData>

It is based on a Jupyter notebook and was developed using Python. However, learners are not required to have any knowledge of Jupyter notebooks or any programming skills.

First, the location data must be read in, which can be done using the Load location data button. In principle, it is also possible to load your own location data or to extend the application with this data, which can be selected at this point.

After loading the location data, the following user interface is displayed:

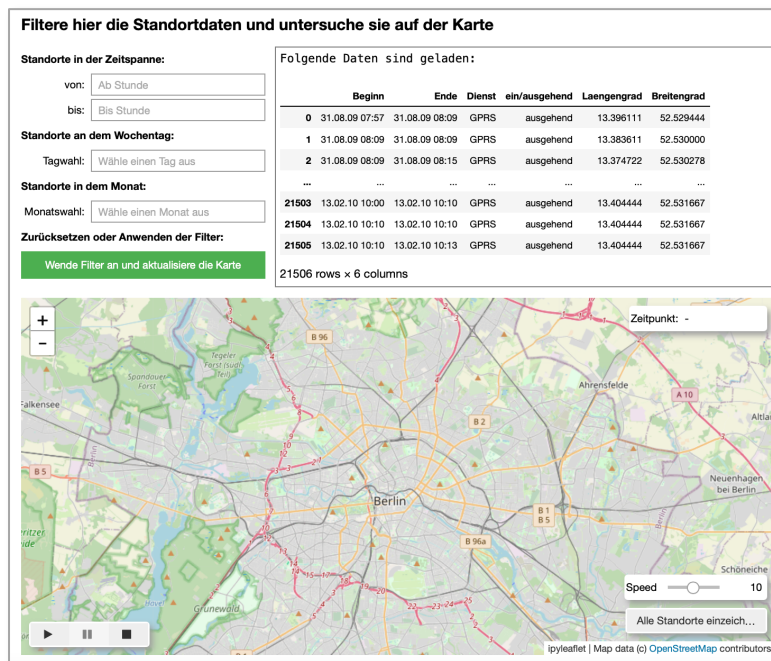


Figure 2: Interactive Web-Application

It essentially consists of three areas: 1) Input screen for setting filters (top left), 2) Display of the processed filters and the table of current location data (top right) and 3) Map for visualizing the location data with several widgets for controlling the visualization (bottom).

Three filters can be set separately or together in the input screen (1). Firstly, you can filter according to a time period, in which case all location data outside this time period is removed. Only whole hours can be entered, such as 16 for 16:00. Secondly, you can filter by a day of the week, such as "Monday", which can be entered as a text form in the text field or selected from the drop-down menu. All location data for days other than the selected weekday will be removed. Thirdly, you can filter by a month, which can also be entered as text or selected from the drop-down menu. All location data from other months will be removed. One or more filters used are applied to the current location data by clicking on the Apply filter and update map button. The map is then reset and can be used to visualize the current location data. Clicking the Reset all filters button restores the original location data so that new filters can be set.

In the display (2), the most recently applied filters are each described with a sentence and the current location data is shown in tabular form.

The location data is visualized on the map (3). This can be displayed dynamically using the play widget at the bottom left. This makes it easy to follow the progress of the location data. This plotting of the locations can be paused with the Pause button in this widget and reset with the Stop button. The speed of plotting can be adjusted using the Speed slider at the bottom right. The Draw all locations button can be used as a replacement for the dynamic visualization of the location data in order to draw all locations at once. For performance reasons, however, this is recommended and regulated for a small amount of location data. At the top right of the map, the time of the last location plotted is shown in a text field. The zoom level of the map display can be changed using the plus and minus buttons at the top left of the map.

Locations are displayed as markers. By holding the mouse over a marker, a tooltip with the time of this location is displayed. If several markers are close to each other, they are combined into a cluster (colored circle with a number). By clicking on the circle or zooming into the map, the clusters are separated again (depending on the zoom level).

Please feel free to contact us if you have any questions, have discovered errors or have ideas for improvements. We are always interested in further developing the web application.

8.4 Example of interpretations of the visualized location data

In the exploration phases of the teaching module with the interactive web application, the learners filter and interpret the visualized location data. As the interpretations are subjective, they will certainly contradict each other in the learning group. Typically, a wide variety of plausible interpretations are offered for the profession. There does not have to be a right or wrong interpretation in the course of the lesson. It is important to be able to argue for your own interpretations with reference to the data.

Example of an interpretation:

Most people sleep at night. In Germany, this would mean that many people sleep between 3:00 a.m. and 4:00 a.m., for example. If you now set the filter for the time period to 3 a.m. to 4 a.m., a total of 294 locations are available. After marking them on the map, 208 locations in Berlin can be localized. If we now take a closer look at these, 189 of them are located near the Rosenthaler Platz subway station (see Figure 3).

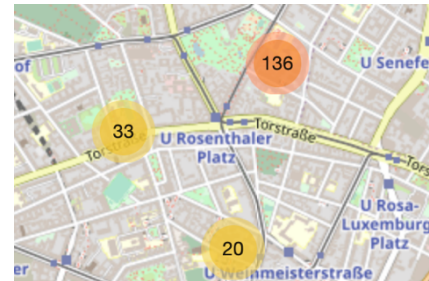


Figure 3: Location data filtered with 3 to 4 am.

Another example of an interpretation:

Students can tend to interpret very strongly. Take the following example from a trial: I filtered the location data and found out that he often stayed at a certain company. The company is active in a certain industry, which I found out using a search engine. So that's why he now works in the financial sector, for example.

One pupil replied to this in a teaching test that the location of a company during the day would only allow the interpretation that he might work there, but not what profession he pursues there. He could also work as a janitor, for example.

Discussions of this kind show a good understanding of how information is obtained through the interpretation of (location) data and can be very beneficial for the course of the lesson. The reliability of the interpretations should therefore be taken into account but not underestimated. The evaluation phase of these explorations is short. It can certainly be extended in favor of discussions of this kind. It can then also be taken up that further data sources are also included in such data evaluations (see optional addition in phase 2b) in order to enable more precise interpretations.

8.5 Further information on the mobile network

Simplified structure of the mobile network

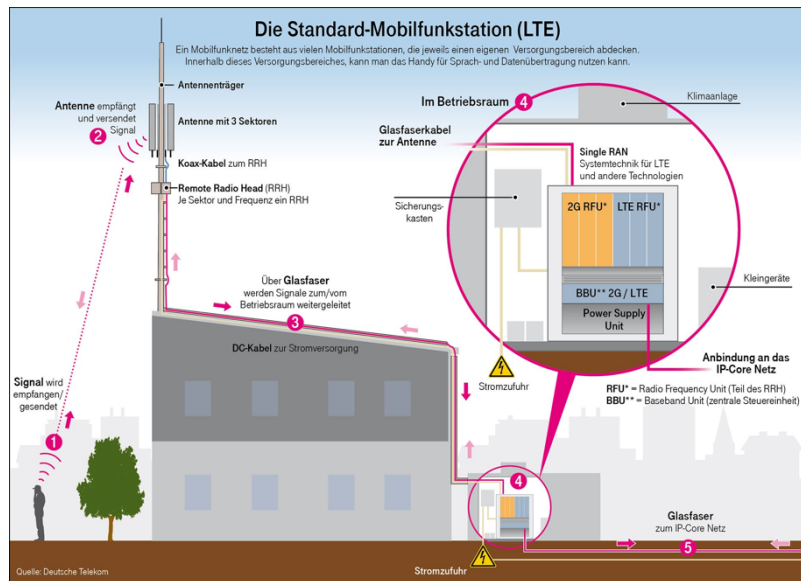
The mobile network is a data-driven digital artifact. It is operated by a network operator, of which Telekom Deutschland, Vodafone and Telefónica Deutschland are active in Germany. Since the first mobile network was set up in the 1980s, the architecture and relevance of the mobile network as well as the patterns of action of the participants have evolved astonishingly. At the end of 2020, the three network operators mentioned above had a total of over 150 million subscribers. For the teaching module, the structure and functioning of the mobile network is reconstructed, omitting details that only play a subordinate role in the course of the lesson. The mobile network consists of the following components

Mobile station or also end device or cell phone:

The mobile station consists of a cell phone and a SIM card. The SIM card is assigned a unique global identifier, the International Mobile Subscriber Identity (IMSI), so that every subscriber to the mobile network can be identified with it. In addition, each SIM card is assigned a unique telephone number that can be used as a unique identifier for the subscriber as part of the teaching module. The identification of the mobile station therefore does not depend on the actual end device itself, but on the SIM card. As an alternative to the physical SIM card, eSIM has been offered since the end of 2015, which is a procedure for storing the relevant information of participants on the end device. This information is stored on a built-in, non-replaceable module of the end device (also known as a subscriber identity module).

Base station or mobile radio station:

The base station, also known as a mobile radio station, is often referred to colloquially as a transmitter mast, but in addition to the antenna it also includes a computing unit. An antenna consists of several parts, known as sectors, which can receive and transmit in different directions. A schematic representation is shown in Figure 4. The antenna transmits and receives the electromagnetic waves between the base station and the mobile subscriber, which is necessary for making calls. The received signals are forwarded from the base station to the exchange via cable (usually fiber optic). A base station covers a radio cell with its transmission and reception area. Each base station in its radio cell is assigned a cell ID, which allows identification, e.g. with the location.

*Radio cell:*

The radio cell represents the area in which a connection can be established between a terminal device and the respective base station. Cell sizes vary in diameter from a few meters to a few kilometers. Neighboring radio cells normally overlap, which enables an uninterrupted transition from one radio cell to the next. Various factors such as the weather, houses and trees or geographical conditions influence the size of a radio cell. All radio cells together form the coverage of the mobile network.

Figure 4: Construction of a base station or mobile radio station from <https://www.telekom.com/de/blog/netz/artikel/heptaband-antenne-5g-530618> (level of detail is not suitable for teaching).

Control station:

A control station manages several base stations and handles data transmission between a base station and a switching center. The control station is not explicitly identified for the teaching module in order to reduce the complexity of the setup.

Switchboard and central database:

A switching center is called a mobile services switching center (MSC) and handles the switching of connections between mobile subscribers. It also establishes a connection to other networks, such as the fixed network. Other tasks of a switching center include checking the authorization of a mobile station for a mobile connection and billing fees. An exchange has numerous interfaces to other components of the mobile network. One of these is the connection to the distributed Home Location Register (HLR) database. Data about the mobile stations is stored in this database. This enables a switching center to localize a mobile station or the respective mobile subscriber in the mobile network with the currently dialed base station. The teaching module focuses on this recorded location data, which is necessary for the efficient establishment of a mobile connection. The teaching module is reduced to one exchange with one database, although a mobile network actually contains many exchanges.

Methods for locating in the mobile network

There are various ways to localize or locate a mobile station or a terminal device, which are briefly outlined here. The teaching module looks at the location data of the base station connected to the mobile station. These can be inaccurate due to the size of the respective radio cells. Over time, various measurement methods have been developed that allow more precise positioning. For example, the Timing Advance method measures the propagation time of the signal between the mobile station and the base station in order to estimate the distance between these two elements. This makes it possible to achieve a more precise location. Another method is Enhanced Observed Time Difference (E-OTD), in which a transit time is measured between the mobile station and the base station for several neighboring base stations. The location data of the teaching module is often

Data awareness through the exploration of location data from the mobile network

compared with relatively accurate GPS data. In fact, the Global Navigation Satellite System (GNSS) method uses satellite-based positioning by the end device, which forwards this data to the base station. With this method, accuracies of a few meters deviation can be achieved.

8.6 *Examples of secondary purposes for the use of location data from the mobile network*

Example 1: Analysis of movements during the corona pandemic

Various measures (including a lockdown) were introduced to contain the Covid-19 pandemic. The effect of the measures can be examined based on the mobility of the population. For example, mobile phone data from the mobile phone provider Telefónica was provided, which was previously anonymized and aggregated. A report on this can be found at destatis. It also contains interesting visualizations that could be used in the classroom.

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

Example 2: Analysis to optimize public transport

Mobile communications providers use collected mobile communications data (especially location data) for various projects in the context of traffic optimization or the optimization of local public transport. Telefónica reports on several projects relating to public transport in Leipzig and Munich, traffic planning in Germany and the evaluation of traffic jams on German roads.

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

Similarly, the mobile communications provider Deutsche Telekom reports on a project with transport companies in Nuremberg, where anonymized location data is used to optimize transport services. The second link also contains a video that describes the secondary use of location data and could possibly be used for teaching purposes.

<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>

9 **Glossary of relevant terms:**

Data:

Data is (digitally) represented information and can be stored and processed in computer systems, for example. With regard to the concept of data awareness, it should be particularly emphasized that contexts, phenomena or even interests, emotions or actions of a person are modeled on the basis of selected characteristics. Personal data in particular, which plays a major role in data awareness, originates from a context in which the respective person is or was involved. When addressing data, the modeling aspect should not be neglected, as the contexts, phenomena or persons are not fully represented. In the context of data awareness, this means that the personal data collected does not represent an image of a person, but is merely modeled on the basis of selected characteristics. This can also result in a distorted representation of a person. It should also be noted that certain information, such as emotions or interest, is operationalized in terms of the characteristics for the collected data, which is often not trivial (What does it mean when a data-driven digital artefact stores the “interest” of the user? - What is that? How is this determined?)

Digital artifacts and data-driven digital artifacts:

In the data awareness framework, the term data-driven digital artifacts (ddA) was introduced. This describes a special type of digital artifacts. Digital artifacts are a collective term for digital tools, computer systems of all kinds, their components and their interconnection. They include hardware, software, data and objects as well as algorithms and data structures. Data-driven digital artifacts are digital artifacts that change themselves or their feedback when interacting with them by processing collected data. These often use machine learning methods, for example.

Architecture and relevance of digital artifacts:

The dual nature of digital artifacts, or duality, describes an analytical separation of aspects of a digital artifact (see above). According to this view, a digital artifact can be described from the perspective of architecture and

relevance. Architecture refers to all technological structures and mechanisms. Relevance, on the other hand, refers to the intentions, functions, opinions, interpretations and context of a digital artifact.

Explicitly and implicitly collected data:

The concepts of explicitly and implicitly collected data were introduced in the data awareness framework. These are generally linked to the user - they often represent personal data. Explicitly collected data is data that the user intends to create with their action, i.e. that they have directly and actively entered or generated. Users are usually aware of this. These are, for example, texts and images posted on social media platforms, the search term in a search engine or the telephone number of the person you want to call when making a call via the mobile network. In contrast, implicitly collected data is collected and generated indirectly through observation (tracking) or processing of already collected data alongside the actual action of the user. Users are often not aware of this data collection. In the example of the social media platform, this includes likes and clicks, in the case of the search engine, for example, clicks on search results or, when making calls via the mobile network, location data of the connected base stations.

Primary and secondary purposes of processing and use:

In the data awareness framework, the terms primary and secondary purposes of processing and using collected data were introduced. These refer to the processing and use of data about a user that is collected when using data-driven digital artifacts. Primary and secondary purposes refer to the intention with which this previously collected data is processed and used. Primary purposes include processing and using the collected data in order to be able to offer the data-driven digital artifacts with the features. These relate to the processing and use from a user perspective: The data is processed and used in order to be able to offer features to users. In the example of the search engine, this is the display of search results. This would also include personalized ordering of the search results. In terms of the feature for the user, this would mean that the user is shown the results that are ideally relevant to them. Secondary purposes means that the data is processed and used to pursue other/further purposes - e.g. further commercial or scientific purposes. This "secondary use" of the data refers to the processing and use of the collected data from a provider perspective: What else can a provider of a data-driven digital artifact use the collected data for? In the context of streaming services (e.g. Spotify), this could include using usage data (e.g. music listened to) to analyze the emotions of users.

Data Moves:

Data moves are used to describe data operations. These include the followingⁱ:

- Filtering: Forming a subset of the data
- Grouping: Dividing data into subgroups
- Summarize: Aggregating data
- Calculate: Create new attributes based on existing data (e.g. create a third column based on two columns)
- Merging/joining: merging data records
- Reorganize: for example, changing the presentation of the data

ⁱ 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).