

Guideline for the “Big data” media package

This guideline will provide an overview of the contents and context of the “Big data” media package and suggestions for methodical, didactic use of the included media. It helps you navigate the media package. It suggests a logical sequence for working with the individual media files and at the same time provides content-related input, such that the guideline serves as a quick reference.

Contents

1	Introduction to teaching this topic	2
1.1	Motivation for the topic.....	2
	Why is this topic important for your students?	2
1.2	Ideas for teaching	2
	Learning objectives of the topic of big data	3
1.3	Methodical suggestions and teaching ideas.....	3
1.4	Media selection.....	3
2	Introduction to the topic of big data (part A)	4
2.1	General introduction with relevance to the everyday world.....	4
2.2	Data tracks, data protection, and data security	5
3	Big-data practical examples (part B)	6
3.1	People and their data tracks	6
3.2	Data collections and their consequences (practical examples)	6
3.2.1	Practical example: Big data and the Internet of Things	6
3.2.2	Practical example: Big data and cloud computing	7
3.2.3	Practical example: Big data and virtual gaming worlds.....	7
3.2.4	Practical example: Big data and personalized online advertisements	7
3.2.5	Practical example: Big data and a data-driven election campaign.....	7
3.2.6	Practical example: Big data and government surveillance	8
3.3	Advantages and disadvantages of data collections.....	8
4	Opportunities and challenges of big data (part C)	8
4.1	Developments of society as a whole	8
4.2	Wearables	9
4.3	Future questions related to big data.....	9
4.4	A look to the future.....	10

1 Introduction to teaching this topic

1.1 Motivation for the topic

Big data is an important term that describes something huge: It includes the collecting, saving, and processing of such a large volume of data that it cannot be handled by normal computers. Where does this huge stream of data come from? In a world where digitalization is becoming increasingly important, every person leaves behind data. People do not even have to own a smartphone for this to happen; an e-mail address, an ATM card, or a health insurance card already suffices. But data that have nothing to do with people, such as data about the weather, particulate matter, or the Earth's movements, are also captured digitally and stored. The topic of big data thus defines a substantial element of digitalization. The linking of data to generate new knowledge is also an important component. Developments in this very broad subject area are therefore significant – for industry, science, government, and society. At the same time, big data is a collective term for which numerous definitions exist and that is closely related to the topics of algorithms and artificial intelligence.

Why is this topic important for your students?

Children and adolescents are confronted with the collection, storage, analysis, and processing of (their own) data every day, sometimes knowingly, often unknowingly. Many children and adolescents have their own smartphone and surf the Internet. The topic of big data has thus long been a part of their everyday world. Since in addition to many advantages it can also entail dangers, the topic deserves a critical approach. Adolescents adopt this approach by understanding what is behind big data, what that means for them, and how society changes as a result. The objective should therefore be to enable children and adolescents to recognize the effect of big data so that they learn to handle their own data as deliberately as possible. They should be equipped to understand the complex correlations, make decisions, and form their own opinion on the topic.

1.2 Ideas for teaching

The media of this package are designed for use in all secondary schools, especially for the subjects of engineering and information technology. However, some individual media files are also suitable for subjects like ethics, civics and politics, and business education.

The media of the “Big data” media package are designed for full development of the topic in context. But the individual media files also work alone and independently of each other, and so it is possible to use only selected media files in class (additional information on this is found in chapter 2). The media files primarily contain elements for teaching, but they can also be used for preparation of and following up with the students.

The “Big data” media package is divided into three parts: **Part A** comprises an introduction, **part B** practical examples, and **part C** opportunities and challenges. The level of complexity increases with each part. Part A provides an overview and an introduction to the topic. In part B the topic becomes more specific and the learned material is reinforced based on practical examples. Part C goes more in-depth.

The media package thus can be used with students without prior knowledge of the topic; on the other hand, students with an advanced level of knowledge can skip a part.

The media package can be used in various grade levels. It is possible to use parts A and B starting from grade 7, while the media from part C are suitable starting from grade 10. As a result, the “Big data” media package can accompany students through various grade levels until they graduate.

Learning objectives of the topic of big data

The overarching learning objectives include the following:

The students will:

- understand what the term “big data” describes and be able to define it. (part A)
- recognize specific connections to their own everyday world and understand the relevance of the topic to different areas of their lives. (part A)
- deal with the subareas such as data collection and analysis as well as their data tracks, data protection, and data security. (part A)
- be able to reflect on their own handling of data and adapt their behavior. (part A)
- become familiar with the possibilities and risks presented by big data. (part B)
- become aware of various perspectives and arguments related to the topic and be able to form their own opinion. (part B)
- be able to name key opportunities and challenges of big data, weigh them against one another, and evaluate them. (part C)
- become acquainted with the terms “wearables” and “self-tracking” and know what is meant by them. (part C)
- have an idea of how the topic of big data may develop in the future and be able to analyze these developments for themselves. (part C)

1.3 Methodical suggestions and teaching ideas

The topic of big data can be introduced by having the students exchange ideas and by conducting a survey (in class) on how they use media and which media they prefer. The topic thus becomes specific and relevant from the start, since it is pulled into the students' everyday world and daily routine. The introduction is followed by a more intense discussion in class that can take place either with input from the teacher or by having small groups or the class as a whole work through the content using the media.

The worksheets included in the media package can be used for individual work in class or as preparation for or follow-up to the respective lesson at home. The contents of the media package can also be worked through as part of a project day on big data. Learning at stations with individual stations that take up different aspects of big data would be practical for this. Such a project can also be organized for multiple classes or grade levels.

1.4 Media selection

The media and their content are based on the students' everyday world, daily routine, and consumption practices. They tie in with (any potential) prior knowledge and experiences of the students. The new knowledge of the topic is then built up systematically. For this reason, the contents are designed to be application- and case-oriented. They begin with the students' need to grasp things, to experiment, and to attain knowledge, insights, and a deeper understanding through questioning and contradictions. The media in the media package consist of various formats (information sheet, video, audio interview, graphic, and information module as well as worksheet with answer). A link list provides links for additional information.

The big data topic is closely interconnected to the topics of artificial intelligence and algorithms, which each have their own media package on the Siemens Stiftung Media Portal:

<https://medienportal.siemens-stiftung.org/en/artificial-intelligence-112778>

<https://medienportal.siemens-stiftung.org/en/algorithms-in-our-everyday-lives-113224>

2 Introduction to the topic of big data (part A)

2.1 General introduction with relevance to the everyday world

The graphic and information sheet for the introduction to big data provide the basis and starting point for the topic. The importance of the topic to the students' everyday world and daily routine is then explained. The objective is for the adolescents to understand the relevance of big data to them and the connection with digitalization and to spark their interest in looking more closely into the topic. The graphic can serve as an introduction to the overall topic of big data by having the students describe the correlations they can recognize in the graphic and augment them with any prior knowledge. The information sheet presents the correlations once again.

In everyday life, every person generates a lot of different data and thus leaves behind data tracks that are recorded. But data that do not originate from people, such as weather data like temperature or precipitation, are also collected. The largely automated acquisition, collection, storage, and processing of these data raise various questions that need to be considered more closely in further work with the media package: How exactly do social networks, for example, acquire data about their users? What type of data is being collected? What insights can be gained from these data? How can big data help? What risks are involved? What about privacy, thus the protection of data? What can each individual do to protect his or her own personal data?

Media: Graphic "Big data – Introduction"
 Information sheet "Big data – Introduction"

After the introduction, the term can be defined so that the topic is delimited. Big data means "masses of data". The definition is limited to the essential aspects that the term "big data" describes to create a basic framework and a starting point for successively looking at the complexity of this topic more closely over time. The graphic then specifically illustrates the description of big data with the help of the three Vs: volume (very large volumes of data), velocity (high speed to process the data), and variety (diversity of the data).

Media: Information sheet "Big data – Definition"
 Graphic "Big data – The three Vs"

The two media described below add further explanations to the definition and involve a certain level of abstraction. However, the content can also be skipped. It is possible to continue directly with the explanatory video after the definition.

The term "big data" stands for a variety of technological applications and methods. One form of data acquisition is tracking on websites using cookies or tracking through social networks. Fingerprinting is a new form of tracking that does not use cookies. The data are then evaluated, for example, using the data mining method. Data mining is the use of statistical methods to systematically search for and evaluate information from large volumes of data.

The graphic and the interactive information module illustrate data mining with the typical analysis types clustering (identify similarities and form groups), association analysis (establish dependencies of characteristics and rules), and regression analysis (determine correlations from available data).

Media:	Information sheet “Big data – The methods behind the term” Graphic “Big data – The data mining method” Information module “Big data – Data mining”
--------	--

After this, the term “big data” can be illustrated once more with the explanatory video and the students’ views can be expanded again. The video’s main character named Loki and the integration of applications as examples again create a direct connection with the students’ everyday world. The video answers the question of how data tracks that count as big data are left behind. This happens, for example, via social networks, location services, and fitness and tracking apps on smart-watches. These processes are often invisible. Private companies and scientific institutions then work with the obtained data. Advertisements and sales pitches are geared toward these data, or insights that can serve society are gained. In some countries, personalized advertising is not (yet) very pronounced, but the principle should be identified. There are diverse types of use and players and both positive and negative consequences, which can be illustrated with the video. This video consists of three distinct parts, and so the teacher can also play and discuss it in three parts. A teaching idea that suggests using the video in class as part of a discussion of the pros and cons is also provided. Should Loki’s data be collected, processed, and analyzed? What are the consequences of this? The teaching idea provides information on the teaching plan and suggests possible arguments as an aid for the discussion.

Media:	Explanatory video “Big data – Loki and her data” Teaching ideas “Big data – How does it work and why is it important in my life? (discussion of pros and cons)”
--------	--

2.2 Data tracks, data protection, and data security

After the video, it makes sense to pick up on the everyday reference once more and to motivate the students to delve deeper into the aspects that are particularly significant to them: data tracks, data protection, and data security. The worksheet focuses on the individual level, and so the topic is within the grasp of each student.

Assignment A queries the knowledge that the students learned from the introductory part and simultaneously requires the students to transfer the learned knowledge to their own lives. If the students have difficulty doing this assignment, the teacher can provide the answer sheet with examples to them. These examples can serve as an aid to the students. Assignment B prepares the students for an in-class discussion that forms a bridge to the topic of “data protection”. The perspective here is European; data protection is taken very seriously in Europe and there are strict legal regulations. Elsewhere, the regulations concerning data protection are not as strict. Data protection is an extremely important aspect of the topic of big data that affects the students very directly and that frequently is not reflected upon. Data protection is important because personal data can be used to create microprofiles about persons. Big data brings enormous changes to the topic of data protection. A private person can protect himself or herself, for example, through data minimization by carefully considering what data to place on the Internet. Practicing data hygiene is also important. This means deleting data after some time. An important criterion for supporting data protection is ensuring that the members of a society are aware of the issue.

The discussion reveals how difficult it is to perform some activities without leaving behind data tracks via services or apps. An activity quickly becomes “inconvenient” without services and apps because their offers often help us by providing recommendations and suggestions. However, these recommendations and suggestions are based on the data that we have given to the services and

apps. All people are thus caught in a dilemma. The goal here is not for the teacher to state a solution. Instead, the goal is to make the students aware of this dilemma and the difficulties related to the topic.

Medium: Worksheet “Big data – Data tracks” (with answer sheet)

3 Big-data practical examples (part B)

3.1 People and their data tracks

Part B highlights various practical examples. An information sheet and a graphic that generally introduce the topic serve as the starting point.

To start with, the graphic presents the practical examples with symbols and headings. The teacher can use it as an introduction to part B and first lead an open discussion on the following questions: What examples are you familiar with? What do you know about the topic? Can you describe the connection between one of the examples and big data?

The information sheet aims to provide information on data collection and evaluation in greater detail that builds on the introductory part of the media package but that has not yet been considered. The practical examples it contains, which are also pictured in the graphic, substantiate the information, and the connections between the examples are also shown.

Media: Graphic “Big data in practice”
 Information sheet “Big data in practice”

3.2 Data collections and their consequences (practical examples)

The six practical examples work independently of each other, and so it is possible to use only one, several, or all examples. The related graphics visualize the functioning of the respective big-data practical example or key parts thereof, especially the aspects that are connected with data collection and evaluation. They can each be used as an introduction for an open discussion. The information sheets once more explain each practical example more comprehensively in words based on the following key questions:

- What is the example about?
- What does this have to do with big data? For which subarea does it serve as an example?
- How are data collected/analyzed in this case?
- What data are transmitted? For what purposes can the data be further utilized?
- What is worrying about this?

3.2.1 Practical example: Big data and the Internet of Things

In the so-called Internet of Things, devices and objects collect measured data and input. They can be controlled remotely and can interact with each other. They are designed to make people's lives easier and better. In the process, they collect lots of personal data about the users.

Media: Information sheet “Big-data practical example – The Internet of Things”
 Graphic “Big-data practical example – The Internet of Things”

These media are also suitable for physics class (as a spotlight).

3.2.2 Practical example: Big data and cloud computing

Many individual pieces of data are now stored in a so-called cloud using services like Dropbox and Google Drive. Such cloud computing services are characterized by the fact that they rely on a comprehensive IT infrastructure via the Internet. Online translation services also provide increasingly precise results because they use such structures and big-data applications.

Media: Information sheet "Big-data practical example – Cloud computing"
Graphic "Big-data practical example – Cloud computing"

These media are also suitable for ethics class (as a spotlight).

3.2.3 Practical example: Big data and virtual gaming worlds

Computer games normally consist of a completely virtual world; everything is programmed and exists only as code. Pokémon Go is an online game that uses geodata of the real world as a gaming backdrop and combines a virtual world including fantasy creatures with reality. Players can work together in real time.

Media: Information sheet "Big-data practical example – Virtual gaming worlds like Pokémon Go"
Graphic "Big-data practical example – Virtual gaming worlds"

These media are also suitable for ethics class (as a spotlight).

3.2.4 Practical example: Big data and personalized online advertisements

Advertising appears not only in magazines, on billboards, and on television, but also on social media platforms, in search engines, on news websites, and in weather and sports apps. Unlike traditional advertising, generally not every user sees the same advertising banners or ads online. The advertisements are in fact tailored to the profiles of the respective users.

Media: Information sheet "Big-data practical example – Personalized online advertising"
Graphic "Big-data practical example – Personalized online advertising"

These media are also suitable for civics and politics and ethics classes (as a spotlight).

3.2.5 Practical example: Big data and a data-driven election campaign

During election campaigns, political parties and their campaign aides increasingly rely on big data. They analyze profile databases that are offered for purchase, set up their own apps and websites that acquire data, and make use of available voter registries. For example, they use the data obtained and analyzed for social media communication and can selectively address definable groups there.

Media: Information sheet "Big-data practical example – Data-driven election campaign"
Graphic "Big-data practical example – Data-driven election campaign"

These media are also suitable for civics and politics and ethics classes (as a spotlight).

3.2.6 Practical example: Big data and government surveillance

The government also makes reasonable use of big data in many areas. When it comes to citizens' security and averting dangers, authorities, the police, and intelligence services occasionally turn to surveillance and analysis methods based on large volumes of data.

Media:	Information sheet "Big-data practical example – Government surveillance"
	Graphic "Big-data practical example – Government surveillance"

These media are also suitable for social studies/politics and ethics classes (as a spotlight).

3.3 Advantages and disadvantages of data collections

In the first assignment, the worksheet queries the students' knowledge about the practical examples based on the listed key questions. The second step is about the students' learning to form their own opinion on the topic with the aid of a new scenario that is closely related to their everyday world. The assignment is about tracking students as they perform tasks at school.

The students gather arguments in favor of and against the action and discuss these arguments with a classmate. By being aware of and weighing various positions and arguments, they can form their own opinion. Potential arguments in favor of and arguments against tracking are provided on the answer sheet. These points serve merely as examples and as orientation. The students' own ideas may differ from these points and are still legitimate.

Medium:	Worksheet "Big data – Tracking during class" (with answer sheet)
---------	--

4 Opportunities and challenges of big data (part C)

4.1 Developments of society as a whole

Part C starts with a general introduction. The graphic provides an overview of the opportunities and challenges and clarifies them with keywords. The illustration is based on two main areas that influence the opportunities and challenges: the legal framework and the responsible handling of data. The graphic can be used as an introduction to part C to first lead an open discussion about the students' prior knowledge and their associations with the terms. Afterwards, these associations can be substantiated by the opportunities and challenges addressed in the information sheet. The information sheet mentions key opportunities and challenges that arise in the areas of sustainability, industry, government, law, and society. After the class has discussed the keywords, this information sheet provides additional specific content and context based on examples.

The big challenge for companies is handling the data responsibly and transparently. In addition, the particular challenges for the government and society are establishing and adhering to frameworks and limits. The example given here is the European Union General Data Protection Regulation (GDPR) with extensive rules and regulations that went into effect in March 2018. This poses a dilemma: On the one hand, data must be protected against misuse, and on the other, innovation should not be impeded. For example, the knowledge gained using big data has the potential to be used in early warning systems for natural catastrophes and as information on climate change. Especially in the medical field, big data provides opportunities in the research of various diseases.

Media: Graphic “Big data – Opportunities and challenges”
Information sheet “Big data – Opportunities and challenges”

4.2 Wearables

Devices known as wearables are addressed in an infographic and an interactive information module. These devices are wireless systems worn on the body that transfer data to smartphones, tablets, and computers either directly or via radio protocols. This example substantiates opportunities and challenges. The presented subareas can be looked at and discussed individually or as a whole. A possible work assignment for the students is taking the presented information as a basis to discuss whether they can imagine using wearables in the future and giving reasons for their decision.

A graphic and information module introduce wearables based on examples and areas of application. They include smartwatches and step counters, which use self-tracking. Self-tracking focuses primarily on users’ bodily functions and are intended to improve their health and fitness. Several parties are interested in the acquired data: In addition to the person using the wearable device, interested parties include companies, health insurance companies, and scientific institutions. Opportunities include users’ ability to look after their own health better and facilitation of medical examinations. Challenges include an increasing shift of responsibility from the company to the individual and misuse of health data with commercial value.

The link list provides some sources on wearables and on health and fitness trackers that can be consulted for more detailed information.

Media: Board content “Big data – Wearables: Examples, application areas, opportunities, and challenges”
Information module “Wearables”
Link list “Big data – Link list”

4.3 Future questions related to big data

We must think about the questions that society must ask in the future regarding big data, both in terms of laws and ethics. The students should realize that freedoms also require structure, which is the responsibility of, for example, civil society including the students themselves.

Industry and science, especially medicine, have discovered the relevance of big data and are significantly farther ahead than, for example, the government, which often still lacks the understanding of what strategies or (legal) rules are necessary. The big challenge is thus finding answers to the following questions: Where does the government have to intervene? What laws are needed? However, it must be kept in mind that freedom must be preserved, for example, for research so that it can be determined what is possible with big data.

Aside from the challenges, there are also opportunities, for instance, to combat climate change using a larger basis of data. Ethical questions arise to the extent that existing prejudices in society will be confirmed, intensified, or reinforced by assumptions that are transferred to big-data methods. New algorithms and artificial intelligence are crucial technological developments because they make it possible in the first place to draw conclusions from big data.

The following key questions can be discussed with the students:

- Who has an interest in working with data? Why does this interest exist?
- What risks arise from a lack of regulations?
- What challenges are there? What questions arise?
- What areas are afforded opportunities through big data? What do these look like?

4.4 A look to the future

As a way to encourage reflection about possible developments, a look is taken to the future as a conclusion or as follow-up to the topic of big data. The students imagine possible future scenarios and thus evaluate current developments. This also once again illustrates that creative freedom exists currently and in the future, and the shaping of this creative freedom can have a direct effect on the students' own everyday lives. A worksheet presents five fictitious future scenarios from the areas of grocery shopping, aging, traffic, nutrition, and online shopping. The students rate the scenarios from their points of view and then give reasons for their rating. The answer sheet shows possible aspects that could play a role in the rating.

Medium: Worksheet "Big data – Future scenarios" (with answer sheet)