Data Storytelling to communicate Big Data internally - a Guide

Oberascher, Lisa; Ploder, Christian; Spiess, Johannes; Bernsteiner, Reinhard; Van Kooten, Willemijn

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

![gesis Leibniz-Institut für Sozialwissenschaften logo]   Mitglied der ![Leibniz-Gemeinschaft logo]

JEL Classification: D7, D8

Data Storytelling to Communicate Big Data Internally – a Guide for Practical Usage

L. OBERASCHER ‡,
C. PLODER ♯,
J. SPIESS♯♯,
R. BERNSTEINER♯♯♯,
W. VAN KOOTEN♯♯♯♯

Purpose: Data is collected from all aspects of our lives. Yet, data alone is useless unless converted into information and, ultimately, knowledge. Since data analysts, in most cases, are not the ones in charge of making decisions based on their findings, communicating the results to stakeholders is crucial to passing on information of data-driven insights. That is where the discipline of data storytelling comes into play. Often, data storytelling is considered an effective data visualization. Creating data stories is a structured approach to communicating data insights as an interplay of the three elements data, visuals, and narrative. Sharing data-driven insights to support better business decisions require data storytellers skilled in the "art of storytelling".

Design/Method/Approach: In this paper, the authors discuss the use of data storytelling in business to communicate data to stakeholders for improving decision-making. The findings are derived from (1) an extensive literature review and (2) a qualitative analysis of 13 expert interviews with people incorporating data storytelling into their daily work within their jobs in international companies.

Findings: These interviews revealed the importance of providing a flexible tool to support knowledge sharing for people communicating complex data to internal stakeholders. Combining literature with qualitative research enabled the authors to create the "data storytelling cheat sheet", a guide for practical data storytelling.

Theoretical Implications: Theories like the Psychological distance or the idea of the theory of dual processing dual are used to base our research idea on. There was no new theory built in this paper.

Practical Implications: One of the results is an implementation systematic cheat sheet that helps practitioners to implement data storytelling in their daily business.

Originality/Value: The theory of data storytelling is overwhelming the first time to use and based on an empirical study with experts in the field a guideline for hands on use was developed under a based on a cleanly defined empirical study.

Research Limitations/Future Research: The paper focus on internal data storytelling – maybe with external stakeholders it might be slightly different. The results the data communication part in any data analytics project.

Paper Type: Empirical

‡Lisa OBERASCHER,
MCI – Management Center Innsbruck - Internationale Hochschule GmbH, Austria
e-mail: lisa.oberascher@mci4me.at

♯Christian PLODER,
MCI – Management Center Innsbruck - Internationale Hochschule GmbH, Austria
e-mail: christian.ploder@mci.edu
https://orcid.org/0000-0002-7064-8465

♯♯Johannes SPIESS,
Joint Systems Fundraising- & IT-Services GmbH, Austria
e-mail: johannes.spiess@sos-kd.org

♯♯♯Reinhard BERNSTEINER,
MCI – Management Center Innsbruck - Internationale Hochschule GmbH, Austria
e-mail: reinhard.bernsteiner@mci.edu
https://orcid.org/0000-0002-8142-3544

♯♯♯♯Willemijn VAN KOOTEN,
MCI – Management Center Innsbruck - Internationale Hochschule GmbH, Austria
e-mail: willemijn.vankooten@mci.edu
https://orcid.org/0000-0002-9784-444X

_____

Reference to this paper should be made as follows:
Oberascher, L., Ploder, C., Spiess, J., Bernsteiner, R. & Van Kooten, W. (2023). Data Storytelling to Communicate Big Data Internally – a Guide for Practical Usage. *European Journal of Management Issues*, 31(1), 27-39. doi:10.15421/192303.

Сторітелінг даних для внутрішньої комунікації великих даних – посібник для практичного використання

Ліза ОБЕРАШЕР[‡],
Крістіан ПЛОДЕР[‡],
Йоганнес ШПІСС[#],
Райнхард БЕРНШТАЙНЕР[‡],
Віллемін ВАН КУТЕН[‡]

[‡]*MCI - Центр менеджменту Інсбрук - Міжнародний університет прикладних наук, Австрія*
[#]*Joint Systems Fundraising- & IT-Services GmbH, Австрія*

Мета роботи: Дані збираються з усіх аспектів нашого життя. Проте самі по собі дані марні, якщо їх не перетворити на інформацію і, зрештою, на знання. Оскільки аналітики даних у більшості випадків не відповідають за прийняття рішень на основі своїх висновків, донесення результатів до зацікавлених сторін має вирішальне значення для передачі інформації, що ґрунтується на даних. Саме тут у гру вступає дисципліна сторітелінгу даних. Часто сторітелінг даних вважають ефективною візуалізацією даних. Створення історій даних - це структурований підхід до передачі інсайтів на основі даних як взаємодії трьох елементів: даних, візуалізації та наративу. Для того, щоб ділитися аналітикою на основі даних для підтримки кращих бізнес-рішень, потрібні оповідачі даних, які володіють «мистецтвом розповіді».

Дизайн / Метод / Підхід дослідження: У цій статті автори обговорюють використання сторітелінгу даних у бізнесі для передачі даних зацікавленим сторонам з метою покращення процесу прийняття рішень. Висновки зроблені на основі (1) широкого огляду літератури та (2) якісного аналізу 13 експертних інтерв'ю з людьми, які використовують сторітелінг даних у своїй повсякденній роботі в міжнародних компаніях.

Результати дослідження: Ці інтерв'ю показали важливість надання гнучкого інструменту для підтримки обміну знаннями для людей, які передають складні дані внутрішнім стейкголдерам. Поєднання літературних джерел з якісним дослідженням дозволило авторам створити «шпаргалку зі сторітелінгу даних» - посібник з практичного сторітелінгу даних.

Теоретична цінність дослідження: Теорії, такі як психологічна дистанція або ідея теорії подвійної обробки даних, були використані для побудови нашої дослідницької ідеї. У цій роботі не було створено жодної нової теорії.

Практична цінність дослідження: Одним із результатів є систематизована шпаргалка, яка допомагає практикам впроваджувати сторітелінг даних у своїй повсякденній діяльності.

Оригінальність / Цінність дослідження: Теорія сторітелінгу даних є надзвичайно складною для першого використання, і на основі емпіричного дослідження за участю експертів у цій галузі було розроблено посібник для практичного використання, який базується на чітко визначеному емпіричному дослідженні.

Обмеження дослідження / Майбутні дослідження: Документ зосереджений на внутрішньому сторітелінгу даних - можливо, із зовнішніми зацікавленими сторонами він може дещо відрізнятися. Результати є частиною передачі даних у будь-якому проекті з аналізу даних.

Тип статті: Емпіричний

Ключові слова: сторітелінг даних, великі дані, звітність, візуалізація даних.

1. Introduction

Recently, the term data storytelling and the rise of big data rapidly gained popularity (*Segel & Heer, 2011*). "Today, we are drowning in data and starved for information" (*Brown, 2014*). Businesses mostly rely on quantitative data KPI (Key Performance Indicators) to measure the success of their business processes and decide on future strategies. According to *Segel and Heer (2011)*, numbers alone often do not help communicate relevant information beyond the data since data alone is useless unless converted into information and, ultimately, knowledge (*Brown, 2014*). While data visualization uses different visualization tools, techniques, and know-how to present information better understandable to stakeholders, data storytelling goes even one step further. It gives the data meaning to support the decision-making process. Data storytelling can be considered the "why" to the "what" (*Segel & Heer, 2011*) and can especially support decision-making.

Different user groups in the company as well as outside of the company need different types of information and can rely on different background experiences. For decision-making it is sometimes important that some of these users have to be combined to decide on a particular question. The theory of psychological distance can be used to explain these phenomena.

Psychological distance theory is a concept in psychology that refers to the idea that the perceived distance between an individual and a particular event, object, or concept affects their mental representation of that event, object, or concept. This perceived distance can be physical, temporal, social, hypothetical, or factual, and has been shown to impact cognition, emotion, and behavior. For example, research has shown that psychological distance can affect decision-making by altering the salience of certain attributes of the event, object, or concept being considered, and can also impact the representation of risk and uncertainty associated with those events (*Liberman & Trope, 1998; Trope & Liberman, 2010*).

Playing with colors, highlighting elements, or differing objects' sizes are just examples of creating graphics. However, sometimes a better-looking chart in the wrong context is not enough to effectively communicate information to stakeholders responsible for decision-making. The effectiveness of communication can be explained with the theory of dual processing which refers to the idea that the human mind operates in two different modes of thinking: the intuitive, automatic System 1 and the reflective, controlled System 2. In the context of data storytelling, combining these two ideas means considering both the emotional, intuitive appeal of a story and the logical, reflective analysis of the data. This can be done by presenting data in a visually appealing and easy-to-understand format that can immediately engage the viewer's System 1, while also providing clear and concise explanations for the data's meaning and insights that can be processed by the viewer's System 2 (*Kahneman, 2011*).

Additionally, the narrative portion of data communication, known as "storytelling", is crucial. Stories are remembered up to 22 times more than facts alone (*Aaker, 2021*). Connecting the visualized data with a story helps to focus on the most relevant part of the data and hence reduces the complexity that comes along with the tremendous amount of business-related data. Data storytelling presents tailored information in an informative way and affects a specific audience emotionally. It is closely tied to data visualization, data journalism, and data presentation, supported by modern tools. This modern approach to delivering data to stakeholders can be described as a rather creative process; hence, best practices and experience in this field are considered the most helpful (*Feigenbaum & Alamalhodaei, 2020*).

While the process of gathering and delivering data follows a predefined process, data scientists cannot follow structured manuals when it comes to visualization and narrative. According to Hartmann (2020), director of analytics for Google, a modern data analyst should have three critical roles in today's business world.

First, understanding how the firm operates and influencing its strategy. Second, being an expert in collecting and managing data, and third, having creative skills to create innovative visuals that stick in the mind of people. Blending all three functions, it becomes clear that for the delivery of powerful insights to data, possessing a like art + science mind according to *Hartmann (2020)*, is crucial. By communicating visually, people can understand a message much faster and more in-depth with the help of excellent visualization and effective communication skills (*Ryan, 2016*).

Part one of the paper consists of a literature overview (chapter 2) of the relevant topics around data storytelling. This theoretical basis helped the authors create an interview guideline (chapter 4) for part two, where 13 interviews were conducted with experts to answer the research question (chapter 3). The goal was to consolidate best practices (chapter 6) regarding techniques, methods, and skills used in data storytelling to present insights (chapter 5) of big data to decision-making stakeholders within a company.

2. Theoretical Background

The theoretical background provides all the necessary knowledge for the empirical study and is based on a selection of current literature that was chosen following the methodology of *Snyder (2019)*. Most of the literature has been published after 2015 because the term "data storytelling" is relatively new to scientific research. The interview questionnaire is based on the conducted literature review.

2.1. Data Storytelling

Data storytelling consists of two terms. Whereas data represents all figures and signs required for data analysis, storytelling is the narrative part in charge of telling a good story (*Gadatsch & Landrock, 2017*). Traditional forms of storytelling are, e.g., books, movies, and comics (*Soares de Lima, Feijó, & Furtado, 2020*). Data storytelling is creating stories out of data (*Dykes, 2016*) and presenting the retrieved information in a controlled way (*Lee et al., 2015*). *Harris (2007)* argues that data stories, equal to commonly-known stories, consist of the essential elements of Who, What, Where, When, Why, and How. Other authors describe the term "story" more precisely and state that stories communicate information in a psychologically-efficient format (*Gershon & Page, 2001*) within clearly-defined steps (*Kosara & Mackinlay, 2013*). Often, data storytelling is considered an effective data visualization. Creating visual data stories is a structured approach to communicating data insights with the three elements "data, visuals, and narrative" as a basis (*Dykes, 2016*). Since stories are a powerful abstraction for understanding patterns and conceptualizing threats (*Elias, Aufaure, & Bezerianos, 2013*), storytelling supports the discovery, presentation, and communication of data (*Ramm et al., 2021*). According to *Bladt and Filbin (2013)*, data storytelling can be considered a data humanization.

With the rise of big data, companies nowadays don't struggle with too little information, but deal with the vast amount of data generated daily (*Ryan, 2016*). According to *Gadatsch and Landrock (2017)*, big data often refers to increased data volume, velocity and variety. Technical innovations and steadily increasing computing performance have made organizing, storing, and processing large data sets more accessible. However, as businesses amass data and the information pool continuously grows, managing data and creating data insights becomes more complex. It is not simply about having a lot of data, but about how to generate additional value. Visualizing data in a Business Intelligence context is about building reports and creating dashboards according to KPIs. Big data led to the need for going beyond traditional data visualizations. It is no longer sufficient to visualize, but it is necessary to communicate visually and explain the patterns and relationships of such massive data. Information sharing visually includes sound design principles and interactivity with information, such as with the audience. Data storytelling comprises data visualization stories. It helps to understand data relations more

effectively and quicker and engages the user to make better and faster decisions (*Knaflic, 2015*).

The role of storytelling in data is not just a trend, but it is part of daily business activities and affects internal decision-making. The increasing competition requires faster actions supported by improved data storytelling as a powerful tool (*Ryan, 2016*). Access to and the knowledge of how to use particular tools is not enough to tell compelling stories with data. Data storytelling analysts communicate the information required for decision-making by finding the true story beyond the data and giving data meaning and context (*Knaflic, 2015*). To share data successfully in a business regard and to attract the audience's attention, it is crucial to become skilled in the art of storytelling. According to *Davison (2016)*, storytelling can be learned and appropriate skills acquired. Storytelling is the process of narrative writing, whether fiction or for business data, and should be illustrative and memorable. The audience should become inspired and participate (*Davison, 2016*). Stories help to understand complex topics and evoke emotions, making impressions more memorable. In a business context, data storytelling can help to create and more efficiently understand information out of the increasing data complexity (*Neifer et al., 2020*).

To tell a proper story, it is required to understand and combine the three elements (i.e., narrative, visuals and data) acting as a framework in data storytelling and bringing them together as the value of data story telling in the middle where all of them overlay (*Fig. 1*). Data paired with a narrative (i.e., story) explains the context of the data and why it is essential. Data, combined with visuals, enlighten the audience and clearly show the stakeholders what is happening with the data. Patterns or specific structures can be established this way, which otherwise could not be seen properly. As narrative and visuals unite, stakeholders become engaged and are forced to make decisions (*Dykes, 2016*). The transformation process from data-to-data story involves three steps. First, exploring insights from the data (e.g., patterns), then putting the insights into a narrative, and as a last step communicating this story to the audience (e.g., stakeholders) (*Lee et al., 2015*).

According to *Nagel and Ludwig (2020)*, data analytics is a discipline of data science. It deals with the analysis of raw data to create insights out of it and to support decision-making. This discipline aims to extract information from big data with the help of specific methods (*Nagel & Ludwig, 2020*). Due to the continuing digitalization and big data trend, more and more data is available, resulting in significant opportunities for businesses to create valuable insights. That is where the traditional data science approach reaches its boundaries, and data storytelling comes in. Stories are ever a part of our society and have been told for thousands of years (*Davison, 2016*). Digital storytelling exploded recently, enabling contemporary storytellers to communicate via a broad set of channels and tools (*Sanchez-Lopez, Perez-Rodriguez, & Fandos-Igado, 2020*). In business, complex data sets need to be communicated to stakeholders who are usually neither data scientists nor data affiliates (*Neifer et al., 2020*). Especially since *Gershon and Page (2001)* published their seminal work on narrative visualization, researchers have begun to increase their attention to the communicative aspect of data (*Obie et al., 2019*). Data storytellers need to present the data in a tailored way to respective stakeholders to create understanding (*Neifer et al., 2020*). Lacking knowledge and skills in interpreting data for better decision-making could prevent businesses from using their embedded analytics resources to their complete potential. Therefore, storytelling in data analytics is considered a narrative technique to generate valuable knowledge from a large amount of data available (*Boldosova & Luoto, 2019*).
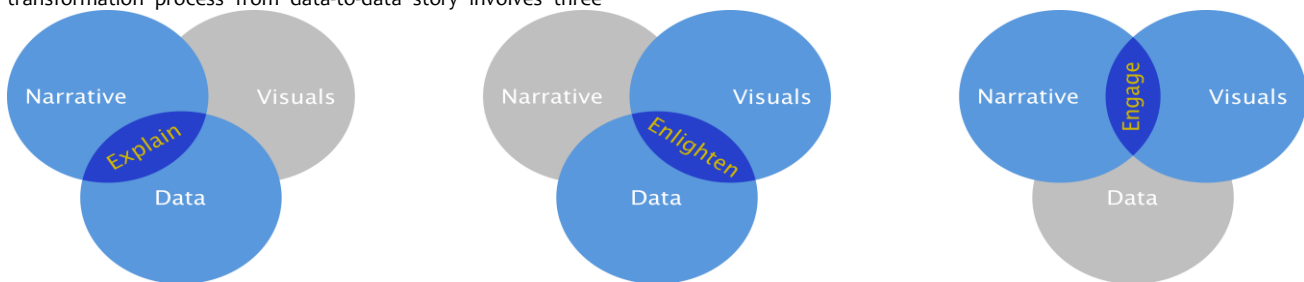

Figure 1: Main three elements of Storytelling. Modified after *Dykes (2016)*.

Source: Dykes, 2016

2.2. Data Visualization

Data visualization can be understood as "the graphic representation of quantitative information" (*Chen, Härdle, & Unwin, 2008, p. 16*) and is defined as abstracted information put into a particular schematic form (*Lankow, Crooks, & Ritchie, 2012*). Visualizations are a storytelling medium to tell compelling data stories (*Lee et al., 2015*). They help explore data, identify patterns, and ultimately get data-driven insights (*Chen et al., 2008*). Graphics reveal a high volume of data in a small space. Traditionally, measured quantities known as data graphics are visually displayed using a combination of numbers, lines, points, symbols, words, a coordinate system, shading, and color (*Tufte, 2001*). Which way data is represented depends on the available information. Most commonly used are line graphs, bar charts, pie charts, scatter plots, or maps (*Friendly & Wainer, 2021*). Infographics in static (i.e., picture, snapshot), interactive, and motion (i.e., multimedia animations) are specific types of data visualization. They are a particularly suitable way to present data to laypeople (*Locoro et al., 2017*).

Human brains are visual, and data analysis aims to find patterns like structures, groups, and trends in data within big data pools. Presenting such patterns is only possible by transforming the data into appropriate visuals (*Ware, 2020*). Measuring the effectiveness of data visualization is challenging and controversial. The scientific approach often relies on studies and tests using eye-tracking tools, which is not feasible in the business world (*Matzen et al., 2018*). The effectiveness of communicating the graphic illustration depends on the creator of the visualization. The information (content) is put into context with other related material, visualized appealingly (design), and constructed in a way that all involved stakeholders understand (*Chen et al., 2008*).

2.3. The Visual Narrative

Narrative visualization combines information visualization with storytelling techniques (*Obie et al., 2019*), guiding the viewer through a particular selected set of information (*Lankow et al., 2012*) to help detect causal relationships (*Roels, Baeten, & Signer, 2017*). Visual stories consist of several visualized story pieces, incl. annotations (e.g., labels, text) connected in a specific order to support a particular communication goal. For example, a visualization could show the energy consumption change over the years, clearly highlighting some aspects of the visual environment, which the author uses to deliver a particular message, like educating the viewer (*Lee et al., 2015*). By combining exploratory information with communication, the intention is to tell a specific story, which can affect the user's interpretation of the conveyed story (*Hullman, 2011*). Data stories consist of several visualization steps, mainly text, and images, based on data (*Kosara & Mackinlay, 2013*). Bringing the

visualization into a narrative means organizing the visual according to different techniques. The author of the visual acts as the decision maker on how to arrange the various elements within the visual environment - giving the creator the power to direct attention to a particular object of interest (*Segel & Heer, 2011*). The visual narrative or narrative visualization can be understood as conveying an intentional story using visual, communicative, rhetorical, and persuasive techniques (*Lee et al., 2015*).

Good data stories imply that designers and researchers understand how visuals communicate (*Hullman, 2011*). Cognitive neuroscience is a field, which made rapid progress in the last years. Nowadays, it is evident that there is "a unified model of how we think visually", named the theory of predictive cognition (*Ware, 2020*). Perception, memory, and visual attention determine how a viewer perceives a visualization design (*Polatsek et al., 2018*). Whether specific techniques to create a visual are good or bad is subjective (*Chen et al., 2008*). However, an excellent visual catches the user's interest and sticks in the viewer's memory (*Polatsek et al., 2018*). Creating compelling visualizations requires data scientists' skills and techniques from designers, artists, and psychologists to decide which components to use in narrative visualization (*Segel & Heer, 2011*).

Traditionally, stories consist of a sequence of causally related events (*Ma et al., 2012*). A particular narrative structure is essential in how the audience perceives a story. Narrative structures are found across different fields, such as drama, literature, and cinematography (*Yang et al., 2022*). Within this chapter, some approaches suggested by various authors are discussed.

To begin with, Freytag's pyramid established by the German novelist and playwright Gustav Freytag, represents a classic narrative structure widely used in literature and film (*Yang et al., 2022*). This particular concept also can be transferred to data storytelling to make the storyline more demanding for the readers or listeners. Therefore, it will be elaborated in more depth in the following paragraphs.

According to the dramatic structure of Freytag's pyramid (*Freytag, 1895*), a story starts with an introduction of the setting and progressively increases the tension until it reaches a peak at the climax. Afterward, stress falls again until the story ends. Data stories show similar structures. Based on Freytag's pyramid (*Freytag, 1895*), *Dykes (2016)* developed a four-stage narrative structure, the data storytelling arc (see Fig. 2), for telling data stories and hence driving business decisions.
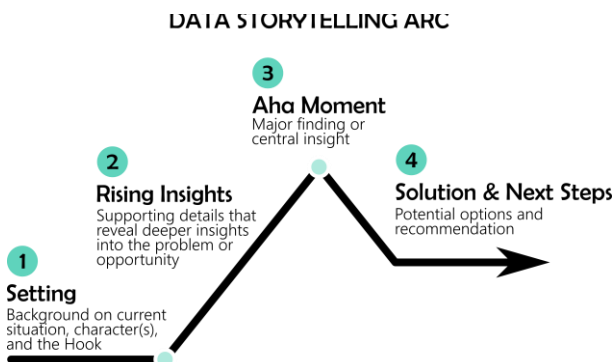

Figure 2: Data storytelling Arc based on Freytag's pyramid. Modified after *Dykes (2020)*.

Source: *Dykes, 2020*

The four major narrative categories classically used in cinematography are Establisher, Initial, Peak, and Release (*Yang et al., 2022*). While during the Establisher-phase, the interaction is set without actions taken by only showing the state of being, the Initial-phase initiates the primary event and the tension. At the Peak-phase, where the things of the highest importance happen, the narrative tension reaches its peak. It is finally released within

the Release-phase as a wrap-up for all previous events, commonly resulting in an outcome or resolution (*Cohn, 2013*).

2.4. The Visual Data Storytelling Process

The described three elements – data, narrative and visual (*Dykes, 2016*) – work together like a rack-wheel to create data stories. *Lee et al. (2015, p. 86)* propose a working model called the "data storytelling process" (see *Fig. 3*) to turn data into a visual story, which is derived from several models in available data journalism literature. Creating a compelling data story engages several roles, from data analysts exploring and analyzing the raw data, journalists building the plot of selected data, to graphic designers preparing the visualizations, presenters responsible for delivering an intended message, and the audience providing feedback. While it is possible to divide mentioned tasks, it is also likely that one person taking charge represents all roles.

In the following chapters, the authors will apply data storytelling in a business context to learn more about the best practice implementation of the day-by-day corporate operations and answer the research question in chapter 3.

3. Research Question

Because data analytics in the context of a "science mind" alone do not meet the requirements to create an appealing data story, the research questions are: Which techniques are considered most effective from a data storytelling point of view, and what type of skills should a data storyteller have to extract and present the most relevant information? How can data be prioritized, and in which way should it be delivered to ensure that the information is easy to understand for the audience?

4. Methodology

For the empirical part of this paper, a qualitative research method was used, based on 13 semistructured interviews (20-30 minutes each) with experts, mainly data analysts and data scientists, in and around the field of data storytelling based on convenience sampling combined with a LinkedIn search on experts. The empirical study was held in the time between March and May 2022 with international participants. This paper focuses on collecting best practices in data storytelling by comparing theory with practice, created a need for qualitative research data directly from the field.

As one quality criterion for choosing proper experts, the authors interviewed professionals across different job positions and hierarchy levels – leading (L), non-leading (NL), and self-employed (SE) – to evaluate whether the job position led to a different approach to data storytelling (*Tab. 1*). To put it straight at this point, there were no notable differences in the statements based on this criterion in the data analysis. Other quality criteria to avoid having one-sided results were that professionals of different genders were interviewed, that the experts had to have at least two years of experience working with and communicating data projects, and that the professionals were working in at least different departments within the same organization or for various companies. The first contact of all the experts was established via e-mail and the appointment for the interview was fixed via e-mail too.

Connecting theory and practical-oriented research enabled the authors to look for patterns and deviations to add value to practical data storytelling. That is how the idea of a suitable data storytelling guide evolved. The outcome of combining an extensive literature review with the conducted interviews is the "data storytelling cheat sheet", which should help data professionals to incorporate data storytelling in their job to communicate data to internal stakeholders better and, with this, make data storytelling a more conscious, rather than subconscious process.
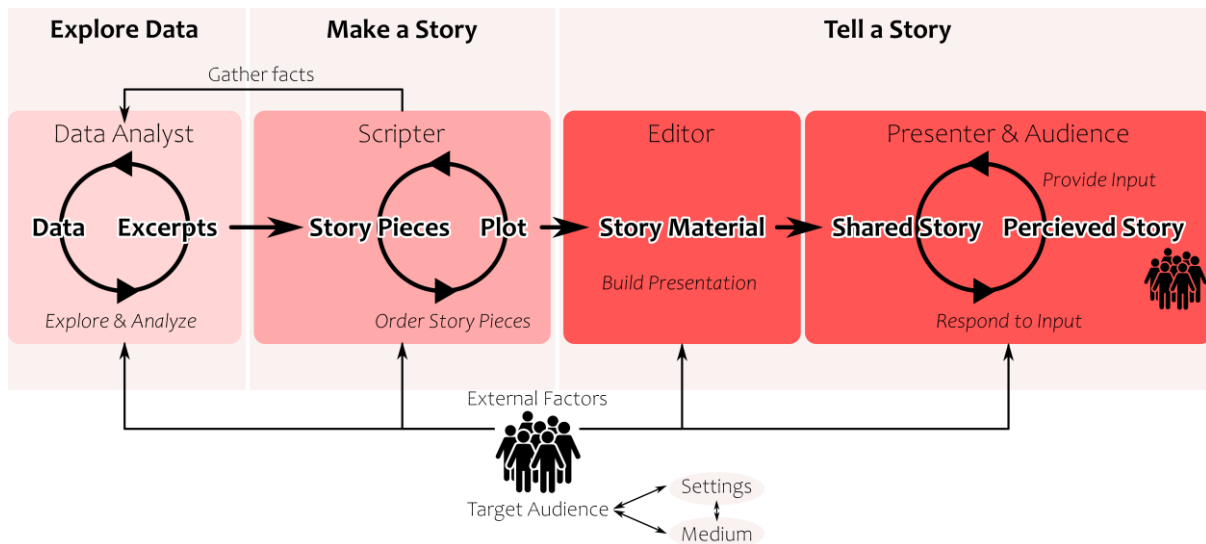
Figure 3: Transformation from data into visual stories. Modified after *Lee et al. (2015)*.

Source: *Lee et al., 2015*

Table 1: Interview partners overview for all 13 interviews (E1–E13) showing working experience in years, duration of interview in minutes, gender (male – M; female – F), job position (leading – L; non-leading – NL; self-employed – SE) and country

Interviewees	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13
Experience (years)	7	8	10	10	10	5	2	12	15	10	2	15	5
Duration (min)	24:12	26:21	28:36	29:10	28:29	31:52	26:02	28:57	28:46	21:39	31:33	25:06	29:04
Gender	F	F	M	M	M	M	M	F	F	M	F	F	F
Job position	NL	NL	NL	L	NL	NL	NL	NL	SE	L	NL	L	SE
Country	AUT	GER	AUT	ESP	USA	USA	ESP	AUT	GER	GER	ESP	AUT	USA

Source: *Research Results, 2022*

Step 1 – Interview guide: First, the authors developed literature-based questions to learn about the expert's experience in the respective fields, discussed in chapter 2. In addition to the interview guide for the interviewees, a one-pager was created to give background information and research explanation to the experts.

Step 2 – Conducting interviews: The interviews were recorded via Zoom and the interviewers tried their best to stay with a formal communication style – which worked out very well.

Step 3 – Interview transcription: Based on the collected audio files, the interviews were transcribed. Some reactions such as "laughing" are noted in the transcripts.

Step 4 – Analysis of interview: To analyze the material, the qualitative content analysis method (*Kuckartz, 2014*) was used and. In this approach, *Kuckartz (2014)* applies three basic methods of qualitative text analysis, namely the thematic analysis, the evaluative analysis, and the type-building analysis. According to *Kuckartz (2014)*, thematic investigation is known among researchers and is proven reliable and transparent. Their relatively open and modern approach closely introduced to the interview material, made the authors consider it the most suitable analysis method.

The analysis process started systematically by creating categories based on literature. Since literature already formed the basis for setting up the interview guide, 21 main categories were formulated based on the interview questions. To analyze the interviews, a matrix was generated using MS Excel and the created categories within rows and the content of each interview within columns were matched. Following this approach, each interview was summed up according to each category. As a next step, the insights across all interviews were summarized per category, and sub-categories were established.

5. Results and Discussion

ab. 2 shows the main categories (sub-categories can be found in the appendix). The categories start with background questions on work experience, current job, and educational and professional background, and become more specific to confident data storytelling-related aspects. In the section below

Tab. 2, the categories are discussed and enriched with information from the analysis. Not all categories are derived from literature (deductive). Some of them were generated during the research inductively.

Table 2: Main categories of the interviews. Deductive (de) categories are derived from literature and are based on the interview guide. Inductive (in) categories were generated during the research process

Categories		Categories	
1) Working experience with data	de	12) Stakeholders	de
2) Current job	de	13) Ethical data storytelling	de
3) Educational background	de	14) Tools	de
4) Career path	in	15) Dashboards	de
5) Work Routine	in	16) Data communication	de
6) Own Definition of data storytelling	in	17) Data journalism	in
7) Data storytelling first contact	de	18) Future of data storytelling	de
8) Recommended Skills	de	19) AI-driven data storytelling	de
9) Data visualization	in	20) Personal improvement	in
10) Explicit data storytelling knowledge	in	21) Data storytelling advice	in
11) Data storytelling within the organization	de		

Source: *Research Results, 2022*

Working experience: ranged from 2 to 15+ years.

Current job: The interviewees are mainly working in the area of data analytics.

Educational background: There is no striking pattern regarding educational background/specific studies. The interviewees have diverse educational backgrounds. "And actually, I end up here in a weird way because first, I studied environmental sciences" (E11). What is noticeable, though, is that the more scientific the background is, the more complex the interviewee perceives data communication/data storytelling "[…] it's also important because I am always, maybe due to my background, rather in detail […]" (E7).

Career path: All interviewees perceive their career as an evolutionary/development process. Learning on the job, trial and error, and doing experiments help to learn how to present the correct data. "I applied for the job, and the analytical part was more of a learning by doing for me" (E1). Communication with the stakeholders is crucial to know how to look at data and prepare and measure it (in terms of which method to use). Listening and observing words, body language, and subtle reactions help understand the stakeholder's needs. "You have to listen to your target group and their expectations […]" (E5). Learning soft skills such as communication skills is as crucial as acquiring hard skills like programming, data visualization, etc. Soft skills can be acquired primarily by being open-minded and curious, and hard skills by attending courses, seminars, training, watching tutorials, searching the internet, and education. "[…] or I learned mainly doing, by reading books, looking at tutorials, etc." (E5).

Work routine: Since all interviewees regularly work on projects, specific habits could not be pointed out. As a rule of thumb, answering the business question comes first. Only then it's about dealing with the numbers. "[…] communication with your clients is the key, so you need to get to know their business goals […]" (E1). The structure should range from a general overview, beginning with an outline of what the stakeholder wants as a result, to researching and getting to know the data in detail (descriptive analysis), "[…] like really understand the data and get the insights out of it. Like the descriptive analysis part" (E11). Getting a business understanding is a process that takes time. To not lose orientation during the project, it helps to remind oneself of the required results (the actual stakeholder/ business need). Feedback loops are considered very effective during this process. "[…] my advice would be to use feedback loops intensively" (E3).

Own Definition of data storytelling: Data storytelling in the broader sense was considered as commenting on data, presenting to stakeholders, and was mostly considered a somewhat subconscious process. In a narrower sense, it was explained as a structured form of data communication to enable better decision-making, as a better service to the audience, and as "Marketing and Sales for Data Science". "I was marketing and sales, turned into data science, and now turned into marketing and sales for data science again. So, this is also my understanding of what data storytelling is a little bit" (E9). Data storytelling translates complex data into simple language for the stakeholder. In a world of data, data storytelling is everywhere. "For me, Data Storytelling is everywhere […]" (E13). Interactivity between the analyst and the stakeholder enables us to embark on a data journey. Data storytelling focuses on one key message, explains why the results are essential, not just showing them, and gives an outlook of what could happen next. Data storytelling makes data "fun to consume" (E3) and does not simply show plain data. Maybe there could be a connection between fun and data quality or content depth, but that was not discussed with any of the experts. The stakeholders and their needs are put at the center. The content is adapted to their qualifications (data understanding). The analyst leads the interpretation of the results.

Data storytelling first contact: Data storytelling is considered a relatively modern term. "The term data storytelling is rather new, but the work behind it is how to communicate data so that it is understood and that the brain of our counterpart can process it correctly and take action is something that I've been working on and interested in for, what, ten years now" (E9). Before hearing about it, data storytelling was done more subconsciously and referred to as simply presenting data. Triggered by the problem that users couldn't interpret the data correctly, data storytelling became more present, and the solution was to make stakeholders understand the data results. "I did not particularly learn it. It just

happened" (E8). Data storytelling now becomes more and more present in society and university education.

Recommended skills: The skillset of a data storyteller is diverse. Data storytelling consists of personal characteristics (soft skills), hard skills, and systematic knowledge. Above all, data storytelling can only thrive within a supportive company culture, allowing us to try out new approaches. "[…] for me, if a new person would come on board, I would say try it" (E4). Being able to put oneself into the stakeholder's shoes and finding the correct language to communicate was pointed out as a top soft skill by every interviewee. Besides data knowledge and analytical skills, being able to think visually and possessing methodical knowledge is essential. On top of that, having a business understanding and putting data into context to answer business needs is considered a must for good data storytelling. "[…] you need some education, really analytical education, as well communication education" (E12). Interest in scientific disciplines like behavioral economics, strategic management, or decision theory helps better understand the stakeholders' thinking. "[…] on how data information is received in psychology, in behavioral economics in learning theory, in whatsoever discipline" (E9).

Data visualization: Whereas data visualization is considered to present the data correctly, data storytelling focuses on presenting the right data. Because visuals are easier for the human brain to process than just looking at plain numbers, visual design is an important part of data storytelling. "[…] you chance that the people remember your message is so much higher when you can build a good story around it […]" (E13). Still, it is only a part of data storytelling. The narrative part brings the actual value to the stakeholder for better decision-making. "[…] the narrative is the actual value of data storytelling […]" (E9). Good data visualization should give an at-a-glance view of what is going on. A rule of thumb is the simpler the visualization, the better.

Explicit data storytelling knowledge: All interviewees stated that in their organization, they have no systematically written-down information on practical data storytelling and that they could imagine a sort of guide which would make it easier, especially for new colleagues, to include data storytelling in their daily business more professionally. Internal knowledge is gathered from previous projects and by asking colleagues who have worked within the department for some time already. "[…] taking information from one project to the other, and we add something […] it's not very systematically written down" (E12).

Data storytelling within the organization: Most of the time, data storytelling doesn't follow a specific structure and is considered "common sense within the team" (E12). That clearly shows that data storytelling is a necessary discipline that needs more attention and to be better incorporated into daily business. Data storytelling can come in various forms, e.g., within Workshops, when presenting a PowerPoint presentation, creating videos, dashboards, or simply a flipchart. For instance, PowerPoint tells the story from the first to the last slide. "[…] it's a presentation which should demonstrate a story" (E12). In any case, it is essential to show the results and the basis for why something was analyzed in a certain way, how it was done, what the results mean, and what can be done next. Data storytelling relevance further depends on the stakeholder's understanding and the data's complexity. Most of the time, data storytelling is teamwork, and every team member working on a specific project presents their part of it. A structural approach would first introduce the topic, perform a background check and clarify the goal before performing the analysis (descriptive part and modeling) and showing the results as a third step. Most of the analysis will not be delivered since the goal is to focus on the main result(s). In the end, the business question should be answered, and the next steps provided. "[…] you have your business question, then you try to understand what's the background of it" (E11).

Stakeholders: A different approach depends on the stakeholder's background and data literacy (qualification). In the first step, it is essential to clarify the task and the needs of the stakeholders. Only

then, the appropriate methods and tools can be chosen. Getting to know the stakeholders and their needs is crucial. Not knowing the target group makes it impossible to design good data stories. "[...] you need to know if they did something already in the field, like what is their previous knowledge? Again, what are their expectations? And so on" (E11). The interviewees claimed that most of the time, they get a feeling for the stakeholders they work with, to know what is needed, how the analysis should be performed, and how the results should be presented. Stakeholders can be clustered when it comes to detail orientation. Two general types are the very number-oriented stakeholders on the one hand and the ones who want to see only the crucial details on the other. "[...] some people are very number oriented and want to know every single number and others are overwhelmed if you present them with a million different numbers, they'd rather have just, these are the most important points" (E6). Stakeholders from a higher hierarchy (e.g., C-level) want data on a higher level, whereas functional departments mainly require detailed and specific numbers. A third type is stakeholders who don't know what they want. In this case, it is recommended not to go into too much detail. No dashboard, report, result presentation, etc., should be overloaded with information. The data that is shown should become self-explanatory.

Ethical data storytelling: Since we are all individual humans, looking at data is always subjective. In the end, the data storytellers show their perspective of things. "And it's always subjective. Because, of course, we pick a focus. I mean, we pick a small focus out of a big amount of data" (E13). The goal is to find a way to be as objective as possible. Having a critical view of things, in general, is required. Using data storytelling in a manipulative way would not mean changing something about the actual outcome but rather focusing on specific areas in which the data storyteller thinks they have convinced someone. Since concentrating on the critical message leaves out other potential results, this means not telling the whole story. Creating awareness and giving transparency hence are essential to prevent ethical implications. Creating awareness refers to making the data storyteller aware that they are tagging their perspective to the audience, also when it comes to visual design, e.g., building legends. Giving transparency refers to showing step-by-step how the analysis led to the results, making the stakeholders understand the data, showing the sources (e.g., study, research, database), and maybe even giving the stakeholders a chance to look at the sources themselves. "[...] show your source and get all the people the possibility to look after it by themselves" (E13).

Tools: The choice of tools depends on the content (simple data – e.g., report vs. complex data – e.g., dashboard), methodology, resources (easy to handle tools – e.g., excel vs. time-consuming tools), the competence of the stakeholder (brief analysis – e.g., pdf's, excel vs. more complex analysis – e.g., dashboard) and required view (static data – e.g., report, presentation vs. dynamic data – e.g., dashboard). Generally, a fancier tool for data storytelling must not be the better solution. "I probably guess there are other tools, which might be shinier and something like this. It's just straightforward to use" (E3). The question is, what fits the purpose of the data better? Whereas R, Power BI, and Excel are used for analysis, tools such as Tableau, PowerPoint, Python, and Qlik view & sense offer great possibilities for visualizations. PowerPoint, whiteboards (primarily for brainstorming within the team), or even YouTube videos are practical tools for verbal presentation purposes. "[...] the difference is in a different kind of presentation... So, in the PowerPoint presentation, I need the explanation in addition. And in the YouTube video, you can explain everything without the presentation" (E12).

Dashboards: Dashboards are considered very compelling graphics, which give a quick overview of what is going on. They are especially suitable for daily business data when looking at KPIs, which should be monitored regularly, and for spotting problems within the daily business. For Dashboards to adequately help stakeholders in decision-making, implementing FAQs or a glossary make the dashboard better understandable and can prevent

misunderstandings. "The users become more independent on playing around with the data and understand it better. Of course, it is helpful to offer a FAQ or help page" (E4). Since the data is usually not presented verbally, it must be shown and commented on in a way that it becomes self-explanatory. Crystal-clear definitions and text supporting the graphs should be available. In addition, data must be fully transparent, include source information and what the dashboard shows precisely. Different target groups require different stories, which need different dashboards. "So, certain departments are talking about a specific topic and using certain terms other departments wouldn't understand or understand it differently" (E5). The advantages of dashboards are automatic updates, delivering accurate data regularly, interactive filtering options, automatically scaling graphs depending on the selection, and the adaptability to the stakeholder's needs, which leads to the independence of the user. On the other hand, such user independency and the many available features can be considered overwhelming, leading to the user getting lost in the data, and thus may discourage the user from actually using it. Keeping it simple is the key. Because the data keeps changing, no data consistency is given, which should also be considered when choosing the right tool for the intended data story.

Data communication: The data presentation represents a story and it should be possible for a person who has never heard about the topic to understand it. The goal is to make the stakeholder realize what was done and why it was done, to transfer proper knowledge from the analysis to support improved decision making. Presentations are mostly executed in teams. In general, stakeholders prefer a shorter presentation with less plain data and more, simpler visuals. The choice of how presentations are done depends on the content and methodology. "[...] it depends on the methodology or what was the analysis you were doing? Was it a descriptive analysis? So, that's easy to get to understand. However, if it is going to data science, predictive analysis, etc. It can be quite complex" (E5). For more accessible data, less data storytelling is required, e.g., descriptive analysis or dashboards. When it comes to complex data, e.g., predictive analysis and data science projects, the challenge of getting the stakeholders to understand the results increases, and data storytelling is highly required. Generally, it is not about the number of results shown but rather about the core information, which influences to decision-making. The goal after the presentation is that everybody should have the same essential numbers available in their head. Putting the data into context is crucial since stakeholders cannot understand the relations (e.g., good or bad numbers related to high or low values). "Never forget to put the data in context. Because the numbers alone are meaningless" (E13).

Data journalism: "Don't believe any data you did not create yourself" (E9), which means being critical of data results is crucial. Some parallels exist between data journalism and data storytelling: they share the same goal of delivering a core message, and transforming complex data to understandable language, leading to enhanced understanding. The process from data to insight is similar, but the depth of the elements (data, visualization & narrative) differs. The 5 W's in journalism (what, where, why, who, when) could be considered an inspiration when creating a data story. "In journalism, you also have that five questions, five W's. So, this is what I also inspire myself to do" (E8). The most significant differences were considered in perspective (media versus business context), the audience (broader, different target groups versus a clear and known audience), and the missing feedback loops with the audience.

Future of data storytelling: Since the amount of data and complexity of analysis are rising, data is getting harder to understand and better data communication is required. The goals will be faster analysis with higher flexibility and better scalability. People will put more time, effort, and resources into big data solutions and applying data storytelling. "I see a growing importance of Data Storytelling in the future because a) the importance of decisions based on data will grow and b) the fields of data, which are analyzed, will get wider and more complex" (E4).

AI-driven data storytelling: All interviewees are curious about AI (Artificial Intelligence) in data storytelling, but at the same time

suspicious of how exactly AI could work in this area. The hurdle of acceptance from stakeholders is also considered an obstacle and maybe a generation topic. "This has to be accepted on the part of the target group" (E5). AI in data storytelling could be seen as AI creating insights for decision-making on an automatized level to support the next steps.

Personal improvement: Data communication in a less rational way by using simple language to convey results to the stakeholders is seen as the biggest challenge, followed by improving their business understanding and getting to know how stakeholders think to discover their needs. "So, the data scientist has his mental model of the world and is very deep into the analytics and the data science methods and has a lot of detailed information and results from one analysis, for example, while his counterpart or her counterpart is flying on a completely different level and has hard times in receiving and processing a myriad of information" (E9). Trying out different approaches (e.g., drama classes as an experimental approach) and keeping an eye on the newest market trends were also mentioned.

Data storytelling advice: Summing up the interviewees' answers, there is no single path to becoming a good data storyteller. Curiosity and the motivation to keep learning are important. Having an "art and science mind", which involves analytical education and communication skills, is the perfect basis. Having a coach, somebody who is skilled in data storytelling, day by day. Communication with and listening to stakeholders is vital. Furthermore, feedback loops are advised. It's about telling the right story instead of many stories. The less data shown, the more it stands out and sticks. Time is precious – getting to the point and keeping the story short is recommended. Data should always be put into context and measured against a benchmark. Knowing the audience gives insights into the stakeholder's data competency and hence which analysis should be performed and how results should be presented so that they stick. Thinking abstractly is recommended to look beyond data and see what implications are hidden there. Constant improvement is inevitable and can be done by, e.g., taking courses, watching tutorials, and searching online forums. Not only the "outer" tools (referring to hard skills) but also the "inner tools" (referring to soft skills) need to be developed. Asking stakeholders and oneself the right questions helps to check whether the information is relevant to the audience and not just oneself. "So, I think it's crucial to have data science and communication knowledge. So, this is the perfect combination" (E12).

The results of the empirical data are compared with literature to show the differences and similarities between science and practice of data storytelling. A data storyteller must be able to cover the entire data storytelling process as proposed by *Lee et al. (2015)*. Data storytelling usually starts with data analysis. In addition, data and information management functions (*Carbonell, Sánchez-Esguevillas, & Carro, 2017*) are required. Deriving insights from data, picking the most relevant data points and presenting them in understandable form for the audience are key skills for data storytellers (*Nussbaumer Knaflic, 2017*). A variety of tools is available to support storytellers in visualizing their story (*Segel & Heer, 2011; Tong et al., 2018*). According to *Segel and Heer (2011)*, narrative visualization can be classified into seven genres, including magazine style, annotated chart, partitioned poster, flow chart, comic strip, slideshow, and video. Newer approaches also integrate paradigms of serious games (*Zhang et al., 2022*). The visual presentation of stories can guide audiences through the plot and improve content comprehension and memory (*Borkin et al., 2015; Brewster et al., 2019*).

Data storytelling often is based on several input data streams. Data must be cleansed and combined before they can be used. Business intelligence, business analytics and AI are approaches that support the data management from various data sources with different data formats.

Boldosova and Luoto (2019) conducted a comprehensive interdisciplinary literature review to explore storytelling as a verbal narrative technique generating memorable knowledge from big

data and business analytics. They found out that data storytelling is "a narrative sensemaking heuristic positively influencing human behavior towards BA [business analytics] use" (*Boldosova & Luoto, 2019, p. 204*). This approach has the potential to improve individual decision-making. Business analytics can be used more often, even on a daily basis, because it increases the quality of decisions (*Boldosova & Luoto, 2019*).

The introduction of AI in the field of storytelling is a relatively new approach. AI can support several activities of storytelling, e.g. monitor data sources, collect data, define characters, develop the plot over time, visualize data, measure success, etc. (*Al-Doulat et al., 2020; Dur, 2012; Kreminski et al., 2020; Thorne, 2020; Yang et al., 2019*).

More and more authors stress the importance of ethical considerations and judgment in all phases of the data storytelling process. This starts with the selection and validation of data sources, data analysis, especially when it is automated with algorithms (e.g. AI), and data visualization (*Diakopoulos, 2018; Ehmel, Brüggemann, & Dörk, 2021; Showkat & Baumer, 2021*).

6. Summary, Limitations and Future Research

After literature and qualitative empirical research, the authors can now conclude that the narrative part of data storytelling is even more critical than assumed in creating convincing data stories. There is a lot of specific theoretical knowledge available within the literature. However, the authors discovered that data storytelling is still considered a relatively modern term in organizations when it comes to its practical daily incorporation into business. The challenge lies in the comprehensive scope of data storytelling application and the uniqueness of every data analysis project.

Data storytelling can be considered a set of tools. Metaphorically, it could be imagined as a toolbox, where each device resembles a different skill and approach. Its usage depends on the business and stakeholder's need, as well as other factors like data competency, software choice or the analysis method used. The required skills are not merely hard skills representing acquired abilities learned through education, but equally or even to a greater extent also soft skills. These soft-skills are even harder to define and hence to acquire. The empirical research revealed that explicit, compact, written-down knowledge around data storytelling is non-existent in organizations but is highly sought-after. Within the 13 conducted expert interviews, a lot of practical know-how was collected. Combined with the literature review in this paper, the authors were able to develop the "data storytelling cheat sheet" – a practical guide for everyone who faces the challenge of proper data communication to internal, organizational stakeholders. The "data storytelling cheat sheet" is available on GitHub: https://github.com/loberascher/Data-storytelling-cheat-sheet.

The cheat sheet is divided into six parts which have to be followed during the practical implementation of data storytelling: (1) **Facts** give a short introduction to data storytelling, (2) **Structure** explains the theory behind data storytelling, (3) **Process** explains which steps have to be implemented by any user of the guideline to make all your stakeholders (4) **Parties** engaged with your story. To be sure you have not forgotten about any important thing, the (5) **Checklist** helps with additional "5 W questions". After implementing data storytelling the first time, additional Material is presented in (6) **Get inspired**.

Limitations: This paper focuses on data storytelling to communicate complex data to internal stakeholders. It explores data storytelling in general and does not have a specific purpose, e.g. change management, innovation management, etc. Communication to external stakeholders is not part of this research. For external stakeholders, we assume that the data storytelling process in general might slightly differ. This paper applied a qualitative research method to gather primary data. With 13 interviewees, the number of participants is limited. Most of them work in data analytics. Thus, the generalization is limited as well.

Future Research: In addition to the qualitative method, a

quantitative approach to collect empirical data should be used as well. Combining the already available qualitative with quantitative data can help to get further results and different views. Studying how data storytelling with external stakeholders differs from internal stakeholders and comparing these two target groups might also reveal further insights. Identifying the eligibility of data storytelling in specific fields (e.g., change management) or with special purposes could also be interesting research topics. For example, in change management, effective change management depends on all involved parties accepting and understanding change process within an organization.

7. Contribution

The contribution of this paper is threefold. First, the practical contribution of the paper is based on the collection, analysis and presentation of primary data. Expert interviews were conducted to get insights on different aspects of data storytelling. The interviewees are business experts who use data storytelling as a means for communication. Because their main job is not data storytelling, their answers might be closer to other business experts than answers of scientists.

Second, the paper contributes to a theoretical understanding of data storytelling, since literature forms the basis for the empirical research, i.e. the interview guidelines are built on categories derived from literature. From the empirical data the authors identified and derived additional categories to complement the data storytelling theory and concepts. The interviews also revealed questions relevant for data storytellers, which are already tackled in literature. Therefore, scientists should increase efforts to communicate new results to interested audiences in an offensive and suitable way.

Third, the methodology used for this research project is suitable to reach the project's aim. Talking to people who apply data storytelling allows for detailed answers. Interviews open the opportunity to ask further questions based on the answers of the interview partners. The personal opinion and understanding of specific aspects of data storytelling can be identified in detail.

8. Funding

This study received no specific financial support.

9. Competing interests

The authors declare that they have no competing interests.

References

Aaker, J. (2021). Harnessing the Power of Stories. Retrieved November 12, 2022 from http://surl.li/fvvzu.

Al-Doulat, A., Nur, N., Karduni, A., Benedict, A., Al-Hossami, E., Maher, M. L., . . . Niu, X. (2020). Making Sense of Student Success and Risk Through Unsupervised Machine Learning and Interactive Storytelling. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Lecture Notes in Computer Science. Artificial Intelligence in Education* (Vol. 12163, pp. 3–15). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-52237-7_1.

Bladt, J., & Filbin, B. (2013). *A Data Scientist's Real Job: Storytelling*. Harvard Business Review.

Boldosova, V., & Luoto, S. (2019). Storytelling, business analytics and big data interpretation. *Management Research Review, 43*(2), 204–222. https://doi.org/10.1108/MRR-03-2019-0106.

Borkin, M., Bylinskii, Z., Kim, N., Bainbridge, C., Yeh, C., Borkin, D., . . . Oliva, A. (2015). Beyond Memorability: Visualization Recognition and Recall. *IEEE Transactions on Visualization and Computer Graphics, PP*, 1. https://doi.org/10.1109/TVCG.2015.2467732.

Brewster, S., Fitzpatrick, G., Cox, A., & Kostakos, V. (Eds.) (2019). *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems.* New York, NY, USA: ACM: ACM.

Brown, E. D. (2014). Drowning in Data, Starved for Information. Retrieved November 12, 2022 from http://surl.li/fvwef.

Carbonell, J., Sánchez-Esguevillas, A., & Carro, B. (2017). From data analysis to storytelling in scenario building. A semiotic approach to purpose-dependent writing of stories. *Futures, 88*, 15–29. https://doi.org/10.1016/j.futures.2017.03.002.

Chen, C., Härdle, W., & Unwin, A. (2008). *Handbook of Data Visualization*. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-33037-0.

Cohn, N. (2013). Visual narrative structure. *Cognitive science, 37*(3), 413-452. https://doi.org/10.1111/cogs.12016.

Davison, R. M. (2016). The art of storytelling. *Information Systems Journal, 26*(3), 191–194. https://doi.org/10.1111/isj.12105.

Diakopoulos, N. (2018). Ethics in data-driven visual storytelling. In N. H. Riche, C. Hurter, N. Diakopoulos, & S. Carpendale (Eds.), *Data-Driven Storytelling* (pp. 233–248). AK Peters/CRC Press.

Dur, B. I. U. (2012). Analysis of data visualizations in daily newspapers in terms of graphic design. *Procedia, 51*, 278–283. https://doi.org/10.1016/j.sbspro.2012.08.159.

Dykes, B. (2016). Data Storytelling: The Essential Data Science Skill Everyone Needs. Retrieved November 25, 2022 from http://surl.li/fvwac.

Dykes, B. (2020). *Effective Data Storytelling: How to drive change with data, narrative, and visuals.* Hoboken, New Jersey: John Wiley & Sons, Inc.; Wiley.

Ehmel, F., Brüggemann, V., & Dörk, M. (2021). Topography of Violence: Considerations for Ethical and Collaborative Visualization Design. *Computer Graphics Forum, 40*(3), 13–24. https://doi.org/10.1111/cgf.14285.

Elias, M., Aufaure, M. A., & Bezerianos, A. (2013). Storytelling in Visual Analytics Tools for Business Intelligence. *IFIP Conference on Human-Computer Interaction, 8119*, 280–297. https://doi.org/10.1007/978-3-642-40477-1_18.

Feigenbaum, A., & Alamalhodaei, A. (2020). *The data storytelling workbook* (1st ed.). Abingdon Oxon, New York NY: Routledge.

Freytag, G. (1895). *Technique of the drama: An exposition of dramatic composition and art*. S. Griggs.

Friendly, M., & Wainer, H. (2021). *A History of Data Visualization and Graphic Communication* (1st ed.). Cambridge, Massachusetts, London: Harvard University Press.

Gadatsch, A., & Landrock, H. (2017). *Big Data für Entscheider: Entwicklung und Umsetzung datengetriebener Geschäftsmodelle.* Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-17340-1.

Gershon, N., & Page, W. (2001). What storytelling can do for information visualization. *Communications of the ACM, 44*(8), 31–37. https://doi.org/10.1145/381641.381653.

Harris, J. (2007). The whale hunt: Metadata. Retrieved December 20, 2022 from http://thewhalehunt.org/interface.html.

Hartmann, K. (2020). *Digital Marketing Analytics: In theory and in practice* (1st ed.). Independently Published.

Hullman, J. (2011). Visualization Rhetoric: Framing Effects in Narrative Visualization. *IEEE Transactions on Visualization and Computer Graphics.* (12), 2231–2240. https://doi.org/10.1109/TVCG.2011.255.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar.

Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals* (10th ed.). Hoboken, New Jersey: John Wiley & Sons, Inc.

Kosara, R., & Mackinlay, J. (2013). Storytelling: The Next Step for Visualization. *Computer, 46*(5), 44–50. https://doi.org/10.1109/MC.2013.36.

Kreminski, M., Dickinson, M., Mateas, M., & Wardrip-Fruin, N. (2020). Why Are We Like This? The AI Architecture of a Co-Creative Storytelling Game. In G. N. Yannakakis, A. Liapis, P. Kyburz, V. Volz, F. Khosmood, & P. Lopes (Eds.*), International Conference on the Foundations of Digital Games* (pp. 1–4). New York, NY, USA: ACM. https://doi.org/10.1145/3402942.3402953.

Kuckartz, U. (2014). *Qualitative text analysis: A guide to methods, practice & using software* (5th ed.). Los Angeles: SAGE.

Lankow, J., Crooks, R., & Ritchie, J. (2012). *Infographics: The power of visual storytelling* (1st ed.). New Jersey: John Wiley & Sons, Inc.

Lee, B., Riche, N. H., Isenberg, P., & Carpendale, S. (2015). More than telling a story: Transforming data into visually shared stories. *IEEE computer graphics and applications, 35*(5), 84-90. https://doi.org/10.1109/MCG.2015.99.

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology, 75*(1), 5–18. https://psycnet.apa.org/doi/10.1037/0022-3514.75.1.5.

Locoro, A., Cabitza, F., Actis-Grosso, R., & Batini, C. (2017). Static and interactive infographics in daily tasks: A value-in-use and quality of interaction user study. *Computers in Human Behavior, 35*(71), 240–257. https://doi.org/10.1016/j.chb.2017.01.032.

Ma, K. L., Liao, I., Frazier, J., Hauser, H., & Kostis, H. N. (2012). Scientific storytelling using visualization. *IEEE Computer Graphics and Applications, 32*(1), 12–19. https://doi.org/10.1109/MCG.2012.24.

Matzen, L. E., Haass, M. J., Divis, K. M., Wang, Z., & Wilson, A. T. (2018). Abstract Data Visualizations // Data Visualization Saliency Model: A Tool for Evaluating Abstract Data Visualizations. *IEEE Transactions on Visualization and Computer Graphics, 24*(1), 563–573. https://doi.org/10.1109/tvcg.2017.2743939.

Nagel, W. E., & Ludwig, T. (2020). Data Analytics. *Informatik Spektrum, 42*(6), 385–386. https://doi.org/10.1007/s00287-019-01231-9.

Neifer, T., Lawo, D., Bossauer, P., Esau, M., & Jerofejev, A. M. (2020). Data Storytelling als kritischer Erfolgsfaktor von Data Science. *HMD Praxis Der Wirtschaftsinformatik, 57*(5), 1033–1046. https://doi.org/10.1365/s40702-020-00662-3.

Nussbaumer Knaflic, C. (2017). *Storytelling mit Daten: Die Grundlagen der effektiven Kommunikation und Visualisierung mit Daten* (M. Kauschke, Trans.) (1. Auflage). München: Verlag Franz Vahlen. https://doi.org/10.15358/9783800653751.

Obie, H. O., Chua, C., Avazpour, M., Grundy, J., & Bednarz, T. (2019). A study of the effects of narration on comprehension and memorability of visualisations. *Journal of Computer Languages, 52*(1), 113–124. https://doi.org/10.1016/j.cola.2019.04.006.

Polatsek, P., Waldner, M., Viola, I., Kapec, P., & Benesova, W. (2018). Exploring visual attention and saliency modeling for task-based visual analysis. *Computers & Graphics, 72*(12), 26–38. https://doi.org/10.1016/j.cag.2018.01.010.

Roels, R., Baeten, Y., & Signer, B. (2017). Interactive and Narrative Data Visualisation for Presentation-based Knowledge Transfer // Interactive and Narrative Data Visualisation for Presentation-Based Knowledge Transfer. *Communications in Computer and Information Science, 739*, 237–258. https://doi.org/10.1007/978-3-319-63184-4_13.

Ramm, S., Kopf, E.-M., Dinter, B., & Hönigsberg, S. (2021). What Makes a Good Story: The Use and Acceptance of Storytelling in Business Intelligence. *HICSS*. (54), 5706–5715. https://doi.org/10.24251/HICSS.2021.692

Ryan, L. (2016). *The visual Imperative: Creating a visual culture of data discovery* (1st ed.). Amsterdam, Boston: Elsevier MK Morgan Kaufmann.

Sanchez-Lopez, I., Perez-Rodriguez, A., & Fandos-Igado, M. (2020). The explosion of digital storytelling: Creator's perspective and creative processes on new narrative forms. *Heliyon, 6*(9), e04809. https://doi.org/10.1016/j.heliyon.2020.e04809.

Segel, E., & Heer, J. (2011). Narrative Visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics, 16*(6), 1139–1148. https://doi.org/10.1109/TVCG.2010.179.

Showkat, D., & Baumer, E. P. S. (2021). Where Do Stories Come From? Examining the Exploration Process in Investigative Data Journalism. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW2), 1–31. https://doi.org/10.1145/3479534.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039.

Soares de Lima, E., Feijó, B., & Furtado, A. L. (2020). Adaptive storytelling based on personality and preference modeling. *Entertainment Computing, 34*(6), 100342. https://doi.org/10.1016/j.entcom.2020.100342.

Thorne, S. (2020). Hey Siri, tell me a story: Digital storytelling and AI authorship. *Convergence: The International Journal of Research into New Media Technologies, 26*(4), 808–823. https://doi.org/10.1177/1354856520913866.

Tong, C., Roberts, R., Borgo, R., Walton, S., Laramee, R., Wegba, K., . . . Ma, X. (2018). Storytelling and Visualization: An Extended Survey. *Information, 9*(3), 65. https://doi.org/10.3390/info9030065.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440–463. https://doi.org/10.1037/a0020319.

Tufte, E. R. (2001). *The Visual Display of Quantitative Information* (2nd ed.). Connecticut: Graphics Press LLC.

Ware, C. (2020). *Information Visualization: Perception for Design* (4th ed.). Cambridge: Elsevier Inc. https://doi.org/10.1016/B978-0-12-812875-6.01001-X.

Yang, L., Xu, X., Lan, X., Liu, Z., Guo, S., Shi, Y., . . . Cao, N. (2022). A Design Space for Applying the Freytag's Pyramid Structure to Data Stories. *IEEE Transactions on Visualization and Computer Graphics, 28*(1), 922–932. https://doi.org/10.1109/TVCG.2021.3114774.

Yang, P., Luo, F., Chen, P., Li, L., Yin, Z., He, X., & Sun, X. (2019, August). Knowledgeable Storyteller: A Commonsense-Driven Generative Model for Visual Storytelling. In IJCAI (Vol. 3, No. 6, p. 7). https://www.ijcai.org/Proceedings/2019/0744.pdf.

Zhang, Y., Reynolds, M., Lugmayr, A., Damjanov, K., & Hassan, G. M. (2022). A Visual Data Storytelling Framework. *Informatics, 9*(4), 73. https://doi.org/10.3390/informatics9040073.

Appendix 1: Category schema

Categories	Sub Categories	
1	**2**	
1) Working experience with data	Years	
2) Current job	Data Analyst	Data Scientist
	Business Analyst	CRM Analyst
	Data Consultant	Reporting
	Data Engineer	Financial Communication
3) Educational background	Management & Law	Healthcare Management
	Marketing	Business Administration
	Linguistics	Mathematics
	Biostatics	Environmental Sciences
	Psychology	Bioinformatics
	Public Relations	Sustainability Management
4) Career path	Learning by doing	Trial & Error
	Communication	Listening
	Soft skills	Hard skills
5) Work Routine	Projects	Knowing the data
	Feedback	Knowing the stakeholders
	Reports	Knowing the business needs
	Dashboards	
6) Definition of data storytelling	The structured form of data communication	
	Translation of complex data into simple language	
	One key message	A red thread within an analysis
	Marketing & Sales for Data Science	
	Interactivity	Data Journey
	Target orientation	Leading the interpretation
7) Data storytelling first contact	Subconsciously	Modern term
	Missing understanding of Analysis results	
	Education	
8) Recommended Skills	Empathy	Goal-orientation
	Creativity	Open mind
	Innovativeness	Curiosity and self-driven
	Resilience	Team player
	Focus	Communication
	Rhetorical skills	Critical thinking
	Patience	Analytical understanding
	Visual Thinking	Methodical know-how
	Decision Theory	Business understanding
	Learning Theory	Strategic Management
	Drama writing	Behavioural economics
	Supportive company culture	
9) Data visualization	Simplicity	Visual design
10) Explicit data storytelling knowledge	Internal best practices	
11) Data storytelling within the organization	Common sense	Workshops
	PowerPoint	Videos
	Dashboards	Flipchart
	Data Journey	Structural approach
12) Stakeholders	Data literacy	Detail-depth
	Simplicity	Stakeholder needs
13) Ethical data storytelling	Subjectivity	Critical view
	Showing perspective	Awareness
	Transparency	Working scientifically
	Governance guideline	
14) Tools	Content	Complexity
	Resources	Competence
	Methodology	Analysis
	Visualization	Presentation

1	2	
15) Dashboards	Compact	Glossary
	Self-explanatory	Transparency
	Target-orientation	Advantages
	Disadvantages	
16) Data communication	Teamwork	Data story
	Data complexity	Core information
	Creativity	Context
17) Data journalism	5 W's in journalism	Data transformation process
	Depth of elements	Perspective
	Audience	Medium
	Feedback	
18) Future of data storytelling	Complexity of data	Faster Analysis
	Higher flexibility	Better scalability
	Resources	
19) AI-driven data storytelling	Curios	Suspicious
	Uncertainty	Hurdle of acceptance
	Generation topic	AI categorization
20) Personal improvement	Trends and Market	Data communication skills
	Different approaches	Business understanding
21) Data storytelling advice	No single path	Art and Science mind
	Coach	Communication
	Feedback	Listening
	Core insights	Context
	Asking questions	Business knowledge
	Data knowledge	Stakeholder knowledge
	Teamwork	Constant improvement

Source: Research results, 2022