DELIVERABLE 3.3
ParCos Storyteller

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 872500.

“ParCos – Participatory Communication of Science”
A HORIZON 2020 RESEARCH AND INNOVATION ACTION

Consortium: Lappeenrannan-Lahden teknillinen yliopisto (FI, coordinator), Katholieke Universiteit Leuven (BE), Vlaamse Radio- en Televisieomroeporganisatie (BE), and Knowle West Media Centre LBG (UK).

Webpage: https://parcos-project.eu
Duration: 1/2020 – 12/2022
Grant: H2020-872500 (Call H2020-SwafS-2019-1)

Contact (co-ordinator):
Asst. Professor Antti Knutas & Asst. Prof. Annika Wolff
LUT University
e-mail: parcos.project@lut.fi

DESCRIPTION OF THE DELIVERABLE

Overview	**Details**
Authors	Sandy Claes, LUCA (Belgium) Catho Van Den Bosch, VRT (Belgium) Nikki Peeters, VRT (Belgium)
Reviewers	Priscilla Van Even (Belgium) Bieke Zaman (Belgium)
Number of Deliverable	3.3
Title of Deliverable	ParCos Storyteller
License	CC BY 4.0, see https://creativecommons.org/licenses/by/4.0/
Attribution	CC BY 4.0 ParCos, http://parcos-project.eu H2020-872500
Dissemination Level	Public
Contractual delivery date	2021-06-30
To be cited as	Claes, S., Van Den Bosch C., Peeters N. ParCos Storyteller, Deliverable 3.3 of the Horizon 2020 project ParCos, EC grant agreement no 872500, Lappeenranta, Finland.

SUMMARY

Like all stories, science stories follow a narrative structure. In this deliverable, we will map different narrative structures and complement them with science and data story examples. We will detail the role of data in these stories, and how data may support the narrative structure. Then, we present 17 design explorations of data stories that concern 1) weather data sets as a familiar type of data for recipients, or 2) sonicated astronomy data as a data set that is not only able to be represented to recipient through visual aspects but also appeal to other senses. Through these explorations and the state of the art examples, we derived a set of participatory data storytelling techniques. These techniques are opened up for both media professionals and academics through the ParCos Storyteller.

Keywords: Data Storytelling, Science Data Communication, Participatory Data Storytelling

TABLE OF CONTENTS

1 INTRODUCTION

In an increasing data-driven society, storytelling is considered a method, even maybe an anti-dote, for making complex technologies (Giaccardi & Redström, 2020) tangible and more human. Storytelling is the result of human activity, which is difficult for machines to learn to master. This might also explain why storytelling has been increasingly of interest of scholars over the last 10 years in the context of academic fields as human-computer-interaction or human-data-interaction. While a large body of work has mainly focused to expose storytelling techniques for data experts and professionals, traditional storytelling areas as film and video has been focusing on participatory storytelling approaches to let the viewers contribute actively in the story. These approaches have been particularly useful to give underrepresented actors a voice as mainstream media often rely on actors that represent a large demographic range (Manni et al., 2019). Such participatory storytelling approaches, however, also holds much promise for the communication of science data, as it **establishes trust in the combat to the spread of misinformation** (Meijer, 2009).

Typically, **science data is objective** and also often quantitative in nature. Yet to add meaning to this data, authors often enrich it with qualitative data such as anecdotes, perceptions and experiences (Liu et al., 2020). We consider the participatory input in the story construction to be mainly quantitative (e.g. with sensor data) yet also allow the addition of qualitative data (e.g. perceptions, imaginations). Such approach to data representation is in line with timely studies on making data meaningful (Liu et al., 2020).

In this report, we will connect existing data storytelling structures with those of participatory media. First, we will define storytelling in its traditional, interactive and data-oriented form. Then, we will elaborate on the role of participation in storytelling. Third, we explored these narrative structures and techniques in 17 lo-fidelity data stories. Finally, we will introduce the ParCos Storyteller, which entails participatory data storytelling techniques.

2 WHAT IS DATA STORYTELLING?

In this chapter, we will define storytelling and the role of data plays through examples in different media. There is a difference between storytelling as a method in which participants construct stories together as a way to voice their opinion and engage with each other (Manni et al., 2019), versus storytelling as a product that is understandable for users that were not involved in the process. Therefore, as mentioned in D2.1 Bristol Approach to Citizen Science, the ParCos Storyteller is mainly concerned with the fifth step of the Bristol approach, i.e. orchestration, or the presentation of information to other stakeholders to trigger additional input data.

2.1 LINEAR STORYTELLING

In its essence, storytelling is considered as a linear sequence of occurrences. The most traditional narrative structure consists of three-acts, i.e. setup (act 1), confrontation (act 2) and resolution (act 3) (see Figure 1). Storytellers continuously reinvent this structure by shifting the order, adding or removing narrative elements and storylines. A storyline is a narrative thread that can be experienced by one character in the story or a subset of characters. Often, there are several storylines in one story that all contribute to the main narrative goal. A typical example is the parallel editing or cross-cut in which you see two scenes that are intertwined, such as 1) the crook and his/her plans to derail a train and 2) innocent passengers that enjoy their train ride.

The crook versus innocent passengers is also an example of a conflict. Without conflicts, there is no story. This basic rule is known by all storytellers, and can take many forms. In data stories, conflicts can be identified as the assumptions that exist in the viewers’ mind versus the objective data that might reveal differently, in outliers or in causal relationships, to name a few. Here, characters are thus data sets or data points, and to ‘tie’ the characters emotionally to the viewer, there needs to be a personal link.

A conflict can be the actual story or one of the supporting storylines. In Figure 1, each climax is a conflict. Assume this story is: Central Europe will get warmer in the future. The main conflict lies in the belief of the viewer: whether he or she thinks it is true, and why. A first supporting conflict may then be found in the historical data and how this supports the main narrative goal or not. Then, another supporting conflict may be focused on the outliers and how they increased throughout the centuries. As you may notice, these conflicts are not as straightforward as the crook versus the innocent passengers. Instead, conflicts are implicitly connected to the knowledge of the viewer.

Figure 1. Traditional three-act story structure: 1) setup, 2) confrontation and 3) resolution.

The linear narrative structure is the backbone for stories in different media (e.g., novels, films, presentations) and for different purposes (e.g., entertainment, informing, inspiring). Examples of data stories that follow this structure include infographic films, explainers, news articles with information graphics, etc. Here, the role of the viewer is to consume the

story that is presented to her. Data representation is used to support the story. The interpretation of the data is made by the storyteller.

2.2 INTERACTIVE STORYTELLING

With the rise of interactive technologies since the 1960’s, authors have adapted the existing narrative structures to this digital realm. In interactive storytelling, **two main structures exist:** 1) interactive narrative structures, and 2) open world.

The first consists of different storylines that are interconnected (see Figure 2) and that allow the viewer to make choices to progress the story. Different forms and shapes of such interactive narrative exist. Yet as our point is to explain how this differs from linear storytelling and this structure is most reminiscent of the three-act structure, we decided to use the branched narrative structure as an example.

This interactive narrative structure is most known as different media – even those that are typically not interactive, such as television – adopted this structure. Indeed, television shows today, such as Dora the Explorer and Bear Grills for children but also the Bandersnatch episode of Black Mirror for adults, brought branched narratives to the TV screen with a simple press on the remote-control button. Yet a branched narrative structure is most known by its application in video games.

Figure 2. Branched narrative structure.

Second, an open world narrative structure is a model that is less tangible for those that are familiar with traditional linear storytelling. Here, the story discovery is literally left to the user with little to no steering of the author towards a particular order in storylines (see Figure 3). This structure is often used in game design that facilitate role playing (RPG), and virtual reality (VR), augmented reality (AR) or mixed reality (XR) experiences.

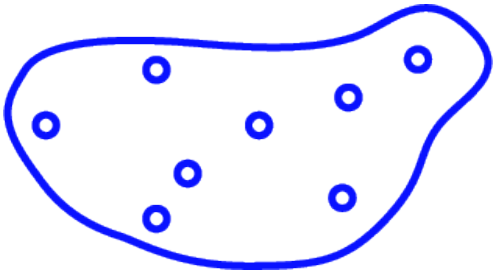


Figure 3. Open world structure.

2.3 What is Interactive Storytelling with Data?

For data stories specifically, the publication of the Narrative Visualization framework (Segel & Heer, 2010) caused an enormous amount of attention for the deployment of narrative techniques to communicate to and engage a large audience with data. The techniques provided by this framework are considered as the basis for our ParCos Storyteller, and therefore we will expose what it entails. It positions storytelling on a spectrum between being author-driven from the one end (i.e. the media professional decides what the viewer sees and when) and reader-driven on the other end (i.e. the user may explore freely). On this spectrum, there are three narrative structures:

1) Martini Glass Structure

Figure 4. Martini glass data story structure.

Figure 4 depicts the narrative structure that resembles the traditional three-act story structure most. It starts with an ‘open’ interface in which the user may interact with settings or data, followed by a linear narrative (that gradually builds up towards the resolution) and ends with an ‘open’ exploration of data points.

2) Interactive Slideshow

Figure 5. Slideshow data story structure.

Figure 5 demonstrates a narrative structure that sequentially offers linear storylines, switched with interaction points that allow the user to decide on the pace. The advantage of this structure is that allows the storyteller to set out the different storylines one by one and as such create a clear overview for complex topics.

3) Drill-Down Story

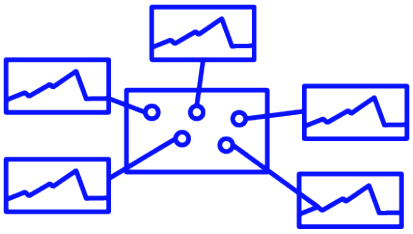


Figure 6. Drill-down story structure.

Figure 6 shows the most ‘open’ structure. Here, the user may freely decide which storyline he or she explores first. The storyteller’s main task is to bring storylines forward that are coherent when connected without chronological order, which is often not evident.

2.4 WHY INTERACTIVITY CONTRIBUTES TO DATA ENGAGEMENT

One of the key reasons why interactive data storytelling is able to engage a wide variety of audiences with data is the ability to find ‘yourself’ in the data. Indeed, **finding personal connection points with the data story** is an engaging factor that is often missing in linear data stories. In traditional storytelling, the author would create a character (fictional or not) with whom we may identify ourselves, and thus find personal links with. However, for data stories this is not always possible to set up. An example that balances this traditional storytelling with personalization through data is Brooke Leave Home (Concannon et al., 2020), see also D2.3. Brooke is the main character in an online film. As she turns 18, she leaves a care institution in the UK and the viewer follows her on her journey onwards. As a viewer, you gain empathy for Brooke as the video shows the girl in close-ups as well as her home, techniques that are typically used by media professionals. Yet Brooke Leave Home is a project that aims their viewers to engage with data. Here, data is used to personalize the film to the actual location you are situated. What Brooke experiences, in other words, can happen next to your own door.

3 WHY PARTICIPATION IN STORYTELLING MATTERS

In storytelling, participation can be interpreted in several manners. Often, participation is understood by professionals as ‘users that generate content for them’ (Varghese et al., 2020). However, true participation in the storytelling process occurs when those users are also enabled to take an active role in the construction of the story and thus have the ability to actively collaborate (Varghese et al., 2020). At the very least, **users may contribute data,** such as media they generated, e.g. videos, photographs, audio and texts. A well-known example in popular media are the photographs of weather phenomena that the weather(wo)man uses in the weather forecast (or news item) to exemplify the current condition (see Figure 7 for an example).

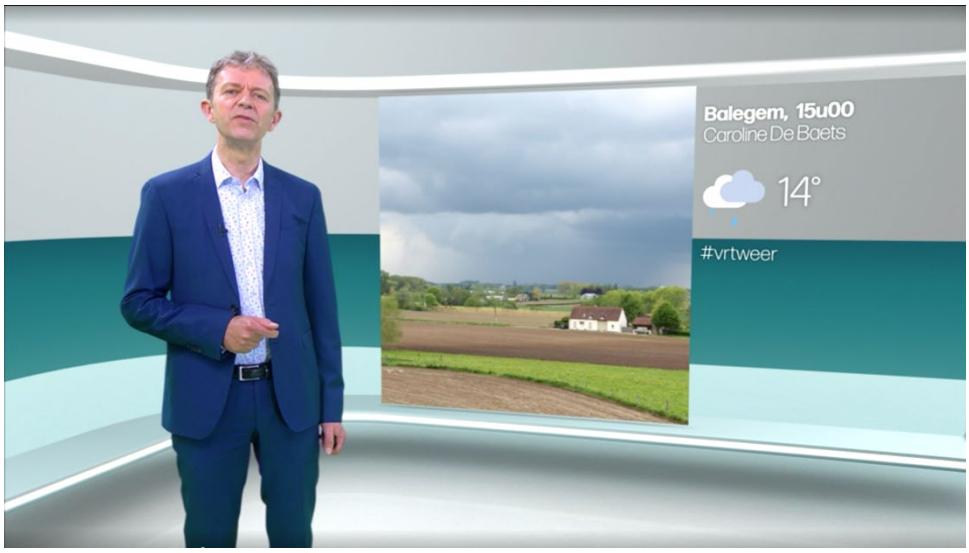


Figure 7. Weatherman Frank Deboosere of VRT discusses the weather of the day of recording (May 20, 2021) with a photo of viewer Caroline De Baets in the background.

Another data contribution by users can take the form of citizen science data, e.g. sensors that citizens installed at their home or office and which results they connect to an online platform where a media professional may base a story on. An example is the project Curieuzeneuzen (‘nosey parkers’) in Flanders, in which 20.000 Flemish citizens participated to measure the local air quality and journalists of newspaper de Standaard regularly publish stories based on this information (See https://curieuzeneuzen.be/).

3.1 COMBATTING MISINFORMATION THROUGH POLYVOCALITY

More contribution by the user may be triggered when the story format allows the **user to explore potential storylines and also contribute insights on the envisioned story**. Such participatory media examples most often occur in social or health contexts in the

documentary genre. Complex problems like mental health benefit from the multi-angle view on the issue that is facilitated through the participatory process (Manni et al., 2019). Here, storytelling is both used as a method and outcome. In community building, the empowering potential of participatory storytelling has been widely recognized (e.g. Lambert, 2006; Bromley, 2010), including the Council of Europe (Lewis, 2008). Also, media professionals, such as those active at public service media, understood it as a tool for democratizing media (Meijer, 2013). BBC, for instance, launched their project Capture Wales in 2003 (Maedows, 2003), and the regional public broadcaster of Utrecht launched a weekly TV show in September 2009, which is still in production today. This project ‘U in de Wijk’ (You in the Neighbourhood) was initially meant to counter the negative impact of journalists that overexposed the problems in that area (https://uindewijk.com/). In its ability **to allow different viewpoints on one topic,** participatory storytelling is a way to overcome such negative framing (Meijer, 2013). However, journalists are not keen to involve others in the composition of their story, which contributes to the sense of misinterpretation of one’s surroundings and **perception of misinformation** (Meijer, 2013).

The storytelling that is used in interactive documentaries (iDocs) is related to that of participatory filmmaking as the interactivity is intended as an array of choices offered to users that are enabled to exercise control over the materials presented in the documentary (Manni et al., 2019). Interactivity can serve a number of functions within the iDocs text: finding information (either within or beyond the documentary), learning, furthering the narrative, personalizing the documentary, adding to the documentary content, play or search “playfully” for hotspots within an image interface (Nash, 2012). In iDocs, non-linearity is perceived as an opportunity that “allows audiovisual projects to provide elements to complement and enrich it, providing several added values to the global experience of the audience” (Castells, 2011, p354-3). IDocs thus embrace a higher level of complexity that **connects to the polyvocality** of participatory filmmaking (Manni et al., 2019).

Media professionals are thus increasingly involving their audience in storytelling. Software tools as Storymaker allow those professionals to process the incoming data that come – in contrast to content produced by experts – in large amounts and in messy structures (Claes et al., 2020). Media professionals act as facilitators and may encourage users to express themselves creatively (Manni et al., 2019). As such, just as **the role of the media professional shifted to this facilitating role** when participatory approaches entered the profession of interface, service and product design in the beginning of the 21st century. Today, the role of professionals in the media domain is broadened to include facilitation, similar to the role of game designers and interactive storytellers that were already accustomed to think of user input.

In Figure 8, the uncertainty of how the data input of the user may affect the story that was envisioned by the author is depicted as interaction points with dotted lines (which differs from the model of interactive data story structures in the middle of the figure). The figure of the participatory data structure on the right of Figure 8 illustrates how participatory data storytelling links to the ‘open world’ story structure that was discussed earlier.

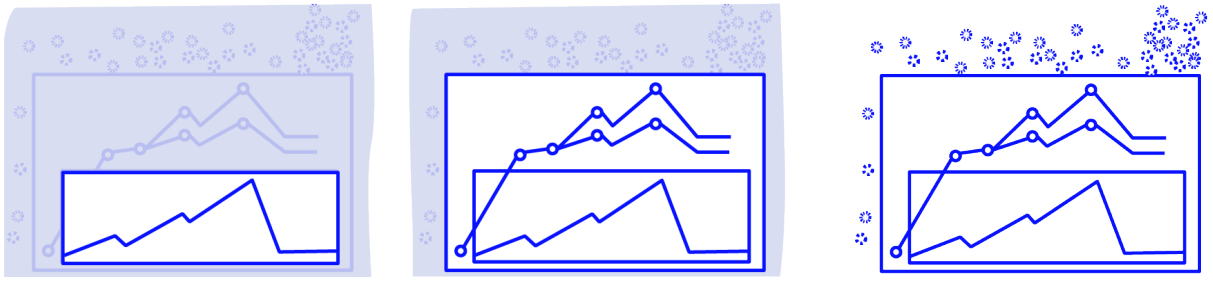


Figure 8. Linear data story structure (left), interactive data story structure (middle) and participatory data structure (right).

Although participatory storytelling is a timely research topic that is also increasingly adopted in practice, its application in data storytelling is still largely unexplored. In the remainder of this report, we will make the concept of participatory data storytelling tangible through examples and practical guidelines.

3.2 PARTICIPATION IN NON-NARRATIVE DATA REPRESENTATIONS

In data visualisation (as it is the main form of representing data), scholars have explored participatory approaches. In interactive data visualisations, participation may take the form of writing text data in a wordle format (Viegas and Wattenberg, 2009) or contributing DIY tracked sensor data on a map (Liu et al., 2020). Also, data visualisation workshops where participants meet face to face to discuss the interpretation of a particular data set (e.g. ParCos workshop at C&T conference in June 2021) can be considered as participatory data representation. Here, data is used as a boundary object to facilitate dialogues about the collective interpretation of that data.

4 WHAT IS PARTICIPATORY DATA STORYTELLING?

Figure 9 presents the different levels of participation in data storytelling, ranging from no participation (consuming) to low participation (actively interpreting data to find new insights) over medium participation (interpreting data to forward the storyline or contributing data to form a new storyline) to actively participating by contributing stories.

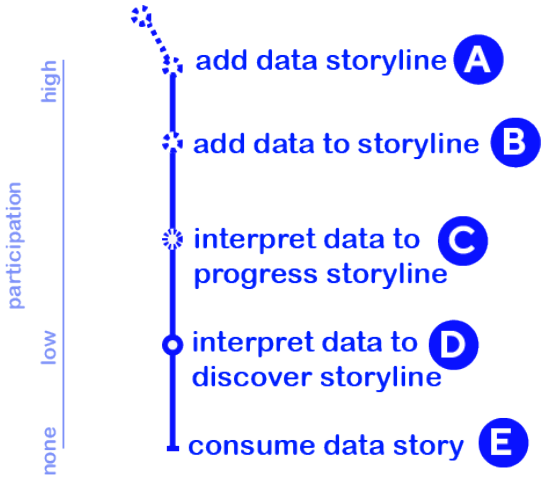


Figure 9. Level of participation in storytelling

4.1 EXAMPLES OF PARTICIPATORY DATA STORYTELLING

4.1.1 Add data storyline

Examples in this category are mostly derived from research projects, both in academic as governmental contexts. Vital Village, for instance, presents 2 examples on their website. The clearest story is about Food Access and combines **scrollytelling** (i.e. a website that you can scroll through to discover the story gradually) **with a map** where participants may upload their own data or insights about data. The latter may feed novel storylines within the larger story (see https://www.vitalvillage.org/data-dashboard/food-access). The narrative structure of this example resembles of the Martini Glass, with the difference that the end now also allows for input and not only exploratory interaction. Another example depicts student transportation equities (see https://www.vitalvillage.org/data-dashboard/customize/student-transportation-equity-map). Here, we are immediately dropped in a map where we, as participants, may add photos and annotations to it and as such add additional storylines. Such ‘photovoice’ is a well-known participatory method to reflect on reality. This story is based on the drill-down structure.

Besides the stories that reside online, also **offline forms of participatory data storytelling** exist. Often such participation is triggered in workshop settings, where different stakeholders are brought together to reflect on a topic through data and their stories. Such

stories may be elicited through existing maps that are augmented by participants with cognitive maps that are drawn with pens to reflect the intensity of perceptions (Liu et al., 2019). Storyboards or scripts, i.e. tools that are typically used in traditional video and film making, are another way to trigger reflection and participation in data storytelling (Wang et al., 2019), while theatre, enactment and play are a more embodied way to experience data and participate in the data story (see D3.1 Guidebook on the use of arts-based methods).

4.1.2 Add data to storyline

We already presented the example of Curieuzeneuzen (see https://curieuzeneuzen.be/) as a cross-over between **citizen science and journalistic storytelling**, which is similar to the project Tuinlab (see https://mijntuinlab.be/), both in Belgium. Worldwide, Zooniverse is the most well-known example (see https://www.zooniverse.org/) with 2.000.000 citizen scientists contributing data and regular data story updates in various news media. Besides the citizen science stories, the popular website pudding.cool often experiments with eliciting responses of their users to feed a data story, e.g. see : https://pudding.cool/2020/07/song-decay/. In this example, users are asked whether they recognize 90’s songs or not, and with these results they presented a story on the popularity of these songs by users that are younger than the song’s origin date. Thus, users that engaged to deliver input were also curious to learn more about the results in a later stage, similar to the citizen science examples.

However, these examples are most often limited to written press that use the contributed data to base their stories on. As a consequence, there is no direct and dynamic relationship between a change in data and the impact on the story(line). Stories are published when it is newsworthy and depending on the rhythm the press envisioned beforehand (e.g. a weekly item, a monthly spread). The underlying storytelling process (e.g. detecting interesting data) is often black boxed. Examples in which the data stories adapt more dynamically when changes in data are made are less common; the digital platform Coronavirus now is one of the exceptions (see https://www.brig.ht/project/comment-representer-les-data-liees-au-coronavirus-dans-une-creation-digitale-en-3d).

Scholars have studied several **workshop tools and methods** to facilitate participants to contribute data, ranging from personal informatics to household log (Elsden et al., 2015). While such workshop settings are closely guided by the moderator, artistic examples as **data sculptures or installations** can take a looser approach in eliciting data from its users. Domestic Data Streamers are famous for their participatory data story installations (see https://domesticstreamers.com/). Here, data participation often takes the form of contributing opinions.

4.1.3 Interpret data to progress storyline

Another way to contribute to citizen science projects – and the stories that evolve from it, is to interpret data that scientists capture. AstroSounds, for example, translates vibrations of space to sound and asks citizens to listen, interpret and discover novel stars (see https://astrosounds.be/). The story, then, is a result of the participation of citizens in interpreting data and being aware for interesting data storylines (e.g., by finding and reporting outliers).

4.1.4 Interpret data to discover storyline

At the very least, the interpretation of data to discover a storyline that was not made explicit by the authors is also considered as participation. In other words, the active discovery of data insights or storylines by the user is a form of participating in the interpretation of data, which contributes to data literacy skills.

Facilitating users to interpret data is not always evident. Data is often connotated to be difficult, too large-scaled, too expert-driven, etc. As a result, scholars have explored ways to present data in more engaging ways, including data comics (Bach et al., 2015), public interventions (Claes et al., 2013) and video (Amini et al., 2015). Recently, scholars studied the impact of data-driven personalised films as a way to familiarise viewers with data concepts (Concannon et al., 2020). Here, users may first consume the story that is tied to their personal living area and afterwards explore how the story changes when moving to other areas.

4.1.5 Consume data story

When data stories are created for consumption, it does not require participation of the user. Yet we consider this level in our model as it is the entry point for users to get acquainted with data stories. Here, examples range from infographic films to supporting charts in a news article. This level also includes science stories that are not literally incorporating data representations in their story yet allow the user to consult their data source.

5 How to Engage Users in Participatory Data Storytelling?

We already defined what data storytelling is, and what the different levels of participation entail. However, participatory design is an approach that demands to actively include all stakeholders (Van der Velden et al., 2014). Therefore, it is not sufficient to bring story structures forward to involve the audience in participatory ways, also the story format should be designed to reach a broad and varied audience. In other words, reaching humans to engage them with the story and become active users is also of our concern.

5.1 Reaching a wide audience in (semi-)public space

Science communication is often oriented at a specific target group of people that are already interested. Often this communication is unlocked at locations that are consulted by that specific target group, such as specialized websites, particular subsections at news websites, TV shows, science musea and libraries. In the participatory mindset, we aim to reach a broader and more varied audience. Presenting data or information in public environments forms one promising way to reach a diverse audience - including humans that otherwise might not be interested. Indeed, (semi-)public places are consulted by various persons that are not particularly looking for information yet stumble upon it in an opportunistic way (Claes and Vande Moere, 2013).

5.2 Facilitating participation

Participatory design strategies such as the Bristol approach (see D2.1), aim to provide tools to a wide audience to tackle civic challenges. Gathering data, making sense of the issue and presenting the issue to others are part of this approach, next to the identification, design and deployment of solutions. As already mentioned previously in this report, such guided participatory workshops form a synchronous way to collaborate in one physical space, which is typically missing from their online counterparts.

5.3 Representing data

When dealing with representing data, we tend to focus on making data visual, such as illustrated by the terms ‘data visualisation’, ‘information visualisation’ or ‘infographics’, thereby excluding those that have a visual disability. Therefore, in our Storyteller, we aim to include ways to represent data through a wide range of senses, including:

The use of audio to represent data, or, in other words *sonification* (Smith et al., 2002). The AstroSounds project that was discussed earlier (as a PDS example) presents data through audio, while Stellar Music no1 is an acoustic composition that makes this audio data tangible through music (Kolláth, 1997).

The use of haptic qualities, such as tangible data representations (Spindler et al., 2010) and data physicalizations (Jansen et al., 2015). The addition of touch to the visual senses supports the data sensemaking process. Data objects are a promising way to make data familiar to the audience at large.

Another way to present data to audiences is via **theatre and play.** Data dramatization is already referred to as an artistic method to facilitate interpretation data in D3.2. In addition, such choreography is also a way to confront audiences with data in new ways, see e.g. https://memoakten.medium.com/data-dramatization-fe04a57530e4

6 PARTICIPATORY DATA STORIES – PARCOS EXPLORATIONS

To explore these data storytelling techniques in various forms, we defined two main data storylines (i.e. Vibration data of stars and historical weather data). These served as starting points for the third-year Bachelor students in Media & Information Design of LUCA School of Arts in Belgium to explore data stories in different media, including online platform, video, podcasts, interactive installations, participatory design methods, data physicalizations, data objects. These explorations were guided by Sandy Claes in the atelier data stories from October till December 2020.

6.1 THE DATA

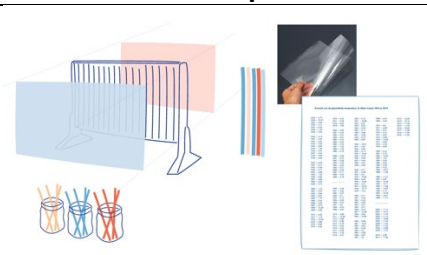
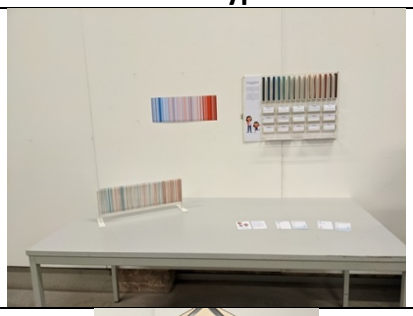
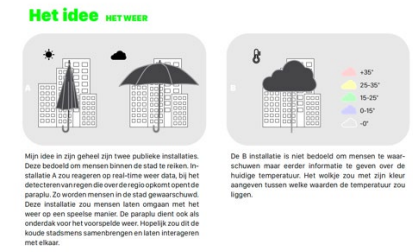
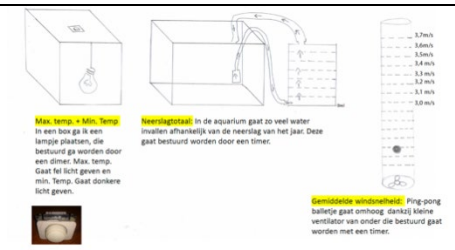
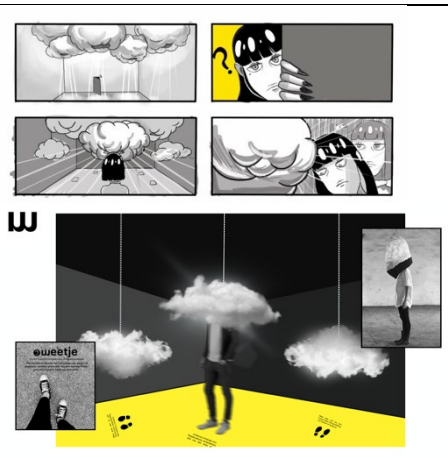
Weather data. This data is already omnipresent in our daily lives; in the form of weather forecasts that are broadcasted on traditional media (radio and tv) and via internet. Moreover, the data have a direct impact on our lives; which mode of transport do I choose? What clothes should I wear?,... However, few people are triggered to explore the data behind those forecasts. We believe that if we give the audience the tools to explore the data via recognizable media, they would be triggered to explore the data for themselves. This would enable them to draw their own conclusions, for example on climate change. Flemish public service media and ParCos partner VRT, and in particular Catho Van Den Bosch and Nikki Peeters, served as professional sounding board for the students.

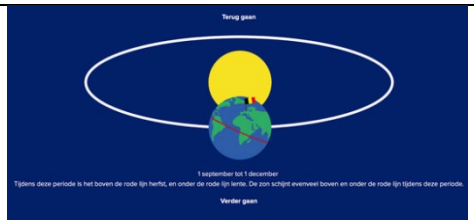
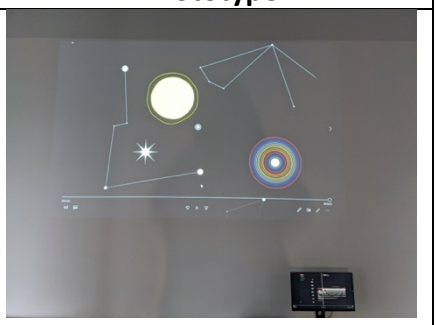
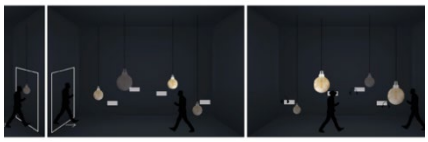
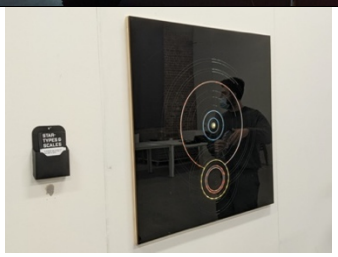
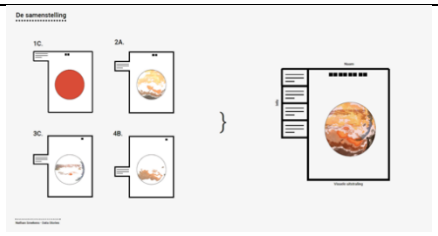
Astronomy data. Unlike weather data astronomical data are far from everyday life. Although space as a topic sparks the imagination of a wide audience, specific astronomical data are difficult to interpret because of a lack of context and the large, ungraspable scale. As a result, the reach of this science data remains limited to people with an interest and deeper understanding of science. We chose this topic because of the challenge to deliver abstract data and subjects to a broad audience. The data that was offered to students originated from the research practice of prof. Katrien Kolenberg (KU Leuven and uAntwerpen). She and her team capture vibrations that are present in space and translate this to sound. Via this sound, it might be easier to discover stars that are not known yet than when visually analysing the vibration data. Besides this data, prof. Kolenberg also presented her science story about the stars and space, which she often brings in Tedxtalk formats (e.g. see https://www.ted.com/talks/katrien_kolenberg_echoes_from_the_cosmos).

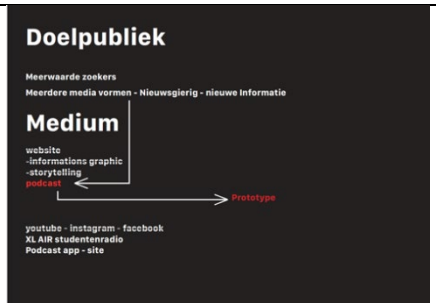
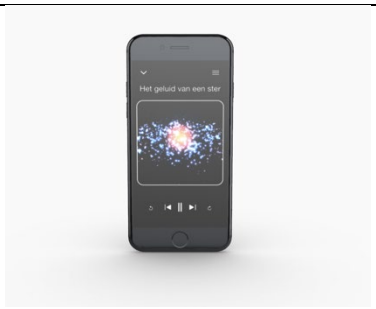
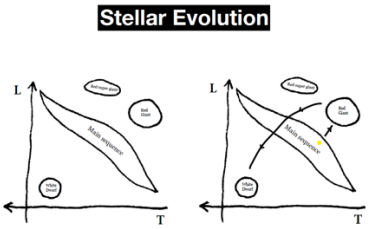
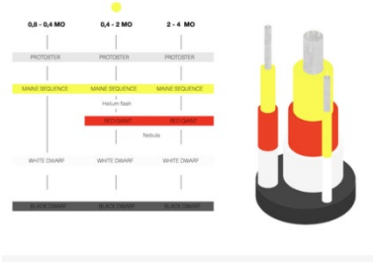
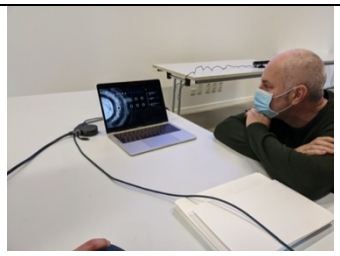
Students were allowed to freely chose the data or science stories that are described above. After 3 weeks, they presented their concepts to VRT and prof. Kolenberg, after which they had another 3 weeks to design and develop their stories to be presented again to VRT and prof. Kolenberg. After that intensive period, students had another 3 weeks to finetune their design.

6.2 DIFFERENT DATA STORIES

In the following, we will give an overview of the outcomes of these explorations and discuss how they provided tangible representations of participatory science communication for other expert stakeholders involved.

Weather Data Stories		
	Concept	**Prototype**
Evolution of the average temperature		
Emotions and weather		
Installation about the seasons		
Thundercloud		

Personalised Weather Forecast		
	Concept	**Prototype**
Music made by stars		
Virtual Exhibition of the Planets		
Installation to illustrate dimensions in the universe		
Walk to illustrate dimensions in the universe		
Composition of the planets		

Children’s book		
Podcast		
Candles that illustrate the Stellar Evolution		
AstroSounds Platform		
The Sounds that the Stars make (educational game)		
Augmented Reality App of the Stars		

Augmented Reality App of the Planets		

7 THE PARCOS STORYTELLER

The ParCoS storyteller provides professionals (e.g., in media, in public institutions) practical insights in how to tell stories with data, and how participation of viewers can be elicited. To serve this goal we (i) explored the needs of our target audience, i.e. the storytellers (see 7.1) and (ii) distilled actionable techniques that are made tangible through examples that are presented an appropriate accessible format (see 7.2).

7.1 EXPLORATORY USER STUDY

A first version of the ParCos Storyteller included examples of data stories that concerned the Covid-19 pandemic (see annex 1.) that illustrated the interactive data storytelling techniques. The aim of this presentation was to discover the needs of media professionals and discover their perspective on their science storytelling practice. The presentation was given to media professionals, 1) internal within VRT and 2) outside VRT.

1) VRT media professionals

The initial presentation was shown in individual meetings to several media professionals within VRT who have expertise in the field of data, as data analyst, data storyteller or innovation specialist. The purpose of this was twofold one the one hand we would inspire them with state of the art examples of data stories and data visualisations. One the other hand we would like to use the presentation as a conversation starter, to discuss their viewpoint on the future role of data stories within a public broadcaster. These discussions have helped us to map the existing expertise within other departments of VRT, for example we received a presentation from the news department about manipulating data. Because of these meetings, we also managed to map out their need and expectations, this will help us to find support in the future if we want to implement the results of the ParCos project in our organisation.

2) Other media professionals

An updated version of the presentation was given at the special interest group on Artificial Intelligence and Data Initiative of the European Broadcasting Union on February 18, 2021. This initiative was attended by data scientists and journalists of different European media organisations, including NPO (The Netherlands), RAI (Italy), France 1 (France), etc. After the presentation, the audience was allowed to give feedback. Here, it was noticeable how different organisations (the three named earlier) struggle with making data engaging. It was also noticeable how these organisations work in a transversal way: data analysts that study content consumption data (e.g. click rates, viewing time) also help journalists to analyse science data.

After the meeting, two media organisations sought contact with us, i.e. RAI (Italy) and RTBF (Belgium, Walloon region). We talked to them in an exploratory meeting in which they presented their existing and upcoming data visualisation work. In both meetings, it was clear the data analysts contribute significantly in the storytelling process of the journalists; they ‘dig’ in the data, search for insights, search for outliers that might form a compelling story and they deliver the right visualisation format.

As a take away of this exploratory user study, we decided the ParCos Storyteller not only to focus on data and science journalists but also open up to data analysts and data scientists as they often collaborate to make stories.

3) Other science communicators

In preparation of the Belgian case study (see D5.1), we also worked together with the storytellers that are involved in the setup of the science festival KNAL (KunstLeuven), the more weather exposition (VRT brand extensions and consultation agency) and the development of citizen science platform AstroSounds (prof. Kolenberg). We do not include these types of storytellers in the target audience for our Storyteller as they are not experts in media. However, the additional perspectives from these professionals allowed us to open up and challenge the science data design space for media professionals with examples that are tied less to screens.

7.2 Data Storytelling Techniques for Participatory Science Stories

Based on the insights of the SOTA, our design explorations and exploratory user study (described above), we defined five steps to compose a data story that can engage a broad audience in the participation of data and science stories. These steps are the basis of our ParCos Storyteller, which can be found in Annex 1.

Step 1. Finding data

Media professionals may consider – next to their traditional sources – to request data from their users, viewers or readers. When those users contribute data, they are already connected to the story the professional envisions.

In order to combat misinformation, storytellers should be as transparent as possible about the data sources they used and the methods they deployed to make sense of the data. In the Storyteller, a good practice is exemplified.

Step 2. Finding the data story

There are two ways: 1) the story ‘pops’ up when analysing data, which may take a long time, and there is a risk that no interesting story will be found; and 2) the data supports a story the professional already envisioned. For the latter, participation of the user may be requested to submit anecdotes, photos or other data that may support or offer new perspectives on the main story.

Step 3. Telling the data story

When interest in the topic at hand is triggered – whether online or offline, the data should follow a narrative structure as this is a way to sustain this interest, whether this occurs in a linear, interactive, or participatory fashion. The techniques that complement these fashions are presented earlier, in chapter 2 of this deliverable. Storytellers may use these structures and examples as a benchmark to cross-check the flow of their data story.

Step 4. Presenting the data story

Data is most often represented through visual means, but also audio or physical qualities may be used to represent data. In fact, complementing traditional data visualisation formats (such as bar charts, line graphs) with representations through other senses, is a successful way to engage a diverse set of audiences.

Step 5. Involving the audience

Participation in data storytelling can take many forms, as illustrated in Figure 8. Storytellers should encourage their viewers to partake in the story process, as a way to build data literacy skills. Viewers may be encouraged to interpret data for themselves, contribute data or help with interpreting data stories. Ultimately, participation offers new perspectives on a (data) story, which results in unforeseen storylines. Storytellers should therefore be aware

for additional story angles, for which they might draw on interactive narrative structures as ‘open world’.

We aim to involve an audience that is perhaps not immediately triggered by science stories (i.e. do not consciously search for it). There are several ways to reach such audience. One way is to enable humans **to stumble upon the stories,** e.g. when reading another news article and noticing a science story is also highlighted in online spaces, or when they do not expect to encounter such stories at offline, public locations, e.g. at the bus stop or in the public library. Storytellers may consider such offline ways to present their story when they want to engage an inclusive audience.

7.3 THE PARCOS STORYTELLER IN PRACTICE

The ParCos Storyteller aims to open up the knowledge we derived from this project. Therefore, we will submit this work to academic venues in Human-Data-Interaction to reach an audience of academics, experts and practitioners. The Storyteller will be extended with our ParCos case studies, and us such we will continue to rework the Storyteller and the presentation that is attached to this deliverable as an accessible document for media professionals. We will promote via professional media-oriented venues, such as the Artificial Intelligence and Data Initiative (AIDA) group of the European Broadcasting Union (EBU).

8 CONCLUSION

In this deliverable, we presented an overview of relevant state-of-the-art practices in data storytelling and the involvement of audience in this through participatory approaches. We exposed the narrative structures behind participatory data storytelling and illustrated this by diverse examples. We also introduced a new model for participatory data storytelling, which forms the heart of the ParCos Storyteller, i.e. a presentation with existing examples that can be shared with and by media professionals. We presented a first version of the ParCos Storyteller to Media and Information Design students who used it as a framework to explore data story designs on astronomy or weather data. In addition, we presented this version to media professionals to detract their existing needs and wishes. Then, we refined the ParCos Storyteller, and we will continue to refine it with our ParCos case study insights and other relevant findings. The ParCos Storyteller will be integrated in the ParCos platform.

9 REFERENCES

Amini, F., Henry Riche, N., Lee, B., Hurter, C., & Irani, P. (2015, April). Understanding data videos: Looking at narrative visualization through the cinematography lens. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1459-1468).

Bach, B., Wang, Z., Farinella, M., Murray-Rust, D., & Henry Riche, N. (2018, April). Design patterns for data comics. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 1-12). Bromley, Roger. 2010. Storying Community: re-imagining regional identities through public cultural activity. *European Journal of Cultural Studies*, 13(1): 9–25. doi:10.1177/1367549409352546

Castells, A. G. (2011). The interactive documentary. Definition proposal and basic features of the new emerging genre. In *Mcluhan Galaxy Conference, Conference Proceedings* (pp. 367-378).

Claes, S., Bauwens, R., Baert, K., & Libot, C. (2020). The Storymaker. Supporting Storytelling with User-Generated Videos through Object Recognition. *https://www. ibc. org/technical-papers/the-storymaker-supporting-storytelling-with-user-generated-videos-through-object-recognition/6761. article*.

Claes, S., & Vande Moere, A. (2013, June). Street infographics: raising awareness of local issues through a situated urban visualization. In *Proceedings of the 2nd ACM International Symposium on Pervasive Displays* (pp. 133-138).

Concannon, S., Rajan, N., Shah, P., Smith, D., Ursu, M., & Hook, J. (2020). Brooke leave home: Designing a personalized film to support public engagement with open data. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-14).

Elsden, C., Kirk, D., Selby, M., & Speed, C. (2015, April). Beyond personal informatics: Designing for experiences with data. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2341-2344).

Giaccardi, E., & Redström, J. (2020). Technology and more-than-human design. *Design Issues*, *36*(4), 33-44.

Jansen, Y., Dragicevic, P., Isenberg, P., Alexander, J., Karnik, A., Kildal, J., ... & Hornbæk, K. (2015, April). Opportunities and challenges for data physicalization. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3227-3236).

Kolláth Z, Keuler JO. Stellar acoustics as input for music composition. *Musicae Scientiae*. 2006;10(1_suppl):161-183. doi:10.1177/1029864906010001081

Lambert , Joe 2006 [2002] *Digital Storytelling: capturing lives, creating community* , Berkeley , CA : Digital Diner Press.

Lewis, Peter M. 2008. *Promoting Social Cohesion. The role of community media*, Strasbourg: Media and Information Division, Directorate General of Human Rights and Legal Affairs, Council of Europe.

Manni, S., Ursu, M., & Hook, J. (2019, June). Stepping Through Remixed: Exploring the Limits of Linear Video in a Participatory Mental Health Film. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video* (pp. 83-94).

Meadows D.C. (2003) ‘‘Digital Storytelling: research-based practice in new media’’, Visual Communication 2(2), pp. 18993.

Meijer, I.C. (2013) WHEN NEWS HURTS, Journalism Studies, 14:1, 13-8, DOI: 10.1080/1461670X.2012.662398

Nash, K. (2012). Modes of interactivity: analysing the webdoc. *media, culture & society*, *34*(2), 195-210.

Smith, D. R., & Walker, B. N. (2002). Tick-marks, axes, and labels: The effects of adding context to auditory graphs. Georgia In- stitute of Technology

Spindler, M., Tominski, C., Schumann, H., & Dachselt, R. (2010, November). Tangible views for information visualization. In *ACM International Conference on Interactive Tabletops and Surfaces* (pp. 157-166).

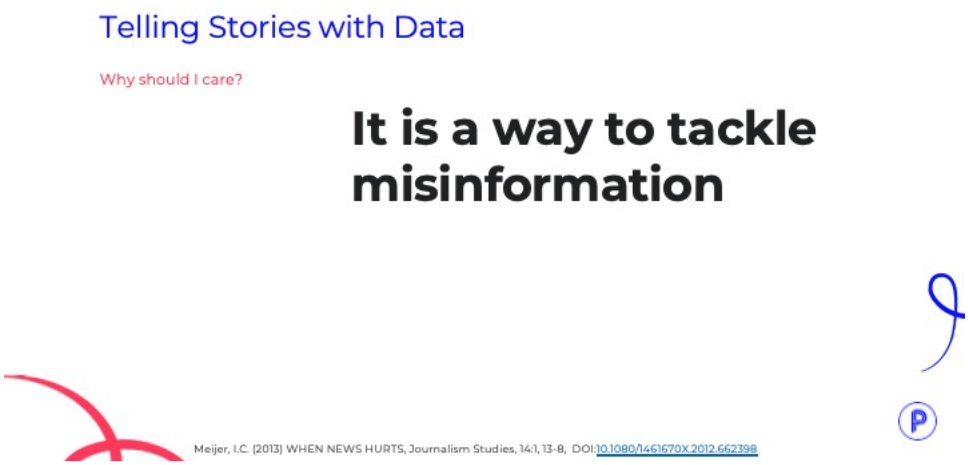
Van der Velden, M., Mörtberg, C., Van den Hoven, J., Vermaas, P. E., & Van de Poel, I. (2014). Participatory design and design for values. *Development*, *11*(3), 215-236.

Varghese, D., Olivier, P., Bartindale, T., & Baillie Smith, M. (2020, April). Towards Participatory Video 2.0. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).

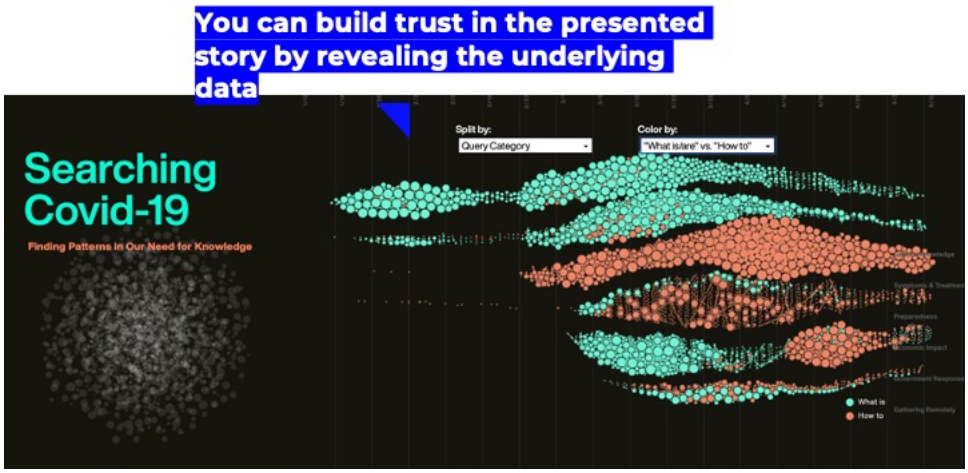
Viegas, F. B., Wattenberg, M., & Feinberg, J. (2009). Participatory visualization with wordle. *IEEE transactions on visualization and computer graphics*, *15*(6), 1137-1144.

Wang, Z., Dingwall, H., & Bach, B. (2019). Teaching data visualization and storytelling with data comic workshops. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-9).

ANNEX 1 – SLIDES OF PARCOS STORYTELLER

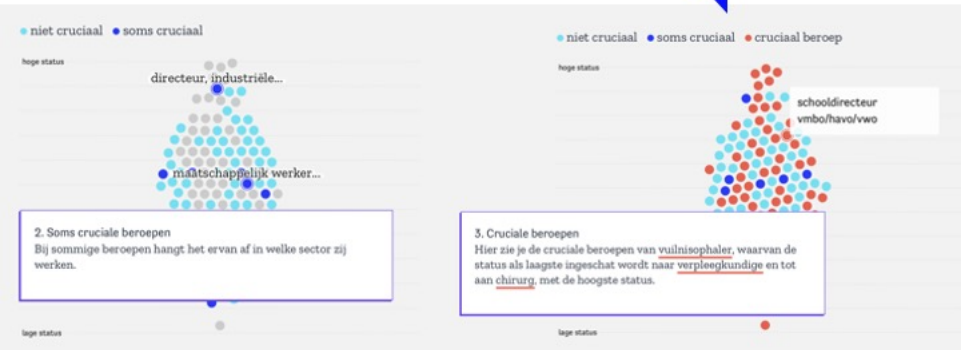


searchingcovid19.com - Finding Patterns in Our Need for Knowledge (https://searchingcovid19.com)



Example: https://pointer.kro-ncrv.nl/artikelen/cruciale-beroepen-zijn-onmisbaar-in-corona-tijd-maar-de-onderlinge-salarisverschillen

2. By presenting the main story in text, you can allow readers to further explore data for themselves



Telling Stories with Data

Why should I care?

It is a way to tackle misinformation

\+ it supports data literacy skills

Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics, 12*(3).

Telling Stories with Data

Why should I care?

It is a way to tackle misinformation

How?

\+ it supports data literacy skills

Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics, 12*(3).

3. Adding data representations to stories contributes to data literacy skills by making data feel familiar

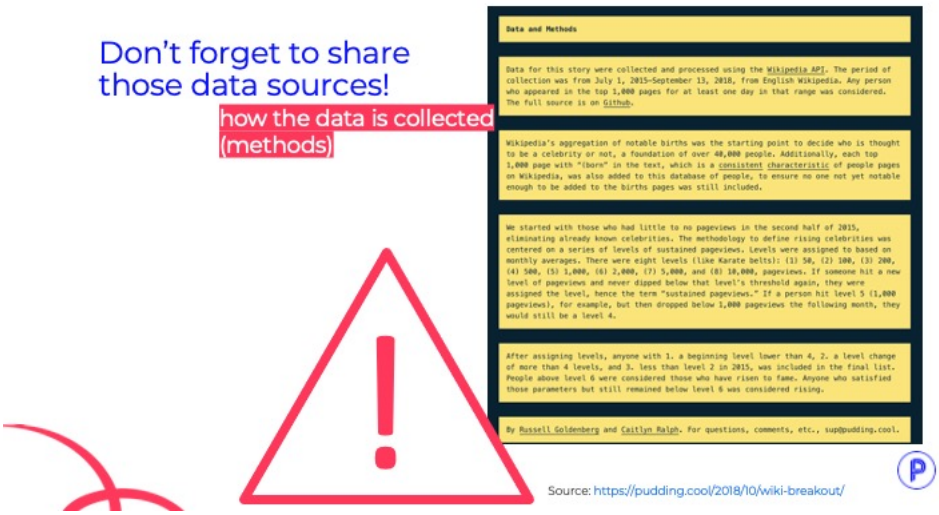
https://www.newscaststudio.com/2019/04/09/weather-channel-climate-change-imr/

OK! I am convinced.
Where do I begin?

Step 1.
Finding data


Besides exploring existing data sources,
You may also ask the audience to contribute data
for example
Collecting citizen science data to inform the story
HOE GEZOND IS DE LUCHT WAARIN U NAAR UW WERK FIETST?
Word één van de 20.000 CurieuzeNeuzen en meet de luchtkwaliteit in uw straat
CURIEUZE NEUZEN
VLAANDEREN
Curieuze Neuzen in Belgian newspaper de Standaard


The collected data does not always need to be about serious topics!
2006
2 of 10
Do you recognize this song?
don't know it
sounds familiar
know it
singing the lyrics
% of People Who Know this Song
18% Boomers
35% Gen X
63% Millennials
45% Gen Z
You knew 11 of 20 songs, which were just added to the 3,438,337 data points we have on how music fades from cultural knowledge.
Percent of people who recognize a clip of "No Diggity," by birth year
11-year-olds
5-year-olds
Born
Birth Year
for example
Collecting opinions


Don’t forget to share those data sources!
how the data is collected (methods)
Data and Methods
Data for this story were collected and processed using the Wikipedia API. The period of collection was from July 1, 2015–September 13, 2018, from English Wikipedia. Any person who appeared in the top 1,000 pages for at least one day in that range was considered. The full source is on Github.
Wikipedia's aggregation of notable births was the starting point to decide who is thought to be a celebrity or not, a foundation of over 40,000 people. Additionally, each top 1,000 page with "born" in the text, which is a consistent characteristic of people pages on Wikipedia, was also added to this database of people, to ensure no one not yet notable enough to be added to the births pages was still included.
We started with those who had little to no pageviews in the second half of 2015, eliminating already known celebrities. The methodology to define rising celebrities was centered on a series of levels of sustained pageviews. Levels were assigned to based on monthly averages. There were eight levels (like Karate belts): (1) 50, (2) 100, (3) 200, (4) 500, (5) 1,000, (6) 2,000, (7) 5,000, and (8) 10,000, pageviews. If someone hit a new level of pageviews and never dipped below that level's threshold again, they were assigned the level, hence the term "sustained pageviews." If a person hit level 5 (1,000 pageviews), for example, but then dropped below 1,000 pageviews the following month, they would still be a level 4.
After assigning levels, anyone with 1. a beginning level lower than 4, 2. a level change of more than 4 levels, and 3. less than level 2 in 2015, was included in the final list. People above level 6 were considered those who have risen to fame. Anyone who satisfied those parameters but still remained below level 6 was considered rising.
By Russell Goldenberg and Caitlyn Ralph. For questions, comments, etc., sup@pudding.cool.
Source: https://pudding.cool/2018/10/wiki-breakout/

It is a non-linear process!

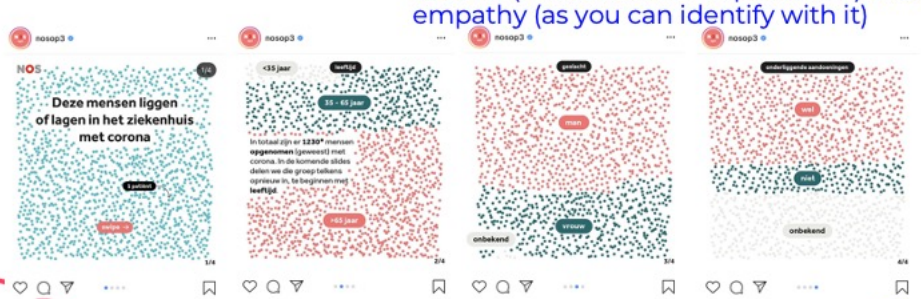
a. The data may be the starting point -> Analyze to find insights

b. The story may be the starting point -> Find data to support the narrative

Source: More than Telling a Story: A Closer Look at the Process of Transforming Data into Visually Shared Stories - Bongshin Lee

Data may be the starting point: the story may reveal itself when analyzing and visualizing the data

This example is seemingly simple yet it incorporates narrative elements as conflict (in the readers’ expectation) and empathy (as you can identify with it)



Example: NOS op 3 - Infographics on Instagram - Demographic information of Covid-19 patients
Also good example: http://rhythm-of-food.net

b. The main story may already be set out, and data may be used to support it, which is often the case in stories that use citizen science

as citizen science data offers both quantitative data and qualitative data in the form of anecdotes, quotes, etc.

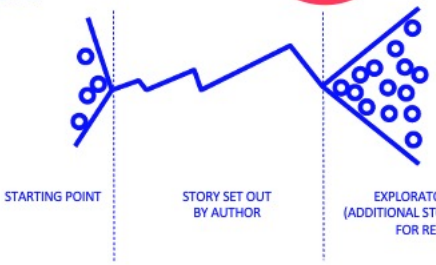
Example: Maai Mei Niet (*do not mow the lawn in May*) in collaboration with Weekly magazine Knack

Step 3. Telling the data story

Think of a narrative structure that facilitates data exploration, such as branched structures

This Martini glass structure is most often used for data stories

It's a Martini glass on its back!



Scrollytelling examples are typical for the Martini Glass structure:

starting with a story that is fixed, scrolling down to the point that the reader may interact freely with the data

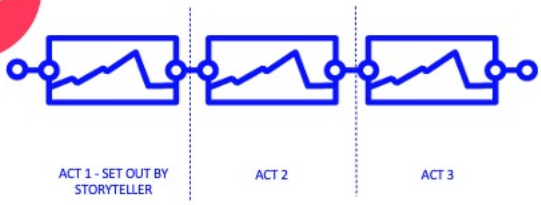
Example: koningshuizen.be (Immo Royal) about the properties of Belgian monarchies

Similar examples: https://www.juiceanalytics.com/writing/20-best-data-storytelling-examples

A second type of narrative structure for data stories is meant for more complex stories

The Slideshow

It allows data to be exposed step-by-step



READER CAN CHOOSE TO CONTINUE, GO BACK, SELECT SOMETHING ELSE

This example repeatedly switches between textual story elements with interactive plot charts that allow the reader to further explore

Example: https://pointer.kro-ncrv.nl/artikelen/cruciale-beroepen-zijn-onmisbaar-in-corona-tijd-maar-de-onderlinge-salarisverschillen

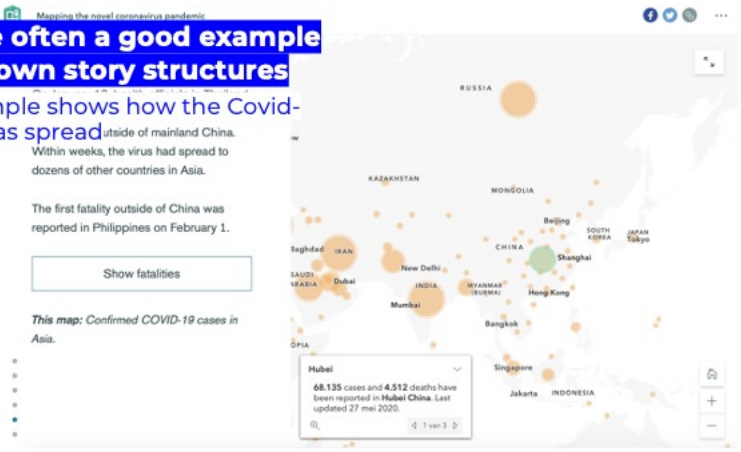
This third structure is most ‘open’. Readers may interact at any point in the story

Readers can zoom in, ask details, explore, etc.

Drill-Down

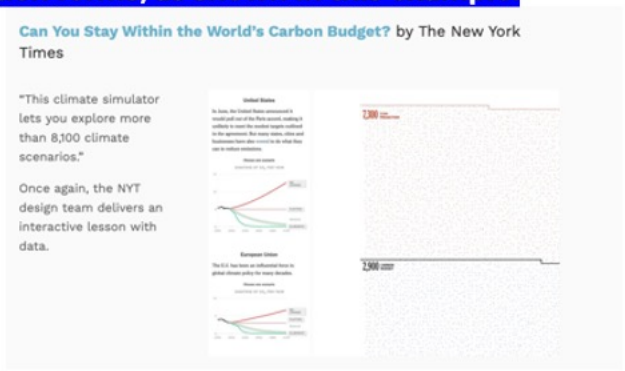
Maps are often a good example of drill-down story structures

This example shows how the Covid-19 virus has spread



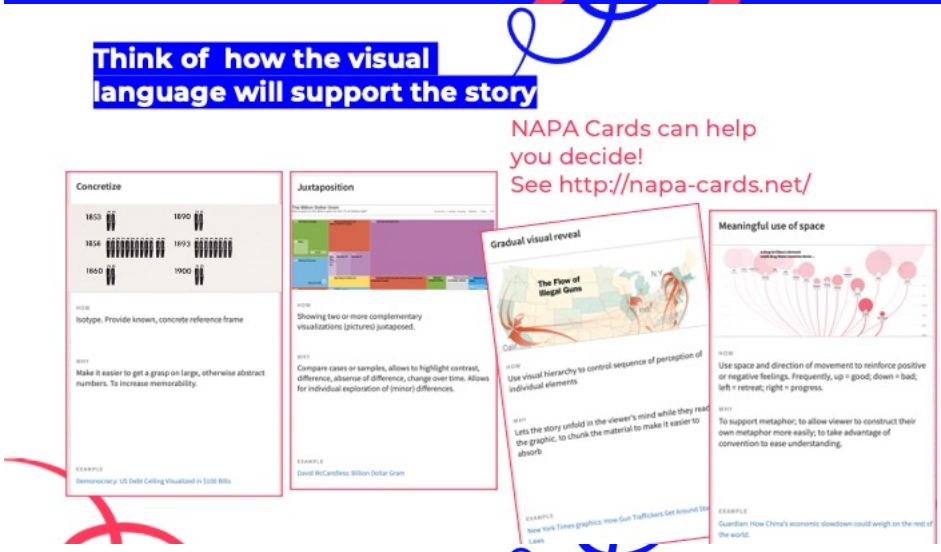
Example: https://storymaps.arcgis.com/stories/4fdc0d03d3a34aa485de1fb0d2650ee0

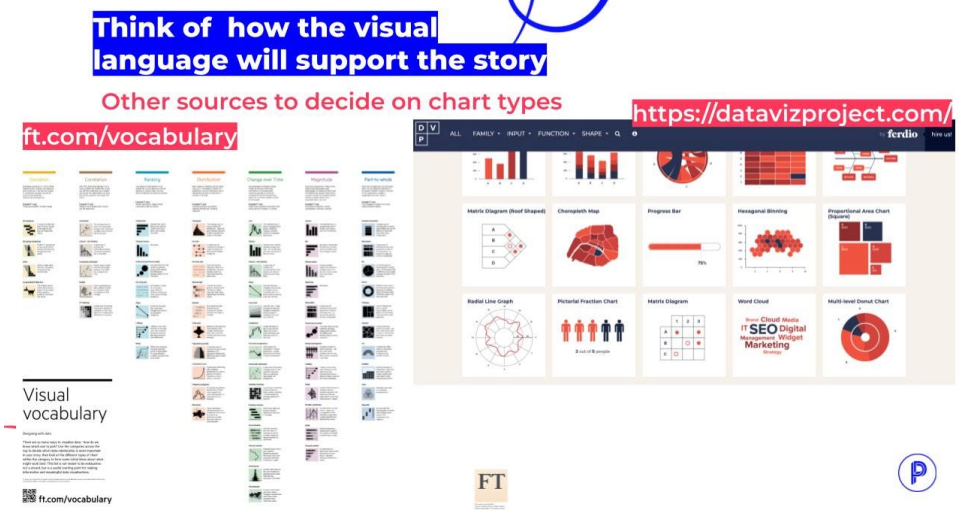
But drill-down story structures can also take other visualization forms, as shown in this example:



Source: https://www.nytimes.com/interactive/2017/08/29/opinion/climate-change-carbon-budget.html?mtrref=undefined&gwh=C51494A17110C96706B65ED1C6CDD9B9&gwt=pay&assetType=PAYWALL

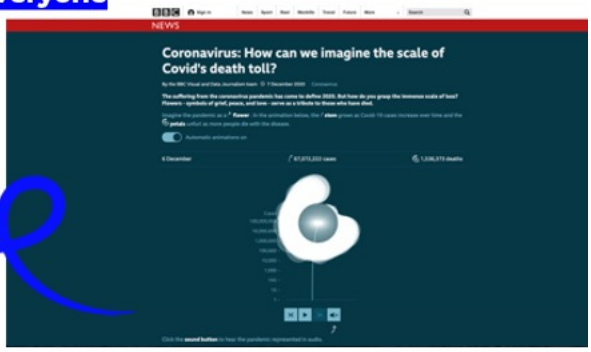

Step 4.
Presenting the data story


Think of how the visual language will support the story
NAPA Cards can help you decide!
See http://napa-cards.net/
Concretize
Juxtaposition
Gradual visual reveal
Meaningful use of space


Think of how the visual language will support the story
Other sources to decide on chart types
ft.com/vocabulary
https://datavizproject.com/
Visual vocabulary
ft.com/vocabulary
FT

Also consider other data representation forms, such as sound, to include everyone

BBC NEWS allows both visuals and sound to explore data



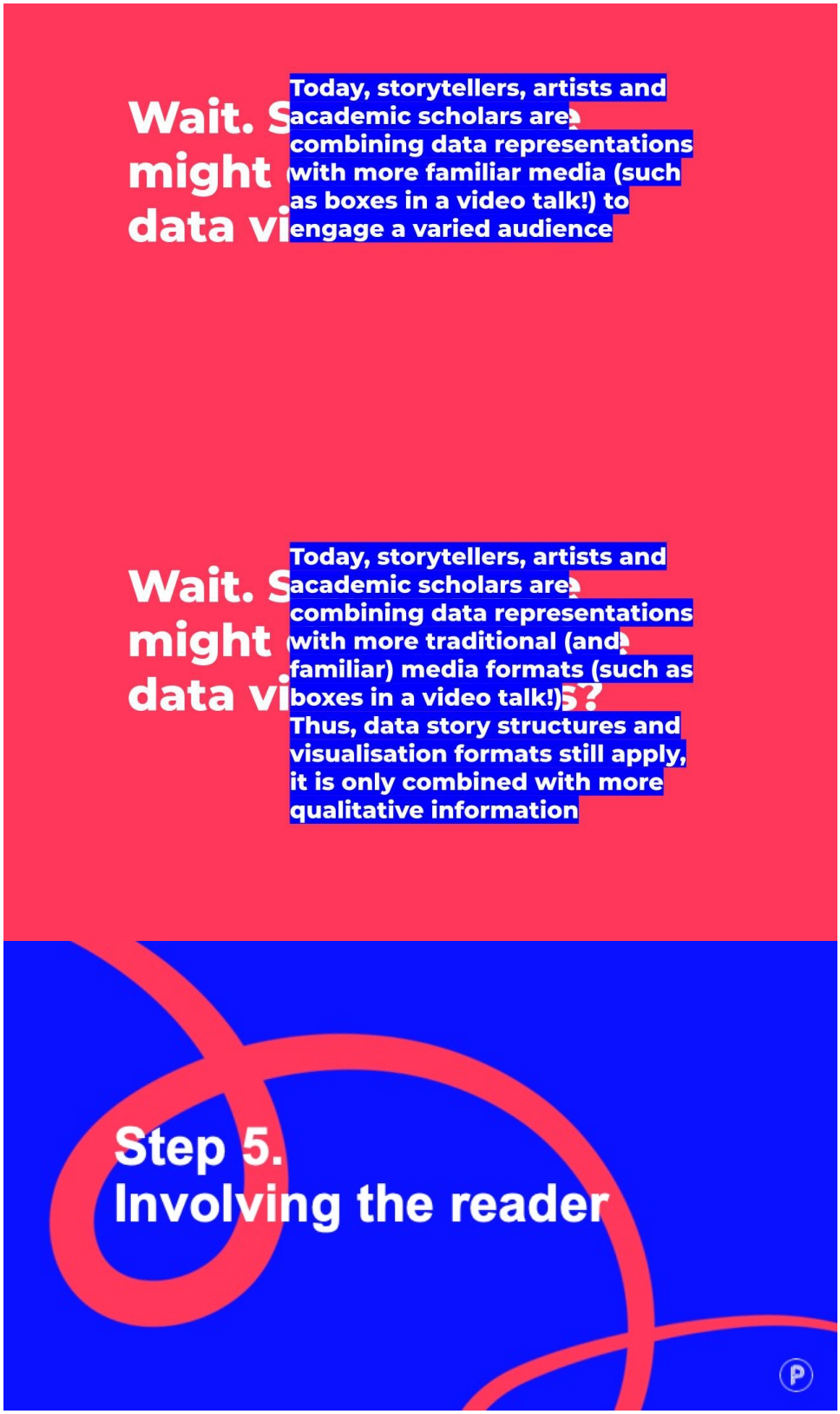
Source: https://www.bbc.co.uk/news/resources/idt-7464500a-6368-4029-aa41-ab94e0ee09fb
More examples can be found on: https://datajournalism.com/read/longreads/data-sonification

Or use physical data representations. These offer a friendly and familiar way to engage people that disregard digital visualizations as bar charts and line graphs

Hans Rosling became famous for telling data stories with boxes

Source: https://www.ted.com/talks/hans_rosling_global_population_growth_box_by_box

Wait. Some people might disregard nice data visualizations?


Wait. S
might
data vi
Today, storytellers, artists and academic scholars are combining data representations with more familiar media (such as boxes in a video talk!) to engage a varied audience
Wait. S
might
data vi
Today, storytellers, artists and academic scholars are combining data representations with more traditional (and familiar) media formats (such as boxes in a video talk!)s?
Thus, data story structures and visualisation formats still apply, it is only combined with more qualitative information
Step 5.
Involving the reader

When presenting data, the reader is already involved in some way of sensemaking

Key is to trigger active participation to truly facilitate data literacy skills and combat misinformation

participation: high / low / none

- add data storyline A
- add data to storyline B
- interpret data to progress storyline C
- interpret data to discover storyline D
- consume data story E

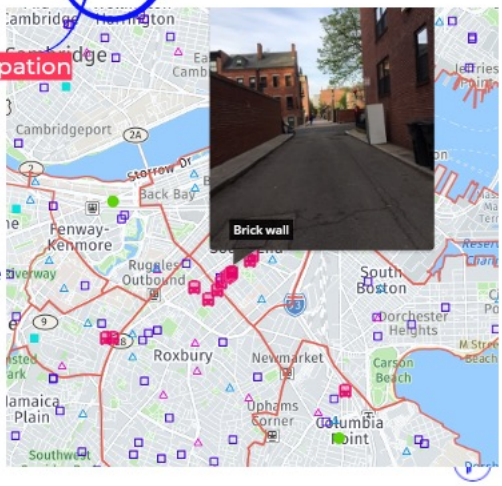
passive participation

Add data storyline

active participation

The ultimate form of participation! Readers may contribute their own data interpretation through anecdotes, photos...

...or both, as in this example that shows the issues with student transportation from students' perspectives.



Source: https://www.vitalvillage.org/data-dashboard/customize/student-transportation-equity-map

B. Add data to storyline

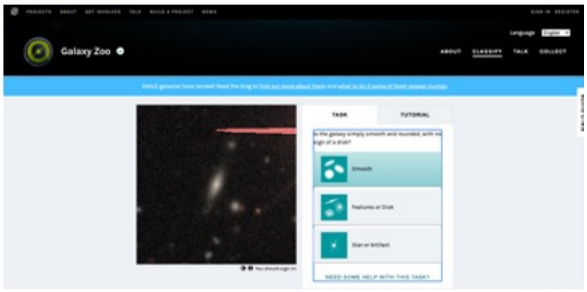
It is another way to establish a personal connection with the story and the presented information

Source: https://viewer.curieuzeneuzen.be/

C. Interpret data to progress storyline

The reader has part in the story, which connects on a personal level

For example, by identifying galaxies, the user is involved in a small part of the larger story that entails the complexity of space



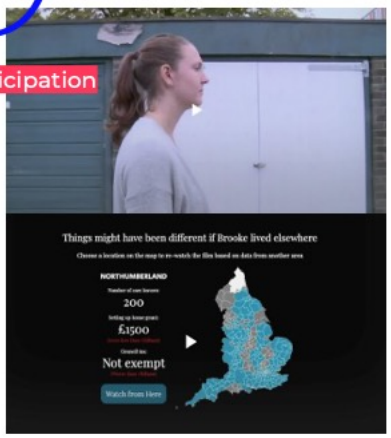
Source: https://www.zooniverse.org/projects/zookeeper/galaxy-zoo/classify

D. Interpret data to discover storyline

passive participation

Make data relevant by connecting it to the reader’s personal situation

in this example, they made a documentary film about ‘Brooke’ who lives near the user, depending on the users’ geolocation data. As Brooke is a ‘real’ person, it is also easier to have empathy for her



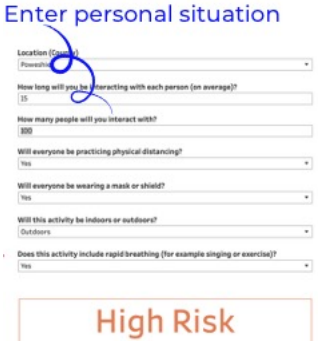
Concannon, S., Rajan, N., Shah, P., Smith, D., Ursu, M., & Hook, J. (2020, April). Brooke leave home: Designing a personalized film to support public engagement with open data. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-14).

D. Interpret data to discover storyline

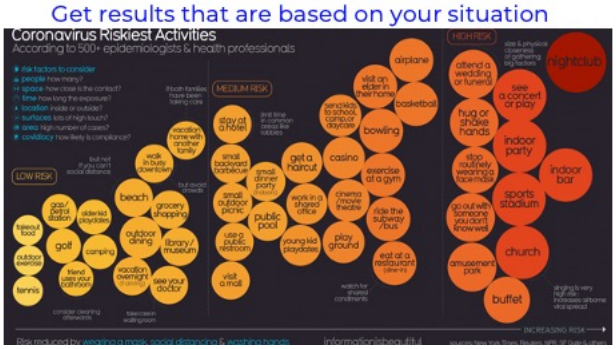
passive participation

Make data relevant by connecting it to the reader’s personal situation

Enter personal situation



Get results that are based on your situation



.