Co-Design of a Voice-Driven Interactive Smart Guide for Museum Accessibility and Management

 Xi Wang ✉

University of Bristol

_____

Abstract

This paper describes the process of co-designing and creating a voice-driven interactive smart audio descriptive guide for Titanic Belfast, a world-leading tourist attraction. This smart audio descriptive guide is intended to enhance museum accessibility and visitor experience, especially for blind and partially sighted (BPS) visitors. A key research question is to explore to what extent museums can conveniently produce their own smart guide to enrich the visitor experience for BPS visitors.

The paper first outlines the necessarily complex set of team functional roles and users in designing the smart audio descriptive guide and then presents the main challenges and opportunities arising from the key user requirements from both BPS visitor and the Museum Management perspectives. The main design features of the smart audio descriptive guide, which address these requirements, are then described.

The paper then outlines the main findings of our evaluative review of the smart guide with a group of BPS participants and from a Museum Management perspective. One of the key benefits of our approach is that the smart audio descriptive guide has the potential to offer museums and cultural venues a new, affordable approach to providing and maintaining a high-quality accessibility experience with lower design effort than traditional audio descriptive guide approaches.

Keywords: BPS visitors, museum accessibility, voice-driven, audio descriptive guide, visitor experience.

✉ xi.wang@bristol.ac.uk, https://orcid.org/0000-0002-0588-1156

Introduction

A new museum definition has recently been approved by the International Council of Museums (ICOM) in 2022. This stresses that museums should be accessible and inclusive (ICOM, 2022). Museums are no longer institutions that only serve the roles of preservation and education but also foster participatory, engaging, and accessible experiences (Black, 2020; Earle, 2013).

Audio Description (AD) is a primary tool used to provide access to museums for blind and partially sighted (BPS) visitors. AD as a form of media accessibility (Remael & Reviers, 2019) is usually aural, verbal descriptions of visual objects (Fryer, 2016). However, the recently published heritage access survey by VocalEyes (2022) shows that only 5% of UK museums provide audio descriptive guides, compared to 3% in 2018 (Cock et al., 2018).

There are various reasons for museums failing to provide audio descriptive guides. The cost of producing an audio descriptive guide can be an important factor, especially for smaller local museums. Consequently, museums can struggle to meet their accessibility aspirations and obligations. One opportunity for alleviating this situation is to consider the use of emerging technologies.

Current assistive technologies used in museums present some disadvantages for BPS visitors. Many museums that provide access facilities offer a keypad or touchpad-based audio descriptive guide (e.g., Titanic Belfast, the British Museum, and the app-based guide in New York's Museum of Modern Art). BPS users need time to become familiar with the AD controller, and keys are sometimes hard to identify. Interaction between the keypad/touchpad-based audio descriptive guides and users is still limited. Another solution is to use QR codes; by scanning the QR code of an exhibit, visitors can receive relevant information on their mobile devices. Installing beacon technology is another way of triggering the AD without requiring contact. Beacons are small Bluetooth-enabled transmitters which can be received by a mobile phone. When the mobile device detects a nearby beacon, visitors receive the corresponding AD in the associated app on their device. Again, there is limited interaction between AD and users.

Chatbots and Artificial Intelligence (AI) are new technologies that are opening up new possibilities for users to explore exhibitions based on their own preferences and interests. Most museum chatbots have been developed using the Facebook Messenger platform (Varitimiadis et al., 2020).

However, the use of such third-party platforms has raised problems of user privacy and data security, given some well-publicised cases of the misuse of personal data on such platforms, e.g., the Cambridge Analytica data mining scandal. However, chatbots can also require a lot of typing, which can be a challenge for BPS users compared with speech input.

Another recent AI-based chatbot is ChatGPT, released by OpenAI in November 2022. It interacts with users in a conversational way and can answer follow-up questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests (OpenAI, n.d.). In March 2023, OpenAI released its new version GPT-4, which can also respond to images.

Another relevant technology suitable for BPS visitors is intelligent speech interfaces, such as Amazon's Alexa, Apple's Siri, Google Assistant and Microsoft's Cortana (Diederich et al., 2022). BPS visitors are likely to have experience in using such devices. Driven by AI-based natural language processing, these devices create the feeling of a human-like interaction (Følstad & Brandtzæg, 2017). Intelligent speech interfaces have been used for different purposes, such as intelligent assistants (Islas-Cota et al., 2022), therapy (Bell et al., 2019), and interactive games. Beyond focusing on the intelligibility and naturalness of such voice user interfaces, studies have also drawn attention to other social factors, e.g., the diversity of people's speech, which also influences the production and perception of speech (Sutton et al., 2019). This is an important consideration for tourist venues with an international visitor profile, where it may be helpful to synthesise local accents, provided these are readily understandable.

This paper will present how we have used smart software-enabled technologies to enhance accessibility and visitor experience for BPS visitors. We have developed a customisable interactive voice-driven smart audio descriptive guide for BPS visitors to museums and cultural venues, and we customised it for Titanic Belfast as a case study (Titanic Belfast, n.d.). Our ultimate goal is to provide an experience as close as possible to having a personal venue guide. This paper focuses primarily on the user requirements of both BPS users and museum staff, which arose while creating our smart audio descriptive guide.

The paper is structured as follows. First, because the end users of our smart guide include BPS visitors, a wider range of skillsets is needed in the design team, including accessibility skills and a larger set of expert roles. We identify these functional roles in Section 1. Section 2 highlights a number of key user

requirements which present unique challenges. Section 3 describes how the design features of our smart guide address those requirements and challenges. Section 4 outlines the main findings of our evaluative review of the smart guide with a group of BPS participants and from a museum perspective. The paper's conclusion includes further developments that we are pursuing.

1. Functional Roles Within the Design Team

The process of designing an interactive smart audio descriptive guide proved to be inherently complex because of the range of skillsets involved. Therefore, we employed a user-centred approach (Greco, 2018) and a collaborative working mode for the smart guide design. Such interactive and collaborative teamwork has been introduced across different projects, including cooperative AD (Hirvonen et al., 2023) and collaborative meaning-making in museums (Neves, 2016). In this section, we highlight the various users and functional roles necessary in the process of co-designing and maintaining a smart guide.

The various users and functional roles are shown in Figure 1. The arrowed lines indicate the main responsibilities rather than merely the interactions. As we will see, the museum itself has significant user roles, which pose some interesting interaction and collaboration challenges. Note that each box in Figure 1 does not necessarily have to equate to a unique person: two (or more) of the functional roles may have to be fulfilled by the same person, particularly in a smaller museum with more limited resources.

Figure 1

Team Functional Roles and Users for Co-Design of a Smart Audio Descriptive Guide
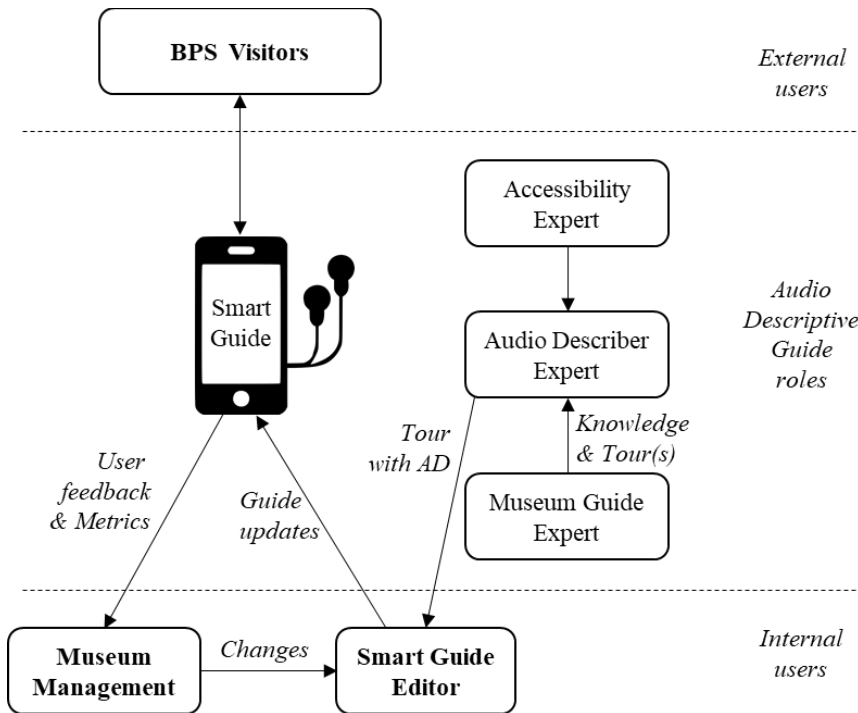
Note. Users Interacting Directly With the System are in Bold

Source: Author's own elaboration

Each of the players in Figure 1, including the users, is involved in the design process to narrow what Greco calls the maker-expert-user gap (Greco, 2018), where "users are not passive recipients but co-creators of meaning, and where each agent – maker, expert, and user – plays an (inter)active role in the semantic construction of an experience" (Greco, 2019, p. 32). With this in mind, we will now elaborate on the role of each of the players identified in Figure 1.

1.1. BPS Visitors

Potential BPS visitors have a valuable role throughout the design process, even before any technology is introduced. For example, in developing our smart guide, we began by first carrying out an extensive visitor experience study in Titanic Belfast (Wang et al., 2020, 2022).

1.2. Museum Management

The use of smart technology opens up new possibilities for Museum Management. As we shall see, it can provide greater flexibility in the layout and organisation of exhibits and exhibitions. In the case of a traditional audio descriptive guide, once it has been developed, Museum Management has little involvement, except perhaps to monitor how many visitors use the device. In practice, updating the audio descriptive guide is not feasible.

In developing our smart guide, we met with key Titanic Belfast staff to discover their aspirations and objectives. Titanic Belfast already had a commitment to accessibility and had already provided a keypad-based audio descriptive guide. Staff in the various galleries were also trained to look out for and to assist BPS visitors. Their involvement in the co-design process was thus well-informed.

1.3. Museum Guide Expert

In any museum, staff will be custodians of the museum's knowledge base and resources. There will also be those who curate this knowledge for the public and design the various ways to tour (parts of) a museum. These two different capabilities may (though not always) be found in the same person, such as the curator. The main contributions of this expert role include the following:

(1) To provide recommended tours through the museum and its galleries suitable for BPS visitors.

(2) To identify points in the tour where optional paths may be selected or omitted by the visitor.

(3) To make the overview information for each exhibition, gallery and exhibit, plus more specific knowledge and background stories for each exhibit in the tour, available.

(4) To select which parts of this information are core and should be presented each time and which parts are optional for the more interested visitor to select.

The Museum Expert Guide may not have specialist expertise in accessibility.

1.4. Audio Describer Expert

The Audio Describer Expert takes the recommended tours, exhibits and information resources from the Museum Expert Guide and creates the AD for each step of the tour.

1.5. Accessibility Expert

The Accessibility Expert's role is to support the team in accessibility awareness and to constantly review the proposed tours and the emerging AD during the early design stages.

1.6. Smart Guide Editor

Once the tour and the AD have been fully defined, this description must then be converted by the Smart Guide Editor into equivalent editable tour scripts, which can be uploaded to the smart guide. Software tools associated with our smart guide then synthesise the actual audio files. In subsequent operations, if changes need to be made to the tour or to the AD, these are made by the Smart Guide Editor, who edits the tour scripts and regenerates the new audio files. This can be done rapidly. This role requires a degree of understanding of structured data representation (see Section 3), and involves collaboration with team members who design the actual tour.

2. User Requirements

In this section, we look first at the main requirements for BPS visitors and then for museum staff. We have numbered the requirements for later reference when we discuss how these requirements are met in our design (in Section 3).

2.1. BPS Visitor Requirements

To discover the requirements of BPS visitors, we first took several groups of BPS visitors on a tour of Titanic Belfast and spent some time during and after the tour exploring their visitor experience and feedback (Wang et al., 2020). Later, we demonstrated a prototype of our smart guide to a focus group (Wang et al., 2022, Wang, 2023). From this, the key BPS user requirements were found to be as follows:

(1) The existing AD device, which required users to enter an exhibit number on a keypad to trigger the relevant AD, was awkward to use. Users required a more convenient way of interacting with the guide.

(2) BPS users would like the ability to ask for more information on an exhibit but do not want to be told everything about the exhibit regardless of their interest.

(3) BPS users would like to be able to ask free format questions in a similar way to addressing a human guide.

(4) The issue of navigation is a major concern. One factor is the safety issue, such as getting lost or unknowingly being near a potential danger point, such as a stairwell.

2.2. Museum Management Requirements

From our discussions with Titanic Belfast and other museum managers, there were four core requirements from a Museum Management perspective:

(1) Managers are responsible for ensuring that the museum meets its aspirations and legal obligations for accessibility and social inclusion. Smaller local museums with limited resources have difficulty meeting their legal obligations. A smart guide should help Museum Management in achieving this.

(2) Creating an audio descriptive guide is normally a rather expensive venture. For smaller museums, the cost can be prohibitive. Therefore, creating an audio descriptive guide for a venue without extensive financial investment should be possible.

(3) Management needs some way of receiving visitor feedback and metrics about using the smart audio descriptive guide in a way that enables them to continuously improve the smart guide and museum services. For example, how long did users spend in each gallery? Would users prefer more AD on a particular exhibit? Or less? Are there topics in which users would be interested that are not covered by the current AD? What questions do visitors ask, and at what points on the tour do they ask them? What types of help do BPS visitors request during the tour? Are there any issues or suggestions that BPS visitors would like to feed back to the museum? Gathering this data (anonymously) through the smart guide would be very

beneficial to managers. However, gathering the data should be as unobtrusive as possible for both visitors and staff.

(4) It should be possible to revise and update the smart guide. This process should be as quick, convenient and as low-cost as possible. Feedback from users may highlight places where the AD should be updated. Refurbishment may need the tour to be re-routed temporarily, or the museum may refresh its exhibitions, perhaps with additional exhibits. In a more extreme scenario, the museum may wish to present an additional temporary exhibition. While Museum Management usually has a commitment to this type of continuous improvement, it sometimes lacks the tools to support this commitment. The task of updating the AD is often so challenging that it is omitted. Therefore, it is a key requirement that the smart guide be capable of being updated rapidly and conveniently. This has far-reaching consequences for the software design.

3. Design Features of the Smart Audio Descriptive Guide to Meet the User Requirements

The user requirements above led to a number of fundamental design decisions in developing the software for our smart audio descriptive guide. In this section, we select a number of features of our smart guide which are intended to directly address these challenges and requirements: one feature which addresses both requirements of the Museum Management and BPS Visitors, one feature which addresses the requirements of the Museum Management, and three which are in response to BPS Visitor Requirements.

3.1. Tour Representation (Museum Management Requirements (2) and (4) and BPS Visitor Requirement (2))

To enable a venue to create and update its own smart audio descriptive guide without writing or modifying the software, a fundamental design decision was taken to enable the smart guide to operate from an editable script which contains all the AD for each stopping point in the tour. Behind the scenes, this tour description is just a structured text file. However, a graphical user interface makes it easier for the Smart Guide Editor to view and edit the tour script.

Recall that one of our BPS user requirements was to increase the autonomy of their visit, perhaps by selecting and/or omitting certain options along the tour, or requesting or declining more information on a particular exhibit. Thus, some parts of the tour will be guide-led (where the guide selects the path through the venue), and other parts will be user-led (where, at certain points, the guide gives the user some options from which they can select). The tour script must distinguish between which parts are guide-led and which are user-led.

To enable the tour script to be processed by the smart guide, it is necessary to formalise the Museum Guide Expert's concept of the guided tour and to represent it in a structured manner. In formalising the tour description developed by the Museum Guide Expert, the Audio Describer Expert, and the Smart Guide Editor, we represent the tour graphically and build the whole tour description using a small number of building blocks.

The first step in defining a tour is to represent each point where the smart guide would speak to the visitor as a rectangular box with the name of the stopping point. This might be the entrance to one of the exhibitions, or it might be a specific exhibit. For example, arriving at Exhibition A (which might represent one of the galleries in Titanic Belfast) would be represented as shown in Figure 2.

Figure 2

Building Block Representing a Tour Section for Exhibition A
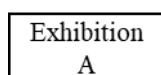
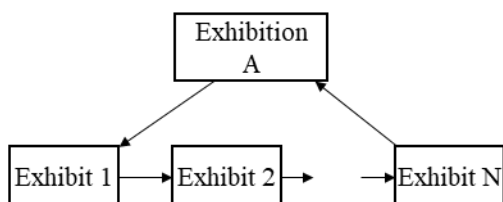Source: Author's own elaboration

Within this building block, we would add the text of the AD, which the smart guide will deliver at point A (perhaps an overview of the gallery that Exhibition A represents; or if A is just a single exhibit, then this would be the first level of AD for the exhibit).

If the tour through Exhibition A consists of several internal exhibits to be covered in some form of sub-tour, then we can represent this by an initial hierarchical decomposition of Exhibition A into the various parts of the sub-tour, as shown in Figure 3.

Figure 3

Initial Hierarchical Breakdown of a Tour Segment A Into Sub-Elements
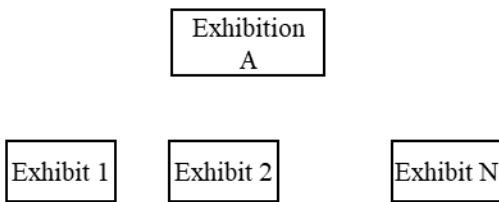
Source: Author's own elaboration

Each rectangular box is called a "node." The next step is to determine whether this internal tour of Exhibition A is to be guide-led or user-led: in other words, will the smart guide lead the user through each Exhibit (1–N) in turn, or will the smart guide give the user a choice as to which, if any, of the options (1…N) they wish to explore? If the tour of Exhibition A is to be guide-led, then we use the first interconnection pattern shown in Figure 4 to indicate the exhibits that are to be visited in the guide-led order. The use of unbroken lines indicates that this tour segment is guide-led.
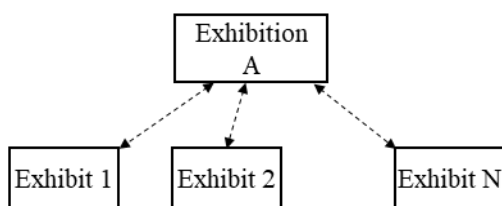
Figure 4

Diagram Representing a Guide-Led Tour Through Exhibition A

Source: Author's own elaboration

Alternatively, if the tour through a particular section or exhibition is to be user-led, we represent this using a second type of interconnection pattern as shown in Figure 5. Here we use dashed lines to indicate that the options visited and their order are not predetermined by the smart guide. The user can choose to visit any, all, or none of the N optional exhibits in Exhibition A, and in any order. The user will be presented with the options by the smart guide; they can select the option they want to visit next or can choose to exit the tour of Exhibition A. This process is repeated until the user either chooses to exit Exhibition A or perhaps until they have visited all the optional exhibits.

Figure 5

*Diagram Representing a User-Led Tour Through Exhibition A*

Source: Author's own elaboration

The user-led pattern is a convenient way to give the user the option of more information on an exhibit should they wish to have a deeper degree of engagement. This situation can be represented as shown in Figure 6. The first box (Exhibit 1) gives the basic AD for Exhibit 1, which is spoken to every visitor, while the optional "More on…" box will, if chosen by the user, give the extra information about Exhibit 1.

Figure 6

Representing the Option of More Information (a Simple Case of a User-Led Tour)
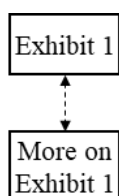
Source: Author's own elaboration

These two basic types of building block can be composed in a hierarchical fashion to any depth. So, for example, we could define an additional "Even more on Exhibit 1" piece of AD, by adding it as an option within "More on Exhibit 1" in Figure 6.

For a complete museum tour, the whole tour will begin at a single box at the top of the diagram. In a computing context, this type of diagram is called a "tree": imagine an upside-down hanging tree with its root (the tour start) at the top, and the branches representing sub-tours in the hierarchy. As mentioned, a node contains the AD to be spoken when the user arrives at the location. The so-called "leaf" nodes have no sub-nodes below them: they are at the very end of the tree branches and

represent a single exhibit. A node has other settings, such as the narrator's voice to be used for the AD, selected from a list of available voices.

To demonstrate the use of a tree to define a larger hybrid tour, we will show the tour representation of an outline of the whole tour of Titanic Belfast, in which we also populate the more detailed tour of the Fit-Out Gallery alone, for brevity. Figure 7 illustrates the resulting (partial) tree representation of the Titanic Belfast tour. Starting the tour at the entrance, the AD for that node gives the visitor a brief welcome and provides basic instructions about using the smart guide. The user will then be given five options for what they can do next, namely Ticket Centre, Gallery Overview (an information-only node), Gallery Tour, Toilets, and a group of options called "Others" (e.g., Restaurant, Café, and Gift Shop). The reason for grouping these "other" options into a single option is to avoid having to read out a list of nine or ten options; if we provide too many options at once, users may easily forget some of the earlier options and get frustrated. Therefore, if there are more than four or five options, using this type of "other" grouping can be advisable. Of course, the Smart Guide Editor has the flexibility to have all nine options read out, simply by not having a group called "Others."

By selecting the Gallery Tour option, the smart guide will provide an AD to commence the Gallery tour. Our slightly shortened Gallery tour consists of visiting seven galleries in a predefined order. The smart guide will then lead the user to the first gallery. Note that this main part of the tour of the galleries is guide-led, because of the sequential nature of the galleries.

For illustrative purposes, we have populated just the fourth gallery in the sequence – the Fit-Out Gallery. Our (shortened) tour of the Fit-Out Gallery has four exhibits, which are visited in order. Again, this is guide-led. However, at each of the four exhibits within the Fit-Out Gallery tour, there is a user-led option for more information. This shows the flexibility in defining different parts of the tour in a hybrid way (guide-led or user-led). If the user selects the option for more information at a particular stop in the Fit-out Gallery, the system then will move to that node in the tour description and play the AD in the node.

Figure 7

A Partial Tree Representing a Tour of Titanic Belfast Focusing on the Fit-Out Gallery
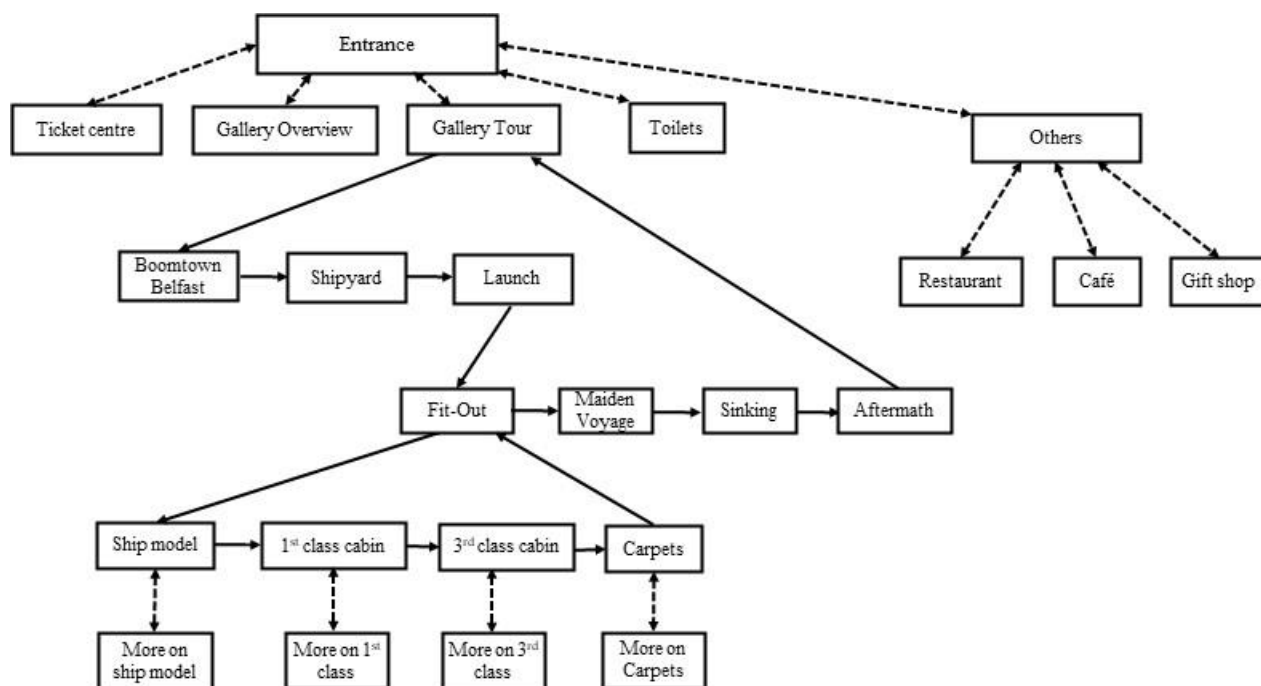
Source: Author's own elaboration

Notice that the Smart Guide Editor should also populate a separate Help tree (user-led), as shown in Figure 7. Visitors can ask for help at any stage of their visit; the help options include how to use the smart guide, the location of the nearest toilets, and information about the café, restaurant and gift shop. When a user exits the Help facility, the guide will automatically go back to where the user was in the tour AD when they asked for help.

Once the tour description for the whole tour has been created (or edited) by the Smart Guide Editor, the description is processed and downloaded to the smart audio descriptive guide device, at which point it will be available to visitors.

3.2. Gathering User Feedback (Museum Management Requirement (3))

An important requirement for Museum Management is to get feedback from users. This is often achieved through post-visit questionnaires, a time-consuming and resource-intensive process for the museum. It is also preferable if visitors can provide feedback during the visit, since visitors may forget

suggestions or comments which occur to them at particular points in the tour. The gathering of feedback should ideally be unobtrusive and immediate.

Fortunately, we can use the software-controlled smart audio descriptive guide to help meet this requirement in ways which would not be feasible with the standard type of AD device. We have implemented two facilities for feeding back information to the management (subject, of course, to the standard anonymisation, user permission and ethical review processes):

3.2.1. On-the-Fly Metrics

The smart guide can maintain a time-stamped log of each user interaction with the smart guide throughout the tour, including when the AD for each exhibit is started and how much of it is listened to. The user's choice of options in user-led portions of the tour can also be recorded automatically. From this information, consolidated spreadsheets can be generated which show which parts of the tour were visited the most, which options for additional information were (or were not) selected, which parts of the tour were more (or less) popular, how often each Help section was consulted, and other usage patterns. The system can also log any questions which the user asks of the smart guide (see Section 3.4 below). This latter information is particularly useful, as it may give an indication of what extra information might usefully be added to the AD. It may also give the museum ideas about how the tour might be modified and improved. An important advantage of this way of gathering feedback is that it involves no effort or time on the part of either the user or Museum Management.

3.2.2. Direct In-Tour Questions

In the tour description above, we have the additional optional ability to include, for any node, any number of "feedback questions" – questions which the smart guide will ask the user just before the user moves on from that point in the tour. Feedback questions might be, for example, "How do you feel about the tour so far?", "What did you think after touching the carpets?" or "Do you have any comments or suggestions about the Fit-Out Gallery?." The users' responses are converted to text and included in the log. This facility can be used to replace questionnaires. Care must be taken not to overload the visitor with such questions. On the other hand, it can be used judiciously to give BPS

users a chance to reflect on their emotional responses in a way which may be beneficial to them. This is where the advice of the Accessibility Expert is valuable.

3.3. Speech Input and Output (BPS Visitor Requirement (1))

Our smart audio descriptive guide is completely voice-driven. Speech recognition removes the need to learn to operate a keypad device and is also more appropriate in a post-COVID-19 world. It does require an internet connection at all times, however. This could be a potential problem, for example, in lifts or other areas with no internet connectivity.

One of the most important features of our smart guide is that it uses high-quality speech synthesis to speak the AD. This removes the need to hire professional (human) readers. This is especially beneficial when small updates to the AD need to be made. If human readers were used for the original audio descriptive guide, it might be difficult to get the same human reader as before.

All the AD and all the various responses that the smart guide can give to the user can be pre-synthesised offline and downloaded to the smart guide before the tour starts. This makes the smart guide less dependent on bandwidth throughout the tour and reduces latency. This also enables the provision of different narrators for different parts of the tour, a feature suggested by our BPS visitor focus group in our current study (see Section 4). They suggested, for example, that the AD for the First and Third Class cabins could be spoken by a hypothetical crew member responsible for looking after the rooms during the voyage. Because speech synthesis can provide a range of different voices and accents, the Audio Describer Expert can create imaginary human guides who can be readily visualised by BPS visitors, thus making the whole experience more person-oriented, engaging and enjoyable. This can easily be defined for each node by choosing the appropriate voice. One current limitation, though, is that the set of voices and accents that can be synthesised is limited and often does not include the local accents we might want to use in specific instances.

3.4. The Question and Answer Facility (BPS Visitor Requirement (3))

As part of our strategy to give as much control as possible to the visitor, we have provided our smart guide with a Question and Answer facility. At any point in the tour, the user can say something like "Can I ask a question?," and the smart guide will then listen to any questions the user asks and

respond as best it can.

The approach we have taken to providing a customisable Q&A facility is as follows. A predefined textual database of questions (and answers) is created. This can be searched by matching the spoken question with the stored question in a database and retrieving the stored answer. We use several natural language processing techniques (described below) to improve the accuracy in recognising users' questions.

We based our Q&A database on a book entitled 882½ Amazing Answers to your Questions about the Titanic (Brewster & Coulter, 1998). This book is a good resource for building a prototype Q&A database since the content corresponds well to the galleries and exhibits in Titanic Belfast. Other questions and answers could, of course, be added from other sources, including visitor questions which staff have received in the past.

At present, if a question cannot be matched against the database, the smart guide simply replies that it doesn't have the answer to the question. A better solution would be to include a human museum expert on call in the loop, to whom unmatched (or incorrectly matched) questions could automatically be sent in real-time, and whose answers would be spoken to the user by the smart guide. These questions and answers could then automatically be added to the Q&A database.

One major problem with free-format questions is that there are many ways of asking the same question, and visitors may not use the exact words as the question stored in the database. For example, our Q&A database includes the question: "How long did it take the Titanic to sink?" But users could ask for this same information in different ways, including: "How long after hitting the iceberg did the Titanic sink?" Therefore, we use four string-matching approaches to help recognise questions from users: fuzzy string matching, string matching based on synonyms, frequency-based word weighting, and alternative question forms.

Fuzzy string matching (also known as approximate string matching) finds strings that match a pattern approximately rather than exactly (Singla & Garg, 2012). When searching the database for the user's question, Fuzzy string matching calculates a matching score between the spoken question and each stored question in the database and selects the option with the highest matching score (provided the matching score reaches a certain threshold).

String matching based on word weighting means that words which occur with high frequency in the Q&A database (such as "the" or "Titanic") are regarded as less significant and given a low score. Words with low frequency occurrences are deemed to be more distinguishing and are given a higher score, e.g., "sink". This approach improves accuracy in recognising users' questions.

Another string-matching approach is based on word synonyms. This approach allows the user to speak a different word from the word in the stored question, but which has a high degree of semantic similarity, like [long, big, length], [design, designed, designer], etc. As part of the Q&A database, we can define a set of words and their associated alternatives. For example, the smart guide may give the option of toilets, but we can define synonyms so that the visitor can ask for the bathroom, the restroom, the ladies, the gents, etc.

Our final string-matching approach is based on providing multiple alternative question forms. We can specify several different formulations of the question for one single answer.

Finally, when a user first asks to put a question to the smart guide, the smart guide encourages users to start their question with certain keywords, such as Who, What, When, Where, Why, How many, etc. This has the benefit of constraining the number of ways of asking each question and can train the user to ask questions more efficiently.

We investigated the use of ChatGPT as an alternative to a customised Q&A facility. For more general questions about the Titanic, the responses were quite appropriate, although they tended to be more verbose than hand-crafted answers. Also, in practice, the answers could not be pre-synthesised, which would introduce an annoying delay. However, when seeking information which is not available online, the answers were unreliable. Future developments of ChatGPT may overcome some of these challenges.

3.5. Navigation Instructions (BPS Visitor Requirement (4))

The smart audio descriptive guide is designed ideally to enable BPS visitors to visit the museum autonomously. In our smart guide, we do not currently make use of any location-finding technology. Instead, just to demonstrate the concept, the smart guide gives navigation instructions such as how many steps are needed to go from the First-Class cabin to the Third-Class cabin, when to turn, and in which direction.

The navigation instructions are stored as an editable Navigation script. Speaking the navigation information is triggered when the user has received all the AD for their current location and is ready to be guided from their current physical location to the next physical location in the tour. The smart guide speaks the pre-stored directions from the Navigation Script. These directions can be viewed as being stored in a matrix, in a row and a column format, theoretically from one physical location to another. In practice, we need to store only directions to those locations which are next on the tour path (in other words, only those paths indicated by lines in the tree need to be described). The obvious limitation of this basic navigation approach is that it relies on users following the directions exactly. If the user gets lost, the system will not be aware of this, which is unsatisfactory in the longer term.

4. Smart Guide Evaluation With the BPS Focus Group

To carry out an initial evaluation of our smart guide prototype, we held an online focus group exercise with a group of BPS participants. We presented a demo of the smart guide, followed by a discussion. We also reviewed the prototype from the perspective of Museum Management. There were 13 people in the focus group, including 10 BPS participants, an RNIB staff member who recruited the BPS participants, a Zoom facilitator, and the researcher herself. Of the BPS participants, four identified as blind and six as partially sighted; three were women and seven men, their ages ranging from 23 to 75. Their educational backgrounds were also diverse, ranging from GCSE or equivalent to undergraduate degrees. All of the participants had visited Titanic Belfast at least once before the reception study and some of them had participated in other elements of our research. Some BPS participants have more advanced experience in using voice-activated devices, such as screen readers, than others.

All our tests involving BPS users received ethical approval from the School of Arts, English and Languages ethics committee at Queen's University Belfast. All participants gave their consent prior to participating. Below we highlight the main findings and recommendations from the evaluation. The first list is from the BPS focus group participants.

(1) The BPS participants were very enthusiastic about the smart guide and its potential. However, they recommended having different narrators for different parts of the tour, perhaps visualising some narrators as crew members of Titanic. Our system already has the

ability to do this, but we underestimated the value for BPS users of presenting each narrator as a person in their own right. This is actually a good practice for all visitors.

(2) In our BPS focus group, one younger BPS visitor preferred a faster rate of speech than the more mature members of the group. We could achieve this by synthesising the complete set of AD at two different rates by the Smart Guide Editor, with the visitor able choose their preferred rate of speech. Most available commercial speech synthesis tools enable the speed of speech to be selected.

(3) When testing the Q&A facility, it became apparent that BPS users are quite likely to come up with unpredicted questions. Some questions from the focus group were more about the human elements of the exhibits.

(4) The personal safety of BPS visitors was a major concern. They recommended that the smart guide be given a "panic button" or a spoken phrase that would put the user in direct contact with a staff member. Even if such facilities are not actually used, their very presence reduces the anxiety which BPS visitors may feel and which can reduce their enjoyment of the visiting experience.

The smart guide was also evaluated from the perspective of Museum Management.

(1) One challenge for museums is that the formal description of a guided tour as a tree structure is not necessarily how the Museum Guide Expert conceptually visualises the tour. The Smart Guide Editor is likely to be in a good position to help in translating the Museum Guide Expert's conception of the tour into the formal tree representation; but the problem should not be overlooked.

(2) Our decision to describe the tour as a tree diagram may, in some circumstances, be too restrictive. A visitor might want to miss out on some guide-led parts of the tour, perhaps because they are running out of time, or go back to a recently visited exhibit. This is a feature which could be added to the current system.

(3) The voice recognition feature of the smart guide requires a constant internet connection for speech recognition. This may require the museum to extend its Wi-Fi infrastructure, including coverage within lifts.

(4) There is a particular benefit in ensuring the emotional engagement of BPS visitors. For example, in Titanic Belfast's Sinking Gallery, the existing AD device for BPS visitors does not read out the visually presented texts of the Morse code messages sent from nearby ships to the Titanic. This turned out to be one of the most poignant moments of the tour and is an obvious example of the desire to update and improve the AD in the light of visitor feedback and experience.

The above observations reinforce the major benefit of being able to easily update the AD or other details of the tour, and to be able to generate the updated smart guide quickly and at little cost. It is unreasonable to expect a venue to get everything perfect the first time.

This discussion shows that this partial evaluation has been beneficial and provides a foundation for further evaluations. Nevertheless, we recognise the following limitations. The reception study was not conducted on-site. Thus, the BPS participants could not test the navigation instructions or use the smart guide themselves. They had to rely on the recorded demo, which was itself limited in terms of the features and capabilities it displayed.

In terms of group dynamics, one strong participant can sometimes lead the rest of the group, either in a positive or a negative direction. This is reflected by a comment from one of the BPS participants about the focus group:

> I found it difficult to get a word in as sometimes these remote discussions tend to be dominated by a few people… maybe I just need to learn to shout a bit louder; however, that is not really in my nature! In hindsight, I should have raised my hand to speak.

5. Conclusions and Future Developments

This paper has focused on some of the unique user requirements and challenges that arise when designing an interactive voice-driven smart audio descriptive guide for taking BPS visitors on a museum tour. We identified several key user requirements, both for the Museum Management and for BPS visitors. We then presented the main design features of our smart guide, which address these requirements. We evaluated our working prototype with a group of BPS participants and a review from a Museum Management perspective. It was clear from our evaluation that the smart guide has great potential, not only to improve the visitor experience for BPS visitors, but also to support the Museum Management in meeting their accessibility aspirations and obligations, and with reduced

investment compared to the traditional approach. From our evaluation, we have also identified several fruitful avenues to pursue in developing our smart guide further.

To develop our custom Q&A facility further, more advanced linguistic analysis techniques will be explored. Our goal is to be able to base the Q&A facility solely on the museum's knowledge base, without having to define all possible questions. We will also monitor developments with ChatGPT (latterly GPT-4).

The challenge of indoor navigation is one we plan to address. Accurate indoor location-finding technology is expensive, whereas beacon technology and Near Field Communication can be used to achieve many of the same benefits at a much lower cost. It can alert BPS visitors to potential danger points or areas of the building which are not on the planned tour path, at which point visitors can be redirected. Navigation assistance is crucial for reducing anxiety and increasing the autonomy which BPS visitors experience in a new environment.

Since the software for our smart guide is not dedicated to any one venue, we envisage that a single app can be used for any participating museum and historic garden (Wang, 2024). The availability of such a guide is in itself likely to give the venue a useful marketing advantage.

Another interesting potential development will be exploring the use of automatic translation for producing a multilingual version of the smart guide. The limitations of automatic translation will likely require simplifying the language used in the AD and the Q&A database.

We acknowledge the limitations of our methodology, specifically that the focus group was conducted online, preventing BPS participants from gaining hands-on experience with the smart guide. In future endeavours, we aim to explore the smart guide's potential for diverse visitors. To address this, multiple on-site focus groups with varied visitor categories will be conducted in future research to provide a more comprehensive understanding of the smart guide's usability and effectiveness.

Funding Details

References

Bell, S., Wood, C., & Sarkar, A. (2019). *Perceptions of chatbots in therapy. Proceedings of CHI Conference on human factors in computing systems extended abstract*, 1–6. https://doi.org/10.1145/3290607.3313072

Black, G. (2020). *Museums and the challenge of change: Old institutions in a new world*. Routledge.

Brewster, H., & Coulter, L. (1998). *882 1/2 amazing answers to your questions about the Titanic*. Firefly Books.

Cock, M., Bretton, M., Fineman, A., France, R., Madge, C., & Sharpe, M. (2018). *State of museum access 2018: Does your museum website welcome and inform disabled visitors?* VocalEyes, Stage TEXT and Autism in Museums. https://vocaleyes.co.uk/state-of-museum-access-2018

Diederich, S., Benedikt Brendel, A., Morana, S., Kolbe, L., & Benedikt, A. (2022). On the design of and interaction with conversational agents: An organizing and assessing review of human-computer interaction research. *Journal of the Association for Information Systems, 23*(1), 96–138. https://doi.org/10.17705/1jais.00724

Earle, W. (2013). Cultural education: Redefining the role of museums in the 21st century. *Sociology Compass, 7*(7), 533–546.

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. *Interactions, 24*(4), 38–42. https://doi.org/10.1145/3085558

Fryer, L. (2016). *An introduction to audio description: A practical guide*. Routledge. https://doi.org/10.4324/9781315707228

Greco, G. M. (2018). The nature of accessibility studies. *Journal of Audiovisual Translation, 1*(1), 205–232. https://doi.org/10.47476/jat.v1i1.51

Greco, G. M. (2019). Towards a pedagogy of accessibility: The need for critical learning spaces in media accessibility education and training. *Linguistica Antverpiensia, New Series–Themes in Translation Studies*, 18, 23–46.

Hirvonen, M., Hakola, M., & Klade, M. (2023). Co-translation, consultancy and joint authorship: User-centred translation and editing in collaborative audio description. *Journal of Specialised Translation*, 39, 26–51. https://www.jostrans.org/issue39/art_hirvonen.php

ICOM. (2022). Museum Definition – International Council of Museums (ICOM). https://icom.museum/en/resources/standards-guidelines/museum-definition/

Islas-Cota, E., Gutierrez-Garcia, J. O., Acosta, C. O., & Rodríguez, L. F. (2022). A systematic review of intelligent assistants. *Future Generation Computer Systems*, 128, 45–62. https://doi.org/10.1016/J.FUTURE.2021.09.035

Neves, J. (2016). Enriched descriptive guides: A case for collaborative meaning-making in museums. *Cultus, 9*(2), 137–154.

OpenAI. (n.d.). Introducing ChatGPT. Retrieved March 6, 2023, from https://openai.com/blog/chatgpt

Remael, A., & Reviers, N. (2019). Media accessibility and accessible design. In Minako O'Hagan (Ed.), *The Routledge handbook of translation and technology* (pp. 482–497). Rout. https://doi.org/10.4324/9781315311258-29

Singla, N., & Garg, D. (2012). String matching algorithms and their applicability in various applications. *International Journal of Soft Computing and Engineering (IJSCE)*, 6, 2231–2307.

Sutton, S. J., Foulkes, P., Kirk, D., & Lawson, S. (2019). *Voice as a design material: Sociophonetic inspired design strategies in Human-computer interaction. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. https://doi.org/0.1145/3290605.3300833

Titanic Belfast. (n.d.). Titanic Belfast. Retrieved December 13, 2021, from https://www.titanicbelfast.com/

Varitimiadis, S., Kotis, K., Skamagis, A., Tzortzakakis, A., Tsekouras, G., & Spiliotopoulos, D. (2020). Towards implementing an AI chatbot platform for museums. *International Conference on Cultural Informatics, Communication & Media Studies, 1*(1). https://doi.org/10.12681/cicms.2732

VocalEyes. (2022). *Heritage access*. Retrieved January 13, 2023, from https://vocaleyes.co.uk/research/heritage-access-2022/

Wang, X., Crookes, D., Harding, S., & Johnston, D. (2020). Evaluating audio description and BPS visitor experience in Titanic Belfast. *Journal of Audiovisual Translation, 3*(1), 246–263. https://doi.org/10.47476/jat.v3i2.2020.124

Wang, X., Crookes, D., Harding, S., & Johnston, D. (2022). Evaluating audio description and emotional engagement for BPS visitors in a museum context. *Journal of Translation Spaces, 11*(1), 134–156. https://doi.org/10.1075/ts.21019.wan

Wang, X., (2023). Design Principles for a Smart Descriptive Guide for Blind and Partially Sighted Visitors. Journal of Identity, Language and Diversity, 1,125-143.

Wang, X., (2024). Translating Gardens into Accessible Multisensory Tours for Blind and Partially Sighted Visitors: An Exploratory Case Study. Journal of Universal Access in the Information Society. https://doi.org/10.1007/s10209-024-01107-0