# ABSTRACT

The Museum Association of the Netherlands aims to increase visits to its nearly 500 museums. Through its partner, Q42, it is looking to increase engagement on its website museum.nl and letting users discover new museums. Of special interest is making it easier for users to discover museums relevant to their interests which are relatively less popular. A design approach of empathising with the user at the moment of search and discovery led us to conclude that the context of the occasion was important. A visit by the same person with her family, alone during an afternoon or during a visit to another city would lead to different user profiles. Therefore static user profiles for a recommender engine may not always give the most relevant results. We developed a prototype - TindArt, where an intuitive, playful interaction with an offered selection of documents lets the user express her implicit preferences that then applies an ensemble method to offer recommendations from three different recommender systems.

# KEYWORDS

museums, visitor experience, recommender systems, information retrieval, content-based filtering, collaborative filtering, ensemble learning

# INTRODUCTION

Recommender systems allow users to discover online content that match their preferences. Such systems look at features of previously selected documents and find new documents with similar features. Such a system could be a solution for a museum association website looking to help prospective visitors find inspiration for their next visit.

Developing such a system has technical and behavioural challenges to address. Visitors who are not logged in, or with cleaned cookies, would not have a profile that the recommender could use. This is the *cold-start problem* in recommender systems [1]. Behaviourally, it is important to understand how a visitor really makes her decision. Specifically, are the purpose of visit, mood that day, willingness to travel on that day etc. consistent enough to use the same user profile information per session. [9] Even if a recommender system does not have a cold start problem and has a user profile available, the data may be irrelevant in that instance.

This proof of concept aims to build a recommender addressing these two challenges. It provides a visual-based interaction that gives user profile information to the recommender. The recommender works on the preferences revealed by the user as in that particular session - allowing flexibility in responding to sessionspecific preferences. In the back end, the recommender system ensembles three models. The *collaborative model* looks at similari-

ties between museums as observed in user online behaviour. The *first content model* uses standard features such as location, category and facilities as used by the Museum Association. The *second content model* uses the text in museum description to find similarities. All three models calculate similarities between different museums in a similarity matrix and use these to make recommendations. Ensembling allows the stakeholders to add new data sources or models to this system in the future.

## Stakeholder context and needs

There are 438 museums in the Museum Association of the Netherlands. In total, they had 32.6 million visits in 2019.[13] There are three categories of visitors - international, domestic visitors and museumkaart holders (annual subscription[[1]](#footnote-1)). These three categories were nearly equal in respective share and are all important for the financial health of the museums. Due to COVID-19 restrictions in 2020, total number of visits was significantly impacted. As restrictions on international travel will continue well beyond 2020, it is important to stimulate visitor interest among domestic and

museumkaart visitors. [13]

Number of museum visits also vary by location and category.

For example, museums in the province Noord-Holland received

158.247,31 visitors on average in 2019. Even within the Randstad area, this average number nearly halves to 79.853,66 in South Holland and 77,933.33 in Utrecht.[13] The Museum Association categorises museums into five categories. Art is the most popular 102 art museums saw an average of 137.137,25 visits. The next most popular category is history and that only saw 51.664,15 visits, nearly a third. Therefore the stakeholders are not only interested in stimulating visitor interest, but also to divert interest to the less visited museums.

Museum.nl is the Museum Association’s website for visitors. It allows potential visitors to browse through different museums and exhibitions. Q42 is the technical partner to the Museum Association and is maintaining museum.nl. The new website was launched on the 1st of July 2020 in time for the relaxation of the first lockdown rules in the Netherlands. However, user engagement has since dropped. Q42 shared that the bounce rate doubled from 15 to 30 percent and average page visits per session dropped from five to three. In addition to stimulating interest and diverting interest to less visited museums, Q42 is also looking to improve user engagement on the site.

A couple of approaches are already in implementation on museum.nl. Visitors can explore museums by themes such as water, money or heroes. This approach however requires hand-made clusters of museums. A second approach is to show content cards at the bottom of all pages with "Discover More" call to action. Users are recommended other museums and exhibits. However in the stakeholder meeting Q42 shared that recommendations in these cards were currently at random, not influenced by user profile or content being viewed.

## Problem Statement

Based on stakeholder interviews, design audit of the website and research into related work, a problem statement and research question were framed. The current interactions on museum.nl were insufficiently capable in creating enough user engagement to motivate visitors to go to (smaller) museums. The subsequent research question therefore was - what improvements could be implemented to provide the right data to a recommender system on museum.nl, that both increases user engagement while interacting with the website and recommending relevant (smaller) museums. This research question guided the development of this prototype.

# RELATED WORK

## Behavioral models of the museum visitor

Museum studies look at various motivations that direct a museum visit. Scholars such as Falk argue that intrinsic factors such as a visitor’s identity is key to understanding visit decisions [9]. Sheng and Chen take a more functional lens and broaden the scope - the visitor is looking to fulfill a need, of which personal identification is one.[18] In their survey of 425 visitors to Taiwanese museums, they identify five needs. In addition to personal identification - fun, cultural entertainment, historical reminisce and escapism are the other other needs. This prototype is informed by both researches. Different users have different self-identities ("I am classicist", "I am a modern art person") but the same user may be looking at different needs ("Today, I want a relaxed outing", "Today, I want to learn about a historic artist I admire"). Therefore for this prototype, the target interaction would build a user profile as of that *user session*.

## Recommender systems

*2.2.1 General approach.*

Sparse matrices is a key challenge to overcome in most recommender systems. Such matrices occur because if we imagine a matrix of different museums and users - and how each user rates each museum - information for most of the cells is not available. Recommender system have evolved to take several approaches to overcome this.

In the initial briefing, the stakeholders defined their problem statement as "[finding] a promising way to make [the website’s museums’ content] available to users is by recommending other pieces of content based on their browsing behaviour." Resnick [16] shows the potential of supplementing such user-based approaches with content-based methods, where the items are given a similarity score based between each other based on specific elements found in those items. This allows the recommender system to not only filter out negatives options for one user profile based on input from other numerous profiles, but also make positive recommendations based on similarity between the different options. This paper also surveyed other approaches in developing recommender systems within the museum domain and identifies that hybrid systems are most often used and perform better.

This prototype therefore takes a similar hybrid approach of supplementing user-driven *collaborative-based system* with *contentbased systems*

*2.2.2 Collaborative-based systems.*

Burke [6] explains that collaborative-based systems are the roots of recommendation systems, often because such user data for from online channels are easily available. This is also true in this case, as Google Analytics data is relatively easy to find and enrich over time.

But this also creates the *banana problem*. In a supermarket, bananas are a high volume product and bought by most visitors. A collaborative recommender would learn to recommend bananas to every visitor. In the context of this project, high traffic museums such as the Rijksmuseum in Amsterdam would be bananas. Diverting traffic to smaller museums and allowing discovery of relevant *hidden gems* was an explicit stakeholder need. Adding content based systems could create a recommender system that also highlights the pineapples in the Museum Association’s member list.

The collaborative system in this prototype can be further improved through two enhancements. Incorporating active learner systems by being selective with known ratings of a user [8] could be relevant given the nature of this system, which mixes historic and instantly processed preferences. Another possible improvement could be the implementation of a metric to measure similarity between user profiles [4], allowing museum.nl to collect data similar to Sheng and Chen’s study [18] in Taiwan and identify underlying user types to their website through vector decomposition.

*2.2.3 Content-based systems.*

Content-based systems forego user preferences and instead focus on information in the content of the items. In the initial briefing, the stakeholders also noted that "[museum.nl] has lots of content. All the museums and specific exhibitions are described." This prototype therefore approaches content based systems not only to overcome the challenges of the collaborative systems - but also explore ways to use museum.nl’s own data in improving its overall recommendation engine. Burke [6] also notes that this approach is often faced with the challenge of the data being hard(er) to obtain and often requiring a lot of processing and preparation.

This guides two design decisions in creating the content based systems.

First - content-type proposed should be readily available and not impose significant additional work on the Museum Association or its members. For instance, data on some standard features of museums already exists - e.g. location, museum categorisation, facilities etc. This should be used. In addition, some other data sources - such as text of museum descriptions - is a more natural part of the museums’ operating processes (new exhibitions and museums usually have descriptions on a website and therefore text is created as a part of regular operations). More complex data sources are technically possible. For example, clustering museums by the visuals of their actual art pieces is feasible through a convolutional neural network and clustering around features, colours etc. This was considered in initial ideation for this prototype as well. However, this involves museum staff creating image databases of their collections. Such databases already exists for larger museums like the Rijksmuseum through the Rijkstudio, but may create significant additional work for smaller museums as it is not part of their regular operations.

Second, ensembling models may provide a better design for continued enhancement of this system. In addition to the model accuracy benefit detailed further in this paper, ensembling allows a more *modular* approach to contentbased recommendation. As new data sources are prepared and processed, new content-based models can be added into this prototype without significant changes needed in the existing models.

*2.2.4 Ensemble methods.*

While Burke [6] suggests the possible combination collaborative and content based systems, Lü et al [11] state that current approaches in recommender systems suffer from a lack of unification and comparison across different solutions.

Ensembling different systems could be a solution. Aggarwal [1] explains that the power of such an approach lies in combining different data sources and allowing one system’s strengths compensate another system’s weaknesses. In this prototype, ensembling therefore allows integration of different data sources (user browser behaviour, standard features, text descriptions), model types (collaborative and content) and creates a modular approach mentioned earlier for continued improvement of the system.

*2.2.5 Recommender systems for museums.* Recommender systems have been applied to museums, providing implications to the design of this prototype. The Cultural Heritage Information Presentation (CHIP) project[[2]](#footnote-2) focuses on semantic enhancement of museums and their art. Pavlidis [14] reviews relevant applications of recommender systems within the museum domain and finds hat almost all model are used, and most often in a hybrid manner. Wang et al [19] compare different recommendation strategies and identify optimal approaches for the development of a user profile for effective recommendation results. Again, the use of hybrid systems appear more successful - although the content-based methods scored significantly better than the collaborative-based methods.

Keller and Viennet [10] focus on the items inside the museums to generate recommendations. They keep track of items liked by users to recommend other items in the museum to the visitor. Results look promising as a significant number of users are observed actually walking towards a suggested item and subsequently liking it as well. This also creates a feedback loop creating meaningful number of cascading recommendations.

After determining that potential visitors of museums are often overwhelmed by the amount of information available, resulting in loss of interest or putting insufficient care and effort in selecting a recommendation that fits, Benouaret and Lenne [3] suggest a solution in the form hybrid recommender system to be tested in a museum - focusing on multiple recommendations, rather than just one.[2].

*2.2.6 Other relevant works.*

In our choice of recommender system approaches, while most sources noted a strong need of thorough and proper evaluation, it is best put into words by Shani and Gunawardana [17], who dedicate a chapter in their book on the subject, noting the need for and make suggestions on suitable validation methods. Our decisions and approach is explained in the *methodology* and *results* sections. Lastly, a mention should be made of knowledge-based systems. These systems make recommendations by focusing on user-specific requirements. [1] While this is also a valid approach in recommendations, this paper decided not to focus on this. Firstly, a filter feature already exists on museum.nl - allowing the website to interact with users with very clear requirements. Secondly, most museums contain a large number of items - therefore a knowledge based system may continue to find a large number of museums relevant for a user even after requirements are specified. Finally, the cold start problem remains. A new user might have insufficient data to generate matching recommendations and hence a starting interaction (such as with the filters on the website) is still needed. This prototype takes advantage of the greater context of the museum visitor and recommendation systems to implement an interface that can do something novel.

# METHODOLOGY

This research was carried out according to the steps as they are described within the Design Thinking Methodology [15]. This concerns a solution-based approach in which the problem context is first explored by empathising with both the user, and stakeholders. Secondly, the identified requirements are defined and documented in the form of system and interaction requirements. Subsequently, as many solutions as possible are produced during an ideation phase leading to a number of user-centred solutions. The most promising solution is then worked out into a prototype during the prototyping phase. Finally, the prototype is tested and improved during an iterative validation process in which multiple stakeholder and user interactions take place. This process should eventually result into a solution that answers the following research objective *"This project aims to realise a recommendation system that allows museum.nl visitors to get a personalised recommendation on which museums or exhibitions to visit in an engaging way. Consequently, the system hopes to establish increased engagement among its users"*.

In order to sufficiently carry out this research, a clear distinction was made between the two main aspects within this study. As this project contains an interaction design part, concerning the required interactions and usability of the system, and a system design part, related to the required recommendation models to establish the output of the system, the research questions have been divided as such.

*Interaction design*

* Which interface do museum.nl users prefer while providing input to any recommendation system?
* Do the proposed interactions lead to an increased degree of user engagement?

*System design*

* To what extend is the proposed recommendation model sufficiently capable of generating a desired recommendation to the user’s input?

For validation of the concept, both formative and summative research were performed. Formative research, research before the implementation of the system, was used to take insights regarding the usability of the proposed interface into account at an early stage. Furthermore, it made validating the theory behind the proposed interactions to possible to some degree, while still shaping the final solution. Consequently, this research supported decision making during the process and helped maximising the user-centricity of the system. Summative research, research after implementation of the system, was used to validate the final shape of the implementation, and to support its quality assessment. Furthermore, the latter helped assessing the acceptance of the presented system.

In order to answer the research questions regarding the interaction design, formative research was performed. During this qualitative research, a comparative usability test combined with thinkingaloud was carried out with a sample group of five participants. The chosen sampling level within this research is typical case sampling. This choice was made because of the assumption that users who fit the typical museum.nl visitor persona will be able to provide more valuable insights than participants who deviate from this profile. The defined sampling criteria are; 1) within the age category of

20-60 years, 2) participants are in no way involved in the study,

3) holding a museum card, or frequently visiting a museum and 4) resident of the Netherlands. Respondents were asked to schedule a museum visit by making use of the two proposed interfaces. Interactions and audio were recorded and striking observations documented. After the usability test the participants were asked to perform a survey in order to validate the prototype. The formative research has been elaborated in more detail in the annexes (see Annex B).

In order to answer the research question regarding the system’s design, summative research was carried out. During this research, a more quantitative approach was used by means of an interactive experiment. During this experiment, a sample group of 15 respondents were asked to perform the same task as during the formative research, only now with the implemented system, equipped with the back-end recommendation models. The same sampling level and criteria were used as during the formative research and afterwards respondents were also asked to complete a survey. Within this survey, participants were asked to rate their top four recommendations via a likert-scale of one to five. A recommendation is considered relevant if a score of 4 or 5 is provided. Average precision scores of the recommender are calculated for the top, first 2, 3 and 4 recommendation. A summary of the results are detailed in the results section and the full panel of responses is in the annexes. (see Annex C).

This research was carefully carried out by following the steps outlined in a predefined validation scheme. The latter can also be found in the appendices (see Annex A).

# SYSTEM AND INTERACTION REQUIREMENTS

This section will start by outlining the potential user of the recommender system so that the needs and context of this user can be described. Furthermore, the requirements that arose from preliminary research and empathising with the user will be elaborated into a sequence of system requirements.

## User persona

This project aims to realise a recommendation system that is able to establish increased engagement among the museum.nl users. In order to design a system that meets this objective, it is important to empathise with the end user. Therefore, this section will describe a typical user persona, by making use of the elements described within the empathy mapping technique [7]. Although the museum.nl website aims to serve a significantly larger group from a broader context than the persona described below, the decision was made to describe only one, with specific demographics as the assumption was made that the typical use case, and the needs of the user are similar to most within the defined target group.

Laura is a 26-year-old journalist living in Haarlem, a city outside of Amsterdam in the northwest of the Netherlands. Her friends describe Laura as calm, with an open character and extremely interested into a wide variety of topics. Hence, her choice to become a journalist. As she lives just outside of Amsterdam, a city with a great number of museums, she holds a museum card subscription for already five years. The main reason Laura likes to visit museums is because she likes to behold and enrich herself through the artworks and showpieces available within the Dutch museums. Although Amsterdam is the city she visited the most museums in, she is becoming increasingly aware that in her own city, but also in the rest of the Netherlands, there exists an incredibly rich offer of museums and exhibitions. Consequently, she decided that she wants to visit at least two museums with varying showpieces and artworks every month.

To orient for her next visit, she uses the museum.nl website. She uses the website on both her desktop, and mobile phone. Laura visits the desktop website in case she schedules a visit upfront, and the mobile website when she incidentally decides to visit a museum already on the way. She estimates that the distribution of her website visits is about 30% desktop and 70% mobile. The website itself allows her to quickly find the museums her museum card subscription provides access to. However, she does not find that the current content, nor the way of recommendation, allows her to easily, and quickly find the museum with the artworks and showpieces that correspond with her interest for that specific visit. The main reason for this opinion is because the content on the website is on a museum level and, at first sight, does not provide any insight into the artworks or showpieces available within that museum. Consequently, Laura feels she is inclined to frequently visit the same, more known museums because the museum.nl website does not sufficiently allow her to base her orientation on the artworks and showpieces within that museum.

Laura’s wish is to have a quick, and easy to use recommendation feature that allows her to quickly select artworks or showpieces that fit her interest for the specific visit she is scheduling. This way, she would be able to also find museums that do not necessarily come to mind at first thought but do have exhibitions she would like to visit. Consequently, Laura is sure she would feel more engaged by the museum.nl website.

## Requirements

The user persona of Laura shows a general experience potentially many museum.nl visitors can relate to. Namely, the most common way of orientation is via museums instead of the preferred orientation of discovering artworks and showpieces, feeling the need to discover and visit a greater variation of museums with varying artworks and showpieces but not having an easy and engaging way to find these museums and feeling inclined to visit certain popular museums more frequently than others. Based on preliminary research (see section 2), the stakeholder context (see section 1.1) and empathizing with Laura (see section 4.1) the following system requirements derived:

R1: The system’s content and interactions should increase user engagement.

R2: The system should provide a personalised recommendation as output after getting showpieces as input.

R3: The interactions should be aligned with the current museum.nl website.

R4: The interface and its content need to be intuitive and understandable.

R5: The output needs to correspond with the interest of the user.

R6: The system needs to allow the user to also discover smaller, less popular museums.

Both the interaction design, as the system design will be based on the defined requirements above.

# INTERACTION DESIGN

As described within section 3, this study was carried out by making use of the Design Thinking methodology [15]. This methodology concerns an iterative process in which the presented solution will be shaped and improved based on interim tests and feedback rounds with stakeholders and potential end users. The following section will outline the required devices, the two interaction models that were used during this study described from the user’s point of view and visual representations of the designed user interactions.

## Devices

This section will provide a brief explanation of the devices on which the system will interact with its users.

*5.1.1 Mobile device.*

As the recommender system will have to be implemented on the existing museum.nl website, the system and its interactions will be accessible via internet browsers across a variety of devices. After empathising with the end-user, elaborated within the user persona of Laura (see section 4.1), the assumption was made that the majority of users will interact with the recommendation system via a mobile device. Therefore, the system’s interface and corresponding

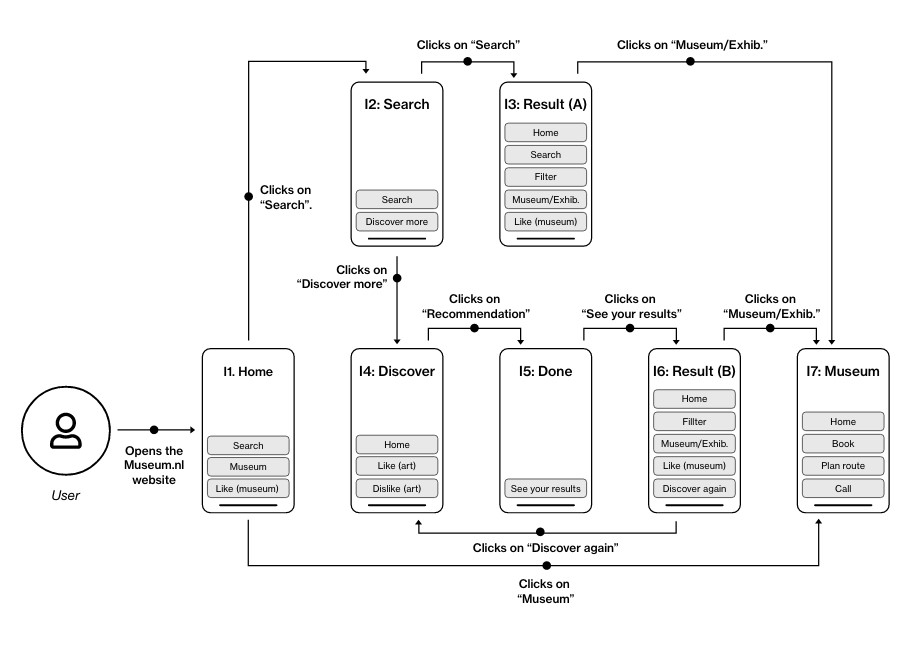


Figure 1: Initial interaction design overview

interactions were designed using a mobile-first focus. In order to optimally interact with the system, mobile phones should be able to display graphical user interfaces (GUI) and be equipped with touchscreen sensors.

*5.1.2 Desktop.*

Although desktop devices did not have the focus during the implementation of this project, the system and corresponding interactions should still be possible and available on desktop devices as the museum.nl website is also available on these types of devices. In order to interact with the system in an optimal way, desktop computers should be connected to a monitor, a computer mouse or track pad in order to see and control the graphical user interface.

## Interaction models

As the system’s interface and interactions were shaped during an iterative process with multiple stakeholder engagements, two different interaction models were designed during the implementation of this system. The first model arose directly after the ideation phase and was a logical corollary of the concept that came about during the brainstorm sessions. The second interaction model was created after presenting the first model to the stakeholder. During this engagement, the stakeholder indicated an alternative interaction model would be desired in order to discover whether or not the concept would also fit the existing interactions and design of the website. In order to optimally investigate which interaction model is most in line with the needs of the user, both were taken to the validation phase (see section 7). This section will explain both interaction models that were used during this study.

*5.2.1 Initial model.*

Figure 1 shows an overview of the initial interaction model, developed directly after the ideation phase. The first step the user takes is opening the museum.nl website. The interface the user arrives on is processed into Figure 1 as I1. This concerns the existing home page of the museum.nl website (see figure 2). The actions the user is able to perform are pressing the “Search” icon, a “Museum” content card and “Like” the museum that corresponds to that particular content card. In case the user chooses to click on the “Search” icon,

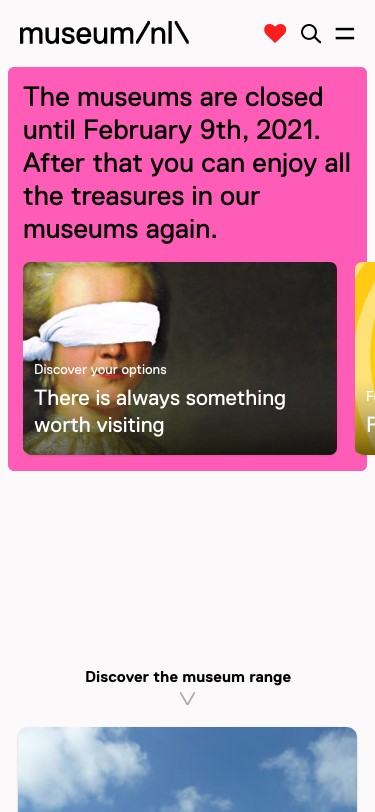


Figure 2: This image displays the Home screen of the museum.nl website.

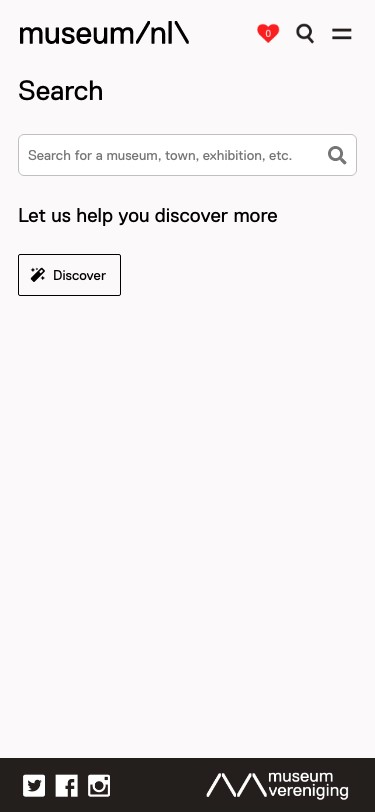


Figure 3: This image displays the Search screen, the button that leads to the TindArt process is positioned under "Let us help you discover more".

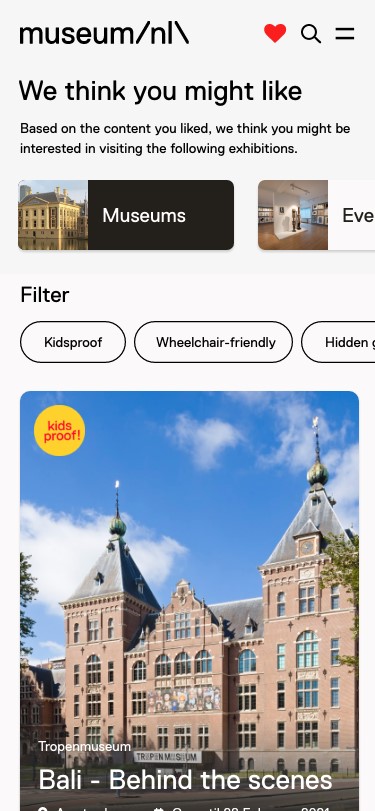


Figure 4: This image shows the Results screen, filtering is possible at the top after which the resulting museums will change below.

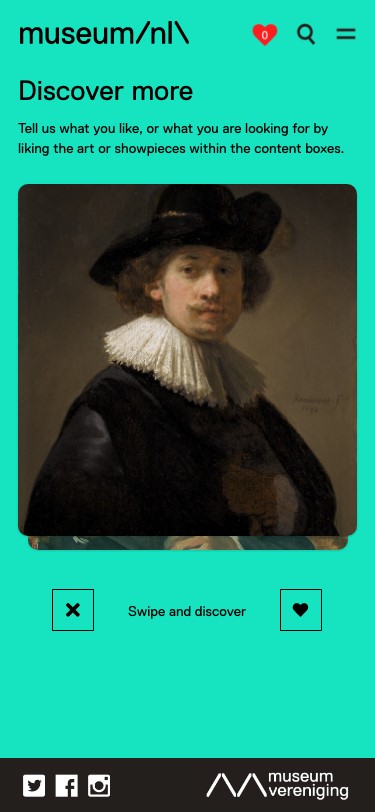


Figure 5: This image shows the Discover interface, the image can be swiped left or right. The cross and heart buttons are positioned below.

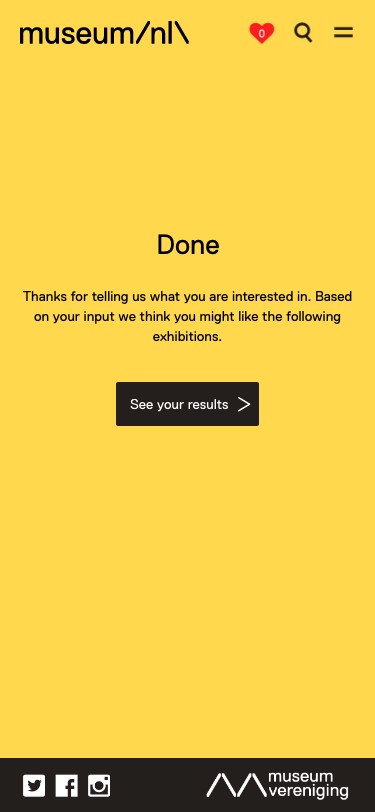


Figure 6: This image displays the screen that appears after the user has liked the desired number of artpieces.

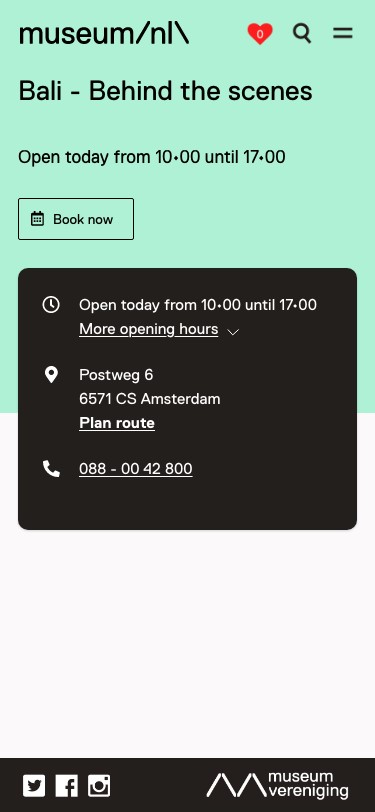


Figure 7: This image displays the Museum page, providing alltherequiredinformationandfunctionalitiesauserneeds to schedule their visit.

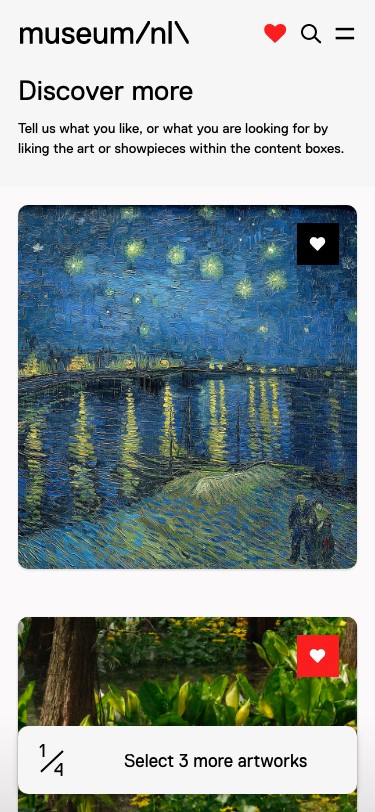


Figure 8: This image displays the alternative discover page which is more aligned with the current museum.nl website.

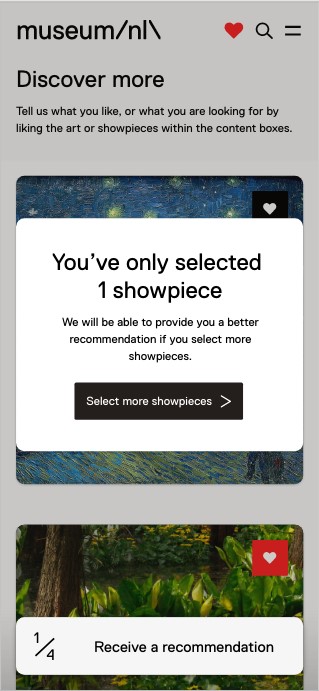


Figure 9: This image shows the modal which appears after the user chose to click on the button before the optimal number of showpieces is selected.

the user is directed to I2. On this screen, the user is prompt to

either insert a “Search” command by using the search input field, or to discover more by pressing on the “Discover more” button decorated with a magic wand (see figure 3).

If the user chooses to enter a search command and clicks on the

“Search” button, he or she is directed to I3 (see figure 1). The user is now able to view a sequence of museums which the system recognised as corresponding to the previously inserted search command (see screen 4). On this page, the user is able to return to the “Home” screen, click on the “Search” icon to enter a new search command, “Filter” on the categories: 1) museums and 2) exhibitions, and the features: 3) kidsproof, 4) wheelchair-friendly and 5) hidden gems. Furthermore, the user is able to click on content cards representing museums and exhibitions (“Museums/Exhib.”) and “Like” the corresponding museums or exhibitions.

In case the user chose to click on the “Discover more” button on I2, the user would be directed to I4 (see figure 1). This screen explains the user how the feature works by displaying the text “Tell us what you like, or what you are looking for by liking the art or showpieces within the content boxes” (see figure 5). In order to meet R1, R2 and R4 (see section 4.2), the interface was partially based on the popular dating application Tinder[[3]](#footnote-3). On this page, the user is asked to either “Like” or “Dislike” a pre-determined sequence of artworks and showpieces. The user can perform these actions by either swiping left or right, or by clicking on the cross and heart button underneath the artwork or showpiece. The latter, was added in order to make the interface more self-explanatory. The assumption was made that users understand the meaning of both a heart symbol and a cross symbol. After the user has liked a total number of four showpieces, the user will be automatically directed to I5 (see figure 1). The choice to automate this interaction was made as this will allow the system to force the user to select a predetermined number of artworks as input to the system. This might be advantageous in monitoring the quality of the recommendation. On I5, the screen will explain the user he or she performed all the required steps and functions as an explanatory guide to the next step by displaying the text “Thanks for telling us what you are interested in. Based on your input we think you might like the following exhibitions.” (see figure 6). The user is able to resume to I6, by clicking on the button “See your results” (see figure 1).

The user is now directed to exactly the same interface as the previously explained I3 (see figure 4). However, now the displayed museums and exhibitions are recommended based on the artworks and showpieces the user provided as input during the previous step. By allowing the user to like and dislike artworks, the system hopes to meet R5 (see section 4.2). Furthermore, by purely shaping a recommendation on related artworks and showpieces, the system hopes to meet R6.

In case the user is not satisfied with the recommended sequence of museums and exhibitions, he or she is able to click on the “Discover again” button after which he or she will be redirected to I4 to start the process over again. In case a user decides he or she likes one of the recommended museums and or exhibitions, he or she is able to click on the content card displaying the museum or exhibition (“Museum/Exhib.”). By clicking on either of the displayed

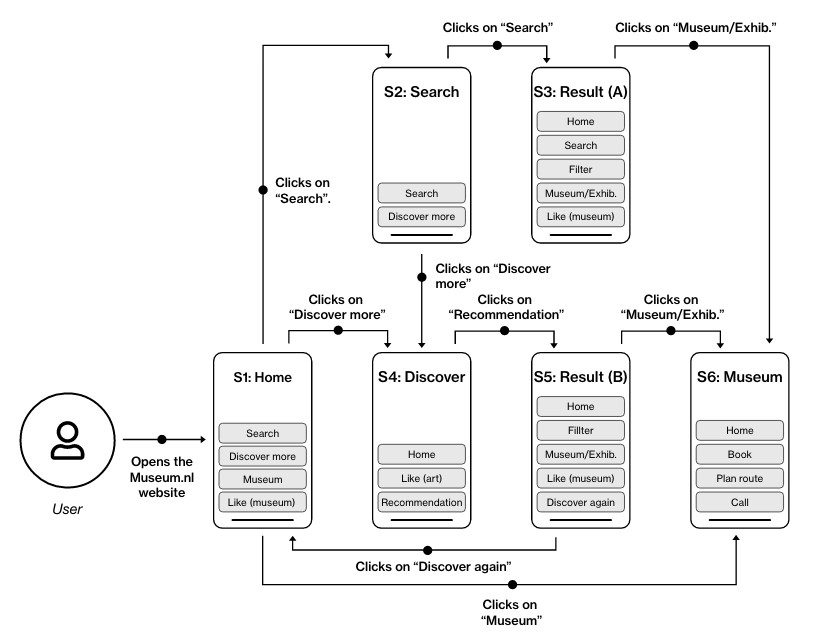


Figure 10: Alternative interaction design overview

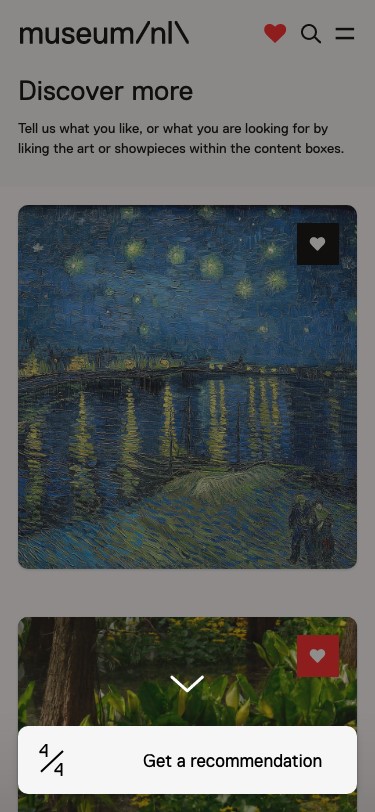


Figure 11: This image shows the focus interaction, guiding the user towards the "Get a recommendation" button.

content cards, the user will be directed to I7 (see figure 1). On this screen, the user is able to “Book” a visit, “Plan route” to navigate to the particular museum or “Call” to the museum (see figure 7).

*5.2.2 Alternative model.*

Figure 10 shows the alternative interaction model, developed after an intermediate stakeholder engagement with Q42. During this meeting, the stakeholder indicated they would like to see a version that is more aligned with the current website, instead of an almost entirely similar interaction as the Tinder application. Although S1, S2, S3, S5 and S6 (see figure 10) are exactly the same as I1, I2, I3, I6 and I7 (see 1), the user is now also able to access the discover page via I1, which is the home screen. When the user enters I4, he or she is explained how the system works in exactly the same way as on S4 (see section 5.2.1). However now, the user sees a scrollable sequence of artworks which are positioned one below the other as a list (see figure 8). The reason behind positioning the artworks as a list, is because now the user will be able to view all the artworks before he or she chooses which ones they want to “Like”. Furthermore, the user will also no longer be able to “Dislike” artworks or showpieces as this interaction also does not exist on the current museum.nl website.

After the user chooses to “Like” the first artwork on S4, a button

appears at the bottom of the screen. This button explains to the user how many artworks he or she selected, and how many more artworks the user should select in order to receive an optimal recommendation by the system (see figure 8). This interaction should make sure the user feels informed during the entire process. In case the user chooses to click on the button before the desired number of inputs are provided to the system, a modal appears explaining the user that he or she only selected a limited number of input fields and that the system will be able to provide a better recommendation if the user selects more showpieces (see figure 9). This interaction was added after evaluating the results of the usability test. One of the insights was that some users chose to directly click on the button after appearing.

After the user has selected the desired number of liked artworks, a focus was added to the screen to make the user aware he or she should now click on the button (see figure 11). Furthermore, the text inside the button changes into “Get recommendation” accompanied with the magic wand icon that was also visible on S2 (see figure 3). The focus was added because during the usability tests it became clear that many users did not immediately notice the change of the button at the bottom. Therefore, the focus was added to provide more visual guidance. After clicking on “Get recommendation”, the user is directed to S5 after which exactly the same process follows as described within section 5.2.1.

All the interactions described within this section are also available in the form of a clickable prototype. This prototype can be accessed via the following link: https://xd.adobe.com/view/0a17a90ffeb1-4896-9b97-693b84fbe05a-5f44/?fullscreenhints=off.

# SYSTEM DESIGN

Section 2.2 describes that this prototype will contain a hybrid approach of supplementing user-driven *collaborative-based system* with *content-based systems*. The following section describes the developed models, the assembling of the models and the final implementation.

## Quantify similarities between museums

For the recommender system to work it must know similarity between museums. (To illustrate, if Laura says she likes museums in the set A, B, C and D, the system needs to know if E, F G and H are similar to this set). Cosine similarity is a common metric used in such systems [14]. Museums are converted to vectors in a multi-dimensional space and the cosine of the angle between two museums - i.e. their vectors is calculated. This is the dot product of the normalized vectors and is a value between 0 and 1, where 0 means no similarity and 1 means completely similar. The outcome of the function is a similarity matrix with on both axes the museums. The similarity matrix contains cosine similarity values between all museum pairs and is used to suggest most similar museums. An example of such a matrix is in table 12

A · B

cos*𝜃* = (1)

∥A∥ × ∥B∥

## Collaborative filtering

|  |  |  |  |
| --- | --- | --- | --- |
| Similarity\* | museum1 | museum2 | museum3 |
| museum1 | 1 | 0.54 | 0.18 |
| museum2 | 0.54 | 1 | 0.61 |
| museum3 | 0.18 | 0.61 | 1 |

The collaborative-based model is a rather simple implementation of the data gathered by museum.nl through Google Analytics in the period from 27-09-2020 til 21-12-2020. The likes a user has given Figure 12: Example of a similarity matrix

on the different events, museums and artworks on the website were extracted from all the data, and each of those possible objects to like was one-hot encoded for each individual user. The result was a table of 665 objects to like and 4933 unique users and their recorded likes. No additional processing was done. To be able to give scores, a similarity matrix is necessary of all the items and how closely they are related to other items, and this is done through the cosine similarity, through a simple absolute function magnitude. To then recommend a list of museums, a user’s preferences feature vector, in the form of a simple list of liked museums, is taken as input. The feature vector is then created by creating a list of likes and unknowns for all likeable objects. The dot product between that feature vector and the similarity matrix is calculated, and the descending sorted output is the list of most similar museums.

## Content-based filtering using features

The first content-based model includes standard features of a museum as collected by the Museum Association. Data provided by the stakeholder at the start of the prototyping (in ’musea.csv’) contained categorization of all museums (into 5 categories), their location (by latitude, longitude, city and province), kids-proof mark and available facilities (for example wheelchair access, parking, proximity to a train station etc.). For the location, this prototype chooses provinces over city for two reasons. Adding city-level detail creates dissimilarity between museums that should be considered similar in this domain (e.g. 2 history museums in different towns in Friesland). Also, users may tend to visit museums in their provinces. All variables were combined into one field, further named ’totalfeatures’. Details of the features can be found in table 16.

|  |  |  |
| --- | --- | --- |
| Nr. | Feature name | Description |
| 1 | Main cate-  gory | Museum has a main category from the 5 values used by the Museum Association. Values: ’BusinessScienceTechnology’, ’History’, ’Art’, ’NaturalHistory’, ’Ethnology’ |
| 2 | Sub category | Museum has a sub category from the 5 values used by the Museum Association. Values are ’BusinessScienceTechnology’,  ’History’, ’Art’, ’NaturalHistory’, ’Ethnology’ |
| 3 | Province | Museum Province in which the museum is located |
| 4 | Kids type | If a museum is kidsproof. Derived from the latestmuseumkidstype |
| 5 | Facilities | Facilities at a museum |

Figure 13: Standard features one hot encoded in the first content based model

The model creates a cosine similarity matrix between museums based on the ’totalfeatures’ feature. This will generate, as described in section 6.1, a matrix with similarity scores. A museum with the same category and the same facilities as the input will have a high score. The model will look up the highest score for each input, sums and takes the mean of the scores. The output of the model is an ordered list on the mean of the scores, with at the top the most similar museums, and at the bottom the most dissimilar.

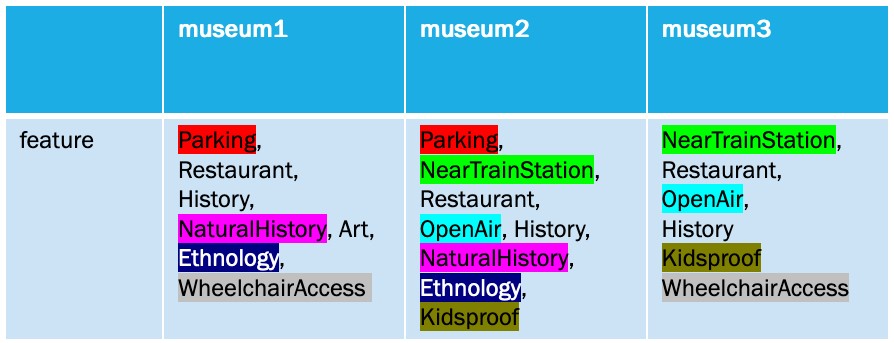


Figure 14: Illustration of standard features based recommender

## Content-based filtering using museum description

The second content-based model uses text in the descriptions and meta descriptions of the museums. An assessment of the corpus provided (in ’musea.csv’) and text on museum.nl indicated that the description and meta description was usually the same as the text on museum pages. The description of a museum describes the content of a museum, for example, the description of a war museum contains terms like ’WW2’. Museums have a higher similarity score when both contain the same terms.

To develop the corpus, main and meta descriptions were merged per museum into a field further called ’totaldescription’. To increase the quality of the algorithm, pre-processing of the description was needed. This included stopword and punctuation removal through relevant Dutch language libraries, as well as other context-specific stop words such as *musea, tentoonstelling* etc. A TFIFD vectorization on the corpus created the dictionary for the text and vectorized the museum texts. This prototype could be improved by enhancement of the corpus which is discussed in future work.

|  |  |  |
| --- | --- | --- |
| Nr. | Feature name | Description |
| 1 | Description | The main description of a museum |
| 2 | Meta descrip-  tion | The meta description of a museum, also called the ’Read More’ on the website |

Figure 15: Text corpus used in the second content based model

Finally, the model creates a cosine similarity matrix between museums based on the ’totaldescription’ feature. This will generate, as described in section 6.1, a matrix with similarity scores. A museum with the same terms in its description as the descriptions of museums in the input will have a relatively high score. The model will look up the highest score for each input, sums and takes the mean of the scores. The output of the model is an ordered list on the mean of the scores, with at the top the most similar museums, and at the bottom the most dissimilar.

## Ensembling recommendations

After the three separate models have generated their recommendations, the ensembler uses these outputs as its input, to create the final set of recommendations to be given to the users. The approach for our implementation is a positional ranking system, rather than using the scores of the similarity of the inputs to the outputs, from the three separate models. If a museum was present in all the outputs of the models, the ensembler will add their positions together. For each time a museum is not present in one of the other sets, an artificial score of ten was added, since each model would output ten recommendations. After all inputs have been processed, the results can now be sorted ascending, resulting in the average most relevant recommendations being returned first.

## Implementation

TindArt implementation was split into a front and back end (FE, BE) with integration (I). The overall logic of the prototype is detailed in the figure below.

|  |  |  |
| --- | --- | --- |
| Nr. | Step | Description |
| FE1 | FE shows a sam-  ple | User is shown a collection A of museums and items in the front end |
| FE2 | User ’likes’ | User likes a subset B from the collection A |
| I3 | Array shared | Step 2 creates an array of preferences that is shared from the front end to the back end |
| BE4 | Recommender systems choose | Based on similarity matrices, the recommender system make a short list |
| BE5 | Ensembler votes | Ensembler makes a final shortlist through weighted votes |
| I6 | Array shared | Array of final results shared from BE to FE |
| FE7 | Results dis-  played | User is shown recommendations based on results obtained |

Figure 16: Overall process flow of TindArt

This process flow allows integration of the BE to multiple FEs. Several web applications or features in native apps could be created through the recommender logic.

*6.6.1 Integration API.*

The Python API allows requests for recommendations via an exposed endpoint through a HTTP POST request. The input for the API is an array (or list) with the *translationSetId*s for the museums in subset B. The response from the API is in JSON with data on recommended shortlist from the ensembler.

The stack chosen is Angular on the client-side and Python with Flask on the server side. "Python, Flask[[4]](#footnote-4), and Angular[[5]](#footnote-5) form a great stack to build modern web applications." [5].

For museum.nl, it would technically feasible to call the recommender system from multiple locations. The TindArt application is one location. Another location could be the "Discover More" content cards mentioned earlier. If a user likes 3-4 museums or exhibits in a session, sending a POST request to the API with the *translationSetId*s of the liked pages could return recommendations to personalise the content cards. The interactions are illustrated below.

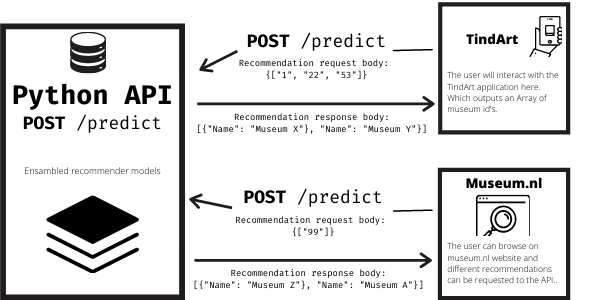


Figure 17: Integration through Python API with multiple FEs

*6.6.2 TindArt FE.*

The front-end -or client-side- basically is the face of TindArt. This is where all the user interaction takes place, and where the user will see the outcome of the predictions. The font end has been built as a Single Page Application (SPA). By using this technique the use of the app will feel closest to a native, app-like experience even though under the hood, it still is a web application. The application is a standalone app, build in the JavaScript framework called Angular (version 11), and as in the spirit of a Single Page Application the entire application runs on a single HTML page, which will dynamically change the DOM contents based on the user interaction. This way, the user does not have to re-render pages as they are moving through the application, giving them a smooth experience. By using this framework it would be easy to eventually plug it into the ecosystem of the current museum.nl website without the need of complex hosting platforms.

The front end itself does not contain any computational logic, and it’s purely used to collect input from the user, send this to the API and then retrieve the recommendations back in JSON format from this API to display it to the user in a user-friendly way. As mentioned the more heavy computational tasks are not performed in the front end, for these model calculations as suggestion Museums and artworks rely on a more powerful server and should not be calculated on the client-side.

*6.6.3 Back End.*

The back end -or server-side- of the application was fully built in Python version 3.8 and exposes an API using Flask. According to Krebs [5] "Applications built with Flask are clearly lighter when compared to Django counterparts. As such, Python developers usually refer to Flask as a microframework."

After this decision was made, the back-end needs to be able to allow our Angular front end application to call this API and let it do recommendations for it, the API was configured to allow Cross-Origin Resource Sharing (CORS) requests. Via a simple HTTP POST request from the front end, including the liked artworks or museums from the user during, for example, the TindArt flow, the recommendation system spits out a recommendation in under one second.

The recommendation consists of a list of 5 Museums. The API takes multiple recommendation functions and then uses a technique called ensambling to take the average of the different functions. This average is going to be used for the final recommendation. This interaction provides the solution for implementing the R2 requirement: "The system should provide a personalised recommendation as output after getting showpieces as input." Since the API uses a recommender-system the API was run on a local development machine and does not need a heavy duty server environment.

# RESULTS

The research question that guided the development of this proof of concept was to identify improvements that could be implemented on museum.nl that increases user engagement while recommending relevant museums. Beyond validation with Q42 and museum.nl during the iterative design process, a sample group of users were asked to interact with the front end and rate the top four recommendations on a likert scale.

## User engagement

The overall feedback on usability is generally positive, with scores higher than 4. Users found the interface intuitive and understandable - and see themselves using a tool like TindArts when scheduling a visit to a museum. These results should be considered a focus group validating the proof of concept. During production, larger scale testing (e.g. A/B testing) could provide further insight for instance, do users offered TindArt have a lower bounce rate or higher page views per session.

|  |  |  |
| --- | --- | --- |
| Nr. | Qualitiave questions asked | Score |
| 1 | During the test I had the feeling I was in full control over the system. | 4.00 |
| 2 | The system’s interface and its interactions are intuitive. | 3.88 |
| 3 | The system’s interface and its interactions were understandable. | 4.35 |
| 4 | I visit a great variation of museums with varying artworks and showpieces. | 3.65 |
| 5 | I see myself using this tool when scheduling a visit to a museum in the future. | 4.12 |
| 6 | The recommendation is a logical output considering my liked artworks | 4.00 |
| 7 | The recommendation output corresponds to my interest. | 4.06 |

Figure 18: User scoring on qualtitative questions and user experience

## Relevant recommendations

Users also rated the first four recommendation on a likert scale. The rationale for limiting this to the top few results is that most users will not scroll through the whole list of recommendations. Therefore the top results must be considered relevant, to keep users engaged and encourage them to click through to booking a visit.

The first recommendation had a mean average precision of 76.5 percent - which is the probability that the top ranked recommendation from the ensembler was considered relevant by the user (and ranked 4 or 5 out of 5) 76.5 percent of the time. There is a drop in the mean average precision when the top three or four recommendations are considered. Potential future work to further improve the models and build a feedback loop to the ensembler so it continues to improve how it adjust weights and create high relevance recommendations are detailed below.

|  |  |  |
| --- | --- | --- |
| Nr. | Number of recommendations | Mean average precision |
| 1 | First recommendation | 0.765 |
| 2 | Top 2 | 0.662 |
| 3 | Top 3 | 0.631 |
| 4 | Top 4 | 0.631 |

Figure 19: Mean average precision of recommendations

# DISCUSSION FUTURE WORK

## Discussion

*No final usability test*: An extensive formative usability test was performed after which several opportunities for improvement were identified. Although these tests showed promising results already, the improved and finalised frontend prototype was not tested on whether or not all of the initial shortcomings were resolved after making the adjustments to the first version. Time constraints associated to this project simply did not allow this final review.

*Participants’ experience level*: All our usability tests were conducted with respondents who held a museum card or confirmed they frequently visit museums.(see Annex C) This choice was made deliberately in order to test the concept with potential early adopters. However, this limited sample group can limit gathering insights from people who have less experience with the existing interactions on the Museum.nl website.

*Participants’ demographics*: The demographics of the test participants show they were well-divided in terms of place of residence and gender but not age.(see Annex C) The research on international visitors are also limited in this research.

*Used recommender models*: The recommender models used in this implementation are, by design, relatively naïve: not much has been done to optimise the resulting recommendations. Therefore, it will be necessary to work on refining those before this implementation is being included on the museum.nl website. The concepts of the individual models and the ensembler are however of quality and do proof that they can produce recommendations which are considered valid by users.

*Varying number of inputs*: Due to time constraints, no extensive research was performed into the ideal number of inputs to the recommender system. Whereas various number of inputs could very well influence the quality of the recommendation in a significant way. If museum.nl were to implement the proposed data system on their current website, it would be advised to research what this ideal number would be so that the user can be guided to provide this number of inputs. The presented interaction model allows the system to provide this guidance.

*Expending the data set*: Lü et al [12] argue that most recommender system suffer from a self-reinforcing pattern of rich-get-richer: "items that were popular in the past tend to be served to even more users in the future." The obvious, yet easily overlooked solution is to not filter data too thoroughly and leave the information ecology as much as is. A related common issue is that of recommendations becoming too general and no longer personalised; if every users likes the Mona Lisa because it is so popular, it will also be recommended to all new users.

## Future work

While the results on this proof of concept shows that an interesting interaction can be created that allows session-specific recommendations on museum.nl, there is significant room for further improvement in the back end as the prototype is moved to production. *Data enhancement*: Data enhancement will improve the similarity matrices in the recommender models. 2 approaches are possible.In the collaborative model, a data pipeline that feeds Google Analytics data to the recommender model could create a continuous improvement process adding more user data over time and improving this model For the text based model, other text sources were also considered. Text scraping from the individual museum websites, collecting any email communications sent, reviews and information from Tripadvisor and google maps could also be added as a continuous data enhancement pipeline that adds to the corpus associated with each museum

*Language model*:This proof of concept is built on a simple Dutch-only model. The dictionary is built from the corpus of text available in the provided dataset. Using a larger language model such as RoBERTa (or RoBBERT for Dutch) will allow the use of English and Dutch text. Their superior understanding of word meanings can also create much more precise similarity structures within museums. The option was considered in this proof of concept but discarded due to implementation limitations at this stage

*Ensembler weights*: The ensembler has equal weights allocated to the votes of the three models. This is currently a hyper-parameter. A future improvement could convert these to learnt parameters. A potential experiment set up is as follows. A training set may be created by storing the arrays of recommendations from the three recommender models (A), the array passed from the ensembler in the BE to the FE (B) and *translationSetId*s of the item a user clicks through or likes in the recommendations (C). The ensembler uses a weight matrix W (currently [1/3, 1/3, 1/3]) to convert A to B. As interaction data is collected, W can be learnt so that A X W = B where C is the top ranking element of B.

*Adding more models*: A By using an ensembler, new content models can be easily integrated into this data system.

*Target audience expansion*: Even though the initial target group is 20-40 years old, two of the older test participants (57 and 63 years old) who also expressed strong interests in this feature. Research and tests of expanding the target audience to a larger age range can increase the return-oninvestment of TindArt. The validation focuses on domestic visitors and museumkaart holders. Filters such as location, hidden gem and kid-proof are used to increase their engagement. However, international visitors can also benefit from TindArt and current filters. Additional filters targeting international visitors when international tourism recovers can be valuable too.

*Accessibility measures*: There are a few feedback from testing participants on text not readable with page background colour. The overall accessibility such as text-background contrast and touch target size can be improved.

*Large-scale quantitative validation*: There are millions of online visitors on Museum.nl every year. It can be valuable to use methods such as AB testings, web page heat-maps and on-page surveys to collect the large amount of user data to validate the concept and to seek improvements.

# CONCLUSION

This study aimed to create a data system for museum.nl that was both a recommender engine and also improved user engagement on the site. The goal was to help the stakeholder Q42 and their customer - the Museum Association of the Netherlands - to increase interest in museum visits through the museum.nl website. Through the use of provided design principles, and validation with Q42, a prototype was created that gave users museum recommendations based on a set of defined preferences as shared in an interaction and is thus our recommendation for a solution to museum.nl and Q42’s problem of user engagement through personalised suggestions. This study expanded beyond the behaviour model of building user profiles only through historical browser data only. Recommendation quality drops with general approaches - especially in a context where drivers of museum visits could change day by day. The user interface design was improved through stakeholder feedback and thinking about integration with the existing website. The improved user interface was further validated through user acceptance testing. To generate the recommendations, models that created similarity matrices between museums were developed. One collaborative model (i.e. using inputs from multiple users) and two content models were developed. Results of the simple models are then ensembled; literature shows it provides greater accuracy than development of complex models in other practical tests. Ensembling also allows further addition of new models and approaches - for instance models created by other teams in this course. The focus of this study was to conceptualise and user-validate this interaction and recommendation approach, and has been left at a successful proof of concept stage. There is still significant scope for future work if this is taken to production by Q42, including data augmentation and weight learning in the ensembler.

# REFERENCES

[1]• Charu C. Aggarwal. 2016. *Recommender Systems The Textbook*.

1. Idir Benouaret and Dominique Lenne. 2016. Personalizing the Museum Experience through Context-Aware Recommendations. In *Proceedings - 2015 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2015*. 743–748. <https://doi.org/10.1109/SMC.2015.139>
2. Idir Benouaret and Dominique Lenne. 2017. A Composite Recommendation System for Planning Tourist Visits. In *Proceedings - 2016 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2016*. 626–631. [https://doi.org/10.1109/ WI.2016.0110](https://doi.org/10.1109/WI.2016.0110)
3. Jesus Bobadilla, Fernando Ortega, Antonio Hernando, and Javier Alcalá. 2011. Improving collaborative filtering recommender system results and performance using genetic algorithms. *Knowledge-Based Systems* (2011). [https://doi.org/10.](https://doi.org/10.1016/j.knosys.2011.06.005)

[1016/j.knosys.2011.06.005](https://doi.org/10.1016/j.knosys.2011.06.005)

1. Bruno Krebs. 2018. Using Python, Flask, and Angular to Build Modern Web

Apps. (2018). [https://auth0.com/blog/using-python-flask-and-angular-to-buildmodern-apps-part-1/](https://auth0.com/blog/using-python-flask-and-angular-to-build-modern-apps-part-1/)

1. Robin Burke. 1999. Integrating knowledge-based and collaborative-filtering recommender systems. *Proceedings of the Workshop on AI and Electronic Commerce* (1999), 37.
2. Rikke Friis Dam and Teo Yu Siang. 2020. Empathy Map – Why and How to Use

It. [https://www.interaction-design.org/literature/article/empathy-map-whyand-how-to-use-it](https://www.interaction-design.org/literature/article/empathy-map-why-and-how-to-use-it)

1. Mehdi Elahi, Francesco Ricci, and Neil Rubens. 2016. A survey of active learning in collaborative filtering recommender systems. , 29-50 pages. [https://doi.org/ 10.1016/j.cosrev.2016.05.002](https://doi.org/10.1016/j.cosrev.2016.05.002)
2. John H Falk. 2016. *Identity and the Museum Visitor Experience*. [https://doi.org/ 10.4324/9781315427058](https://doi.org/10.4324/9781315427058)
3. Ivan Keller and Emmanuel Viennet. 2015. Recommender Systems for Museums: Evaluation on a Real Dataset. In *IMMM ’15: The Fifth International Conference on Advances in Information Mining and Management Recommender*. 65–71.
4. Linyuan Lü, Matúš Medo, Chi Ho Yeung, Yi Cheng Zhang, Zi Ke Zhang, and Tao Zhou. 2012. Recommender systems. [https://doi.org/10.1016/j.physrep.2012.02.](https://doi.org/10.1016/j.physrep.2012.02.006)

[006](https://doi.org/10.1016/j.physrep.2012.02.006)

1. Linyuan Lü, Matúš Medo, Chi Ho Yeung, Yi-Cheng Zhang, Zi-Ke Zhang, and Tao Zhou. 2012. Recommender systems. *Physics reports* 519, 1 (2012), 1–49.
2. Museumvereniging. 2020. Museumcijfers 2019. (2020), 40.
3. George Pavlidis. 2019. Recommender systems, cultural heritage applications, and the way forward. , 183-196 pages. <https://doi.org/10.1016/j.culher.2018.06.003>Modern and Contemporary Art.
4. Hasso Plattner, Christoph Meinel, and Larry Leifer (Eds.). 2018. *Design Thinking*

*Research*. Springer International Publishing, Cham. [https://doi.org/10.1007/978-](https://doi.org/10.1007/978-3-319-60967-6)

[3-319-60967-6](https://doi.org/10.1007/978-3-319-60967-6)

1. Paul Resnick and Hal R. Varian. 1997. Recommender Systems. *Commun. ACM* 40(3) (1997), 56 – 58. <https://doi.org/10.1145/245108.245121>
2. Guy Shani and Asela Gunawardana. 2011. Evaluating Recommendation Systems. In *Recommender Systems Handbook*. Springer International Publishing, 257–297. [https://doi.org/10.1007/978-0-387-85820-3{\_}8](https://doi.org/10.1007/978-0-387-85820-3%7B_%7D8)
3. Chieh Wen Sheng and Ming Chia Chen. 2012. A study of experience expectations of museum visitors. *Tourism Management* 33, 1 (2012), 53 – 60. [https://doi.org/](https://doi.org/10.1016/j.tourman.2011.01.023)

[10.1016/j.tourman.2011.01.023](https://doi.org/10.1016/j.tourman.2011.01.023)

1. Yiwen Wang, Natalia Stash, Lora Aroyo, Peter Gorgels, Lloyd Rutledge, and Guus Schreiber. 2008. Recommendations based on semantically enriched museum collections. *Web Semantics* 6, 4 (2008), 1283–290. [https://doi.org/10.1016/j. websem.2008.09.002](https://doi.org/10.1016/j.websem.2008.09.002) Semantic Web Challenge 2006/2007.

# APPENDIX A - VALIDATION SCHEME

Step 1: Theory

* Considering this project’s briefing, Q42 aims to ease the discovery of Dutch museums and stimulate interest by personalising museum.nl’s content. As a result, they hope to increase user engagement.
* After empathising with the user, we assume the main goal of a user while visiting the museum.nl website is to orientate for a next visit.
* One could argue that this goal is driven by the need to behold and enrich oneself through the artworks and showpieces available in the collections of the Dutch museums.
* Within this research, the assumption is made that it is in the Museum Vereniging’s best interest that visitors of museum.nl choose to visit exhibitions or museums based on the features of their collection, rather than the popularity of the concerning museum.
* Considering the previous two assumptions, we argue that the current way of orientation on the museum.nl website, namely, liking and navigating through content cards displaying various Dutch museums, leads to an orientation based on a museum’s popularity, or the experience of visits in the past; it therefore, does not lead to the discovery of new museums or exhibitions.
* Hence, one could conclude that the current interactions and ways of orientation are not sufficiently aligned with the underlying goal and need of the user, as this is artwork- or showpiece oriented.
* We feel that real engagement and personalisation arises after meaningful interactions triggered by the user him or herself, and the orientation of a museum.nl visitor likely shifts per unique visit of a museum, and thus per unique visit of the website.
* We propose a recommendation system that uses visual content (where available) in the form of artworks and showpieces as input to compile a session specific user persona, and uses this persona to recommend exhibitions and/or museums that fit this profile. Step 2: Conceptualisation

1. We believe our theory is substantially validated if the following subjects are true
   * Users acknowledge their main reason for visiting the museum.nl website is for scheduling a visit.
   * Users acknowledge their interest and orientation for visiting art shifts per visit.
   * Users acknowledge they prefer liking artworks or showpieces over liking museums to get a recommendation.
2. We believe the system is engaging when the following factors are at place:
   * The interface allows the user to be in full control
   * The interface and its content are intuitive and understandable to the user
   * The interface and its content focus on existing a user’s needs and the value dimension.
3. We believe the system’s recommendation is sufficient if the following requirements are true:
   * The system’s output is perceived as a logical result of the provided input.
   * The system’s output corresponds to the interest of the user.

Step 3: Operationalisation

*Problem statement*

The current museum.nl website is insufficiently capable of providing its users with a valuable recommendation on which museums or exhibitions to visit in an engaging, personalised manner.

*Objective*

This project aims to realise a recommendation system that allows museum.nl visitors to get a personalised recommendation on which museums or exhibitions to visit in an engaging way. Consequently, the system hopes to establish increased user engagement among its users.

*Research questions*

In order to sufficiently carry out this research , a clear distinction has been made between two aspects within this study. As this project contains an interaction design part, concerning the required interactions and usability of the system, and a system design part, related to the required recommendation models to establish the output of the system, the research questions have been divided as such.

Interaction design:

1. Which interface do museum.nl users prefer while providing input to any recommendation system?
   * This research question will be answered during qualitative research using a comparative usability test.
2. Do the proposed interactions lead to an increased degree of user engagement?
   * This research question will be answered during both qualitative and quantitative research by means of a user survey after testing the concept- and final prototype.

System design:

(1) To what extend is the proposed recommendation model sufficiently capable of generating a desired recommendation to the user’s input?

• This research question will be answered during quantitative research by considering the mean average precision rate and the percentage of museums with a 3+ rating that were in the low online activity list

*Validation*

For validation of the concept, both formative and summative research will be performed. Formative research, research before the implementation of the system, will be used to take insights regarding the usability of the proposed interface into account at an early stage. Furthermore, it will help validating the theory behind the proposed interactions to some degree, while still shaping the final solution. Consequently, this research will support decision making during the process and help maximising the user centricity of the system. Summative research, research after implementation of the system, will be used to validate the final shape of the implementation, and to support its quality assessment. Furthermore, the latter will help assessing the acceptance of the presented system.

Formative research:

During this part of the research, part 1 and 2 of the conceptualisation will be tested during a comparative usability test combined with speaking aloud and a subsequent structured interview. The steps that will be performed are the following:

(1) Recruit representative users:

1. Sampling level: The chosen sampling level within this research concerns typical case sampling. This choice has been made as the assumption was made that users that fit the typical museum.nl visitor persona will be able to provide more valuable insights than participants who deviate from this profile.
2. Sampling criteria: Considering that the most frequent use case of the museum.nl website is to orientate before scheduling a visit, the most important criterium for participants is that they hold a museum card subscription or visit museums on a frequent basis. Other criteria:
   1. Within the age category of 20-60 years.
   2. Users are in no way involved in the study
   3. Resident of The Netherlands
3. Sample size: 5 users should be enough to discover 80% of the usability problems.

(2) Give them a representative task to perform:

(a) Task: You’re planning a visit to a Dutch museum, however you don’t have a preference for a city, nor a specific museum. You want to discover new exhibitions and get surprised by what the website recommends. Go to the museum.nl website and choose an exhibition that you want to visit by making use of the recommendation feature. (Use the prototype)

(3) Observe users:

1. Interactions will be recorded via screen recording
2. Speaking aloud will be recorded via audio recording
3. What path does the user choose, interface A (TindArt) or interface B (Museum.nl alignment) and thus is likely the most intuitive?
4. Striking observations will be documented and processed
5. Structured interview / survey

(4) Document subsequent conclusions and new requirements that can be derived from these conclusions.

Summative research:

During this part of the research, the improved prototype and its interactions will be tested on part 2 and 3 of the conceptualisation by performing a quantitative experiment with a subsequent user survey.

(1) Recruit representative users (the same sampling level and criteria apply as during the formative research):

(a) 10-15 participants

(2) Give them a representative task to perform:

(a) Task: You’re planning a visit to a Dutch museum, however you don’t have a preference for a city, nor a specific museum. You want to discover new exhibitions and get surprised by what the website recommends. Go to the museum.nl website and choose an exhibition that you want to visit by making use of the recommendation feature. (Use the prototype)

(3) Perform user survey:

1. Usability questions
2. Recommendation questions

(4) Document subsequent conclusions and new requirements that can be derived from these conclusions.

1. We will consider a score of 4+ as a relevant recommendation
2. We will report 2 values
   1. Mean average precision across all interactions
   2. % of museums with a 3+ rating that were in the low online activity list.
3. In the long run, the click through rate on the shortlist will be a better measure to track effectiveness of the recommender.

# APPENDIX B - QUALITATIVE RESEARCH

This appendix contains all the questions, and the answers provided by the respondents during the survey. The screen and audio recordings of the performed usability tests have not been included in the annex of this report, due to privacy reasons.

Respondent 1:

*Q1: Please confirm that you were in no way involved during the implementation of this project, or during the creation of this prototype.* Yes

*Q2: What is your age?* 22

*Q3: Where are you from?*

Friesland

*Q4: What is your gender?* Woman

*Q5: Do you have a museum card subscription?* Yes

*Q6: How frequently do you visit a museum?*

5

*Q7: Do you prefer interface A, or interface B while providing input to the recommendation system?* Interface A

*Q8: Why do you prefer interface A?*

It is easier to understand because you have only two options swipe left or right

*Q9: Why do you prefer interface B?*

n.a.

*Q10: Do you ever use the museum.nl website?* No

*Q11: What is your main reason for visiting the museum.nl website?*

To explore about different museums

*Q12: What interaction would you prefer on the museum.nl website, liking artworks/showpieces or liking museums/exhibitions?*

Museums Exhibitions

*Q13: Why do you have this preference?*

Because im more interested to that kind of art

*Q14: I like to search for the same type of museums, and/or exhibitions every time I use the museum.nl website.* 3

*Q15: During the test I had the feeling I was in full control over the system.* 2

*Q16: The system’s interface and its interactions are intuitive.* 2

*Q17: The system’s interface and its interactions were understandable.* 2

*Q18: I visit a great variation of museums with varying artworks and showpieces.* 3

*Q19: I see myself using this tool when scheduling a visit to a museum in the future.* 3

*Q20: Is there any other feedback you would like to give us regarding the prototype?*

Yes I would recommend the second one! The first one was very unclear and confusing Respondent 2:

*Q1: Please confirm that you were in no way involved during the implementation of this project, or during the creation of this prototype.* Yes

*Q2: What is your age?* 24

*Q3: Where are you from?* Eindhoven

*Q4: What is your gender?* Woman

*Q5: Do you have a museum card subscription?* Yes

*Q6: How frequently do you visit a museum?*

4

*Q7: Do you prefer interface A, or interface B while providing input to the recommendation system?* Interface B

*Q8: Why do you prefer interface A?*

n.a.

*Q9: Why do you prefer interface B?*

In interface B; I got the opportunity to choose from a broad range of art works. I like this process because I took a closer look before ’liking’ the art works. Interface A, I got to the exhibition part faster. But I got the idea that this choice was made for me (after seeing only 3-4 art works); like using a dating app ;-)

*Q10: Do you ever use the museum.nl website?*

Yes

*Q11: What is your main reason for visiting the museum.nl website?*

Planning visits; checking out exhibitions. Finding information about my museum card etc

*Q12: What interaction would you prefer on the museum.nl website, liking artworks/showpieces or liking museums/exhibitions?*

Artworks Showpieces

*Q13: Why do you have this preference?* Liking art works speaks for itself; there’s no doubt in choosing what you like. When choosing an exhibition (before you’ve visited it), you might not be able to make a wellcosidered choice of what you like.

*Q14: I like to search for the same type of museums, and/or exhibitions every time I use the museum.nl website.* 3

*Q15: During the test I had the feeling I was in full control over the system.* 4

*Q16: The system’s interface and its interactions are intuitive.* 4

*Q17: The system’s interface and its interactions were understandable.* 3

*Q18: I visit a great variation of museums with varying artworks and showpieces.* 5

*Q19: I see myself using this tool when scheduling a visit to a museum in the future.* 5

*Q20: Is there any other feedback you would like to give us regarding the prototype?*

In Interface B it wasn’t completely clear for me that I was going to find/choose an exhibition by clicking ’discover your options’.

Respondent 3:

*Q1: Please confirm that you were in no way involved during the implementation of this project, or during the creation of this prototype.* Yes

*Q2: What is your age?* 31

*Q3: Where are you from?* Hengelo

*Q4: What is your gender?* Man

*Q5: Do you have a museum card subscription?* Yes

*Q6: How frequently do you visit a museum?*

4

*Q7: Do you prefer interface A, or interface B while providing input to the recommendation system?* Interface B

*Q8: Why do you prefer interface A?*

n.a.

*Q9: Why do you prefer interface B?*

The interface of B is more user-friendly. It works a bit faster when you already see all available options instead of waiting till you see the next painting.

*Q10: Do you ever use the museum.nl website?*

Yes

*Q11: What is your main reason for visiting the museum.nl website?*

Looking for new museums which I could possibly visit in the future

*Q12: What interaction would you prefer on the museum.nl website, liking artworks/showpieces or liking museums/exhibitions?*

Artworks Showpieces

*Q13: Why do you have this preference?* I think it’s more useful when you can see what is inside a museum instead of looking at the outside of a museum or an unfamiliar name/exhibition. From the outside, you can’t see what is inside. The most important thing is if you like the pieces they have because that is most of the time the main-reason to visit a museum.

*Q14: I like to search for the same type of museums, and/or exhibitions every time I use the museum.nl website.* 3

*Q15: During the test I had the feeling I was in full control over the system.* 4

*Q16: The system’s interface and its interactions are intuitive.* 5

*Q17: The system’s interface and its interactions were understandable.* 5

*Q18: I visit a great variation of museums with varying artworks and showpieces.* 4

*Q19: I see myself using this tool when scheduling a visit to a museum in the future.* 5

*Q20: Is there any other feedback you would like to give us regarding the prototype?*

I really like the concept! I think B works really good and easy. Maybe adding more art pieces would be a good idea. At this time, there are only 8 pieces available to select. If I don’t see more than three pieces that I’m interested in, I have to select one which I maybe don’t really like.

Respondent 4:

*Q1: Please confirm that you were in no way involved during the implementation of this project, or during the creation of this prototype.* Yes

*Q2: What is your age?* 28

*Q3: Where are you from?* Amsterdam

*Q4: What is your gender?* Man

*Q5: Do you have a museum card subscription?* Yes

*Q6: How frequently do you visit a museum?*

3

*Q7: Do you prefer interface A, or interface B while providing input to the recommendation system?* Interface B

*Q8: Why do you prefer interface A?*

n.a.

*Q9: Why do you prefer interface B?* Overall overview of the art I could choose. Gave me the feeling I was more in control of the outcome and could pick the art I liked most

*Q10: Do you ever use the museum.nl website?*

Yes

*Q11: What is your main reason for visiting the museum.nl website?* finding new exhibitions

*Q12: What interaction would you prefer on the museum.nl website, liking artworks/showpieces or liking museums/exhibitions?*

Artworks Showpieces

*Q13: Why do you have this preference?* I believe the art to be of more significant value than the museums

*Q14: I like to search for the same type of museums, and/or exhibitions every time I use the museum.nl website.* 4

*Q15: During the test I had the feeling I was in full control over the system.* 5

*Q16: The system’s interface and its interactions are intuitive.* 5

*Q17: The system’s interface and its interactions were understandable.* 5

*Q18: I visit a great variation of museums with varying artworks and showpieces.* 4

*Q19: I see myself using this tool when scheduling a visit to a museum in the future.* 4

*Q20: Is there any other feedback you would like to give us regarding the prototype?*

No

Respondent 5:

*Q1: Please confirm that you were in no way involved during the implementation of this project, or during the creation of this prototype.* Yes

*Q2: What is your age?* 38

*Q3: Where are you from?* Deurne

*Q4: What is your gender?* Man

*Q5: Do you have a museum card subscription?* Yes

*Q6: How frequently do you visit a museum?*

4

*Q7: Do you prefer interface A, or interface B while providing input to the recommendation system?*

Interface B

*Q8: Why do you prefer interface A?*

n.a.

*Q9: Why do you prefer interface B?*

Interface B allowed me to see all the artworks before I needed to make a selection. Furthermore, I found the colors within A rather distracting.

*Q10: Do you ever use the museum.nl website?*

Yes

*Q11: What is your main reason for visiting the museum.nl website?*

I use it to get some inspiration on which museum to visit next. I also used to use the website for getting some information on opening times etc. but I use Google for that now.

*Q12: What interaction would you prefer on the museum.nl website, liking artworks/showpieces or liking museums/exhibitions?*

Artworks Showpieces

*Q13: Why do you have this preference?* Museums don’t say a lot, artworks do. Besides there are still a lot of museums unknown to me.

*Q14: I like to search for the same type of museums, and/or exhibitions every time I use the museum.nl website.*

1

*Q15: During the test I had the feeling I was in full control over the system.*

5

*Q16: The system’s interface and its interactions are intuitive.*

4

*Q17: The system’s interface and its interactions were understandable.*

5

*Q18: I visit a great variation of museums with varying artworks and showpieces.*

5

*Q19: I see myself using this tool when scheduling a visit to a museum in the future.*

5

*Q20: Is there any other feedback you would like to give us regarding the prototype?*

Maybe you should think about making the button at the bottom of the screen more guiding. So maybe adding a focus when the required number of artworks are selected, or showing a notification in case a user clicks on it before the correct amount has been reached.

# APPENDIX C - QUANTITATIVE RESEARCH

This appendix contains all the questions that were ask during the quantitative research. First, the questions will be outlined together with their label. Then, a figure will be displayed with all the corresponding answers positioned underneath their label. In case a question is a Yes or No question, 1 can be considered a ’Yes’ and 0 a ’No’.

Questions

*Q1: Please confirm that you were in no way involved during the implementation of this project, or during the creation of this prototype.*

*Q2: What is your age?*

*Q3: Where are you from?*

*Q4: What is your gender?*

*Q5: Do you have a museum card subscription? Q6: How frequently do you visit a museum? (Likert-scale 1-5) Q7: Do you ever use the museum.nl website?*

*Q8: What is your main reason for visiting the museum.nl website?*

*Q9: On a scale of 1-5, to what extent do the following statements apply to you: I like to search for the same type of museums, and/or exhibitions every time I use the museum.nl website.*

*Q10: On a scale of 1-5, to what extent do the following statements apply to you: During the test I had the feeling I was in full control over the system.*

*Q11: On a scale of 1-5, to what extent do the following statements apply to you: The system’s interface and its interactions are intuitive.*

*Q12: On a scale of 1-5, to what extent do the following statements apply to you: The system’s interface and its interactions were understandable.*

*Q13: On a scale of 1-5, to what extent do the following statements apply to you: I visit a great variation of museums with varying artworks and showpieces.*

*Q14: On a scale of 1-5, how relevant do you find this recommendation: Recommended museum/exhibition 1*

*Q15: On a scale of 1-5, how relevant do you find this recommendation: Recommended museum/exhibition 2*

*Q16: On a scale of 1-5, how relevant do you find this recommendation: Recommended museum/exhibition 3*

*Q17: On a scale of 1-5, how relevant do you find this recommendation: Recommended museum/exhibition 4*

*Q18: On a scale of 1-5, to what extent do the following statements apply to you: I see myself using this tool when scheduling a visit to a museum in the future.*

*Q19: On a scale of 1-5, to what extent do the following statements apply to you: The recommendation is a logical output considering my liked artworks*

*Q20: On a scale of 1-5, to what extent do the following statements apply to you: The recommendation output corresponds to my interest.*

*Q21: Is there any other feedback you would like to give us regarding the prototype?*

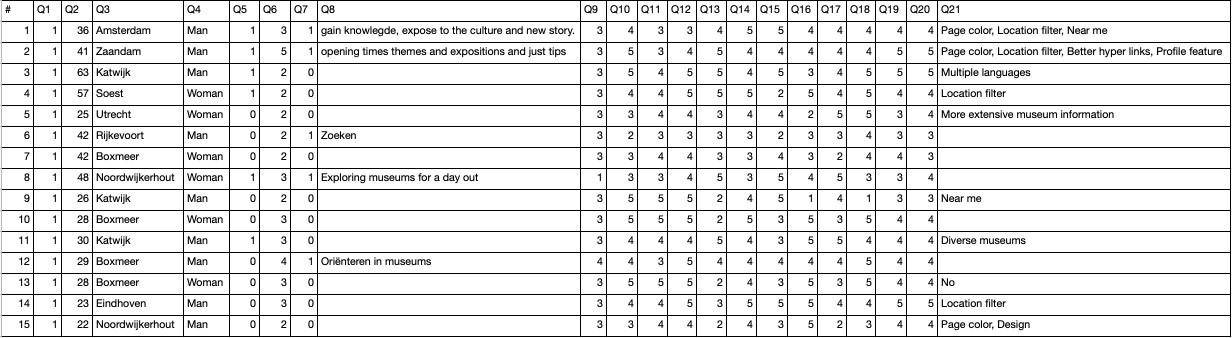


Figure 20: Answers quantitative research

1. www.museum.nl/en/museumpass [↑](#footnote-ref-1)
2. http://chip.win.tue.nl/ [↑](#footnote-ref-2)
3. Tinder: https://tinder.com/?lang=en [↑](#footnote-ref-3)
4. www.palletsprojects.com/p/flask/ [↑](#footnote-ref-4)
5. www.angular.io/ [↑](#footnote-ref-5)