

**TANGIBLE CULTURAL ANALYTICS: THE ADOPTION OF RECOMMENDER  
SYSTEMS WITHIN CULTURAL RESEARCH**

By:

Ben Ashley

Bachelor of Management, Major in Entrepreneurship and Innovation

Rowe School of Business

Dalhousie University

2017

A Major Research Project presented to Ryerson University in partial fulfillment of the requirements for the degree of Master of Digital Media in the program of Digital Media

Toronto, Ontario, Canada, 2019

© Ben Ashley, 2019

## **Author's Declaration**

I hereby declare that I am the sole author of this MRP. This is a true copy of the MRP, including any required final revisions.

I authorize Ryerson University to lend this MRP to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this MRP by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my MRP may be made electronically available to the public.

## **Abstract**

### TANGIBLE CULTURAL ANALYTICS: THE ADOPTION OF RECOMMENDER SYSTEMS WITHIN CULTURAL RESEARCH

Ben Ashley

Master of Digital Media

Digital Media

Ryerson University, 2019

The prospect of implementing recommender systems within the context of cultural research has not been explored nearly as much compared to implementation in e-commerce websites and applications. Recommender systems allow for users to be shown new objects either based upon object similarity or based upon what the algorithm thinks the user will like – which can be derived from user feedback and comparing the user to other similar users. This paper discusses how a recommender system could benefit an augmented reality application that enables 3D viewing of artifacts – as part of the Tangible Cultural Analytics (TCA) project at Ryerson University’s Synaesthetic Lab. This paper outlines four recommender systems: 1) content-based filtering, 2) collaborative filtering, 3) cluster models 4) search based models, and 5) hybrid models; discussing the pros and cons to each. Ultimately, a content-based model without the user profile aspect was chosen for this stage in the prototype. This model showed us just how much potential these recommender systems have when helping cultural researchers uncover new relationships and pieces of history through the study and comparison of artifacts.

## **Acknowledgments**

I would like to thank my supervisor Dr. Ali Mazalek, who provided invaluable guidance to this project. I would also like to thank Alevtina Naumova, who provided crucial feedback as my second reader.

I would like to thank Ahmed Sagarwala and the rest of the MDM staff who provided support and advice when I needed it most. I would like to thank the fellow members of the Synaesthetic Lab who provided advice and feedback to the team. Lastly, I would like to thank my fellow group members, Daniella Kalinda and Zeeanna Ibrahim. Working with them was a privilege.

## **Dedication**

*In dedication to my Mom and Dad. This project would not have been possible without their continued support and belief in me.*

## Table of Contents

<i>Author's Declaration</i> .....	<i>ii</i>
<i>Abstract</i> .....	<i>iii</i>
<i>Acknowledgments</i> .....	<i>iv</i>
<i>Dedication</i> .....	<i>v</i>
<i>Introduction</i> .....	<i>1</i>
<i>Background</i> .....	<i>3</i>
<i>Literature Review – Overall Scope of Project</i> .....	<i>4</i>
Challenges of Accessibility.....	<i>4</i>
Digitization .....	<i>4</i>
New Types of Interactions through Emerging Technologies.....	<i>5</i>
<i>Methodology</i> .....	<i>6</i>
<i>Related Work – Recommender Systems</i> .....	<i>7</i>
Content-Based System.....	<i>9</i>
Collaborative Filtering .....	<i>12</i>
Cluster Models.....	<i>15</i>
Search-based Models.....	<i>15</i>
Hybrid Recommender Systems .....	<i>16</i>
<i>Evaluation</i> .....	<i>17</i>
TCA's Content-Based Recommender System .....	<i>17</i>
Metadata.....	<i>17</i>
Algorithm .....	<i>19</i>
<i>Discussion and Future Work</i> .....	<i>21</i>
<i>Conclusion</i> .....	<i>22</i>
<i>References</i> .....	<i>24</i>

## **Introduction:**

Studying and researching artifacts and ephemera<sup>1</sup> have been a part of learning about history for centuries. Traditional methods of observing the past used by historians, archeologists, and archivists mostly involve the physical study of these artifacts – that is obtaining the artifact and observing its physical properties to understand its significance and narrative. New and emerging technologies however have now opened the door to new horizons and opportunities in studying these artifacts that are full of rich history. The Tangible Cultural Analytics (TCA) team, as part of the Ryerson Synaesthetic Laboratory, has dedicated itself to discovering how new and emerging technologies can help facilitate the research and accessibility of artifacts and ephemera. These innovative technologies include augmented reality, virtual reality, machine learning, artificial intelligence, and recommender systems. Which technologies can serve historical professionals best when it comes to studying artifacts? What are the pros and cons to each? The team’s goal was to figure out a method, using emerging technologies, that would make accessibility and research easier for cultural historians so that more in-depth pieces of history could be discovered. Ultimately, the team decided on using augmented reality as the best method for accessing artifacts that would not be readily accessible. Then, the prospect of including a recommender-type system was discussed and eventually agreed upon.

This paper focuses on recommender systems and how they can be used for an augmented reality application in which its primary function is for studying and researching artifacts. The different

---

<sup>1</sup> Things that exist for only short periods of time that have a certain historical significance such as postcards or tickets; usually written material (Oxford Dictionary, 1949).

types of recommender systems are discussed as well as the pros and cons to each. The prospect of including machine learning and artificial intelligence in future iterations is also discussed. Ultimately the question that is discussed here is how can recommendation systems help with cultural research? What is it that cultural researchers are looking for next when studying artifacts? What common themes do they look for when studying artifacts? This paper will observe the different types of recommendation systems that exist through a literature review and choose one that will help answer these questions. The paper will then discuss a prototype that was developed to help illustrate this type of recommender system and how it can help cultural researchers with their work.

This paper begins with a “Background” section outlining the TCA team objectives. Since this project was done in conjunction with two other Master of Digital Media (MDM) students, their contributions to the project are briefly outlined in this section as well. In the “Related Work – Overall Scope of Project” section, literature on examples of the use of emerging technology in the field of cultural history is discussed. In the “Methodology” section, the formative study that the TCA team conducted is discussed which outline the motivations of this project, which then leads into the second “Related Work” section that discusses recommender systems. The “Evaluation” section goes over the recommender system prototype that was included in the augmented reality application prototype developed as part of our MRP showcase. Lastly the “Future Work” section goes over how this recommender system could improve in future iterations.



## **Background:**

The use of augmented reality and virtual reality, in the field of cultural research, has seen a massive increase in use in the past few years. These exciting new technologies have prompted for increased interest from the general public regarding the study of artifacts and ephemera and their relation to cultural history. This has been shown in museums and cultural sites, where they have experienced improved feedback in visitor experience (He, Wu & Li, 2018). In the context of academic research however, there has not been as much innovation when it comes to the implementation of emerging technologies. The TCA team, comprised of Daniella Kalinda, Zeeanna Ibrahim, and myself set out to solve this problem and come up with a digital solution. Daniella's role was heavily focused on the primary research aspect of the project. She took the lead on conducting a formative user study to understand the current practices of and challenges faced by cultural history scholars. The results of the formative study informed the design of our prototype in Unity – for which Daniella contributed in the development as well. Zeeanna's role focused primarily on the User Experience design aspect, developing the user interface for our multiple prototypes. She designed the entire User Interface, creating low to high fidelity prototypes. She was also involved in the user research process, helping out with the interview process. My role was heavily focused in the recommendation aspect of the app which is discussed in this paper.

## **Literature Review – Overall Scope of Project:**

### **Challenges of Accessibility:**

Since the early 1900s (Anderle, 1978), there have been challenges associated with the discovery, access, and preservation of artist's files (ephemera) and artifacts (Smith, 2016). The time and cost of accessing artifacts that might be halfway around the world often isn't feasible financially. Lots of money is often wasted on simply looking for materials rather than studying them (Anghelescu, 2001). On top of this, there is also the issue of information overload. The problem is not knowing where to start and what to select for acquisition out of the enormous amount of material that is available in the world – digitized or not (Anghelescu, 2001). Often times, researchers simply don't know where to look or even what to look for when working on a project. Overall, there has not been a consistent and systematic approach regarding the collection, preservation, and cataloging of ephemeral materials by libraries, archives, and historical societies (Anghelescu, 2001).

### **Digitization:**

The solution of digitizing these ephemeral materials and artifacts for improved accessibility isn't necessarily as straight forward as one might think. Institutions often lack the financial resources for digitization projects (Smith, 2016). In the case of ephemera, the digitization standards for these types of projects are generally not feasible for collections either (Smith, 2016). This has led many cultural heritage institutions to search for inexpensive ways to make file holdings and artifacts discoverable to researchers (Smith, 2016). There is also the assumption that by

digitizing artifacts and ephemera, they won't be forgotten nearly as much. However, there is danger of the opposite happening, that by digitizing this material it would be forgotten once more (Mussell, 2012). That digitizing might only increase the information overload that already exists. Essentially, the overload and persistence of digital objects within unused or neglected storage media are no different than forgotten pieces of ephemera, lying in boxes in the basement of a library (Mussell, 2012). In order for this to not happen, there needs to be a system or mechanism that prompts users to "choose to remember them" (Mussell, 2012). This means that objects must be well-coded with the appropriate documentation as well as an institution for which they can be appropriately and actively curated (Mussell, 2012).

### **New Types of Interactions through Emerging Technologies:**

There are a number of museums and public installations that are incorporating digitally-enhanced interactive experiences that provide users with a 'multimodal' (Kress, 2009) engagement with the past (Kwan, Chu, Harley, McBride, & Mazalek, 2016). These museums have been increasingly using mono and duo sensory experiences to help convey certain narratives to visitors (Kwan et al., 2016). These technologies have generally included overlaying textual and visual content onto artifacts thus enhancing visual engagement (Kwan et al., 2016). Examples of these include augmented reality technologies such as ScopifyROM at the Royal Ontario Museum (Royal Ontario Museum, 2013) and Jurascope at the Natural Museum of History in Berlin (ART+COM, 2007). The Boston Museum of Science has also developed digitally enhanced puzzle blocks (Horn, Solovey, & Jacob, 2008) to help children understand the basics of programming. Mapping Place (Chu, Clifton, Harley, Pavao, & Mazalek, 2015) is a technology that allows for museum visitors to make and share stories about their family through a multi-

touch tabletop using connection between virtual story beads to tangible story shells. These examples demonstrate that there is a clear market for these emerging technologies in the museum industry, but what about within the context of cultural research?

## **Methodology:**

The team conducted 8 open-ended interviews in a qualitative format with individuals who work in the field of cultural history. The motivations, derived from a literature review, behind these interviews were to discover current working processes of scholars in the cultural history field, to identify pain points regarding interactions toward researching artifacts and ephemera, and to get potential ideas and inspirations for the design of future prototypes. The hypothesis was that researchers are looking for and are open to intuitive ways of researching artifacts.

A formative assessment was written by fellow team member, Daniella Kalinda (Kalinda, 2019), that outlines some of the main takeaways from the study. One of the primary pain points that was relatively common amongst the interviewees was low accessibility. Due to time and money constraints, it can be very difficult to get a physical view of many artifacts around the world. Interviewees also expressed the desire to have more items available online as well as a better cataloguing system for current selections. Overall, they found that augmented reality would be beneficial for teaching purposes, but not necessarily research. It was noted however, that many of these interviewees began their careers before the age of the internet and “this may be a generational thing”. It is also worth noting that many experts in some fields generally have a more close-minded view when it comes to envisioning a new way of working with artifacts.

On top of the pain points of accessibility and the lack of digital cataloging, a common theme that was brought up among interviewees was the narrative of an artifact. Essentially, how an object can “speak” to another and create a larger narrative around that. One person said that he is “always looking at the narrative and how that narrative is connected to other objects associated with it”. The use of graph databases was discussed – as opposed to spreadsheets. Graph databases allow for easily defining relationships and picking out patterns. If patterns are discovered, then one would look for similar patterns in the rest of the graph. This was further explained through an example given by one of the interviewees. “If I pull up a Paleo-Indian artifact which is from 2000 years ago, and I had a system where I could pull up another artifact that has the same morphology but is from Russia and 3000 years ago, then there is something interesting here. Similar styles, similar cultures could be explored. Maybe they even learned from the same people.”. Even though this wasn’t necessarily a pain point, it reinforced the fact that recommender systems could be hugely beneficial toward not only our digital solution, but also cultural research as a whole.

## **Related Work – Recommender Systems:**

Recommender systems have become extremely prominent when it comes to e-commerce sites and mobile applications alike. In a broad sense, recommendation systems are the reason behind the “items we think you’ll like” feature that is part of many services today. They utilize past user preferences, purchase history, demographic information, item popularity, and metadata

characteristics (Bogers & van der Bosch, 2009) to give a user the most relevant and most personalized recommendations for items that the user will want to purchase, or in the case of this paper, observe and research. Recommender systems ideally help the user find items that are relevant to the user's interests. These recommendations might be similar items, they could be items from a similar brand, or they could be complimentary items. The idea behind these systems is to not simply just recommend similar items to the user, but to recommend items that the system thinks the user will like. This is why recommender systems take into account both the content of items that are viewed as well as looking at the user's profile and his or her actions. Within the context of an augmented reality application for tourism, there has been evidence that suggests implementing ratings and recommendations based on feedback from previous users could be beneficial (tom Dieck & Jung, 2018). Augmented reality applications should essentially be built around users' preferences and feedback and that these preferences should result in recommendations (tom Dieck & Jung, 2018). This would likely increase acceptance of such an application (tom Dieck & Jung, 2018).

There is an abundance of artifacts and ephemera out there. Many times, researchers don't know where to look next after finding an indigenous arrowhead because all the information that is readily available to us can be overwhelming and it can be tough to figure out where to look next. In the context of cultural research, there are numerous benefits to how a recommender system could help the research process for professionals who study artifacts and ephemera. Finding artifacts that could help enforce a certain narrative is one, but also discovering new narratives is another. For example, looking at similar patterns on artifacts, similar materials, similar locations all could lead to new discoveries that could uncover new fascinating pieces of history.

## **Content-Based System:**

The content-based recommender model recommends items to a user based upon the description of an item as well as a user profile that contains the user's interests (Pazzani & Billsus, 2007).

This system uses a number of factors to recommend a certain item to a user. It is mostly used for a recommendation system design that uses the contents of items to create features and attributes which are then matched to a user profile (Wei, He, Chen, Zhou, & Tang 2017). Items are then compared with items that have been previously liked by the user, and then the best matched ones are the items that get recommended (Wei et al., 2017). The feedback that is given by the user doesn't necessarily have to be limited to likes. The feedback can include giving an item a certain number of stars out of five, or it can include the amount of interaction that the user has with the specific item. In some cases, the content can be represented through the use of keyword-based models, in which the system creates a type of Vector Space (VSM) representation of item features or metadata (Deldjoo et al., 2016). Generally, a relevance score is then developed for each item that is being compared, and then the items that produce the highest scores are recommended to the user.

Generally, in a content-based system, items that can be recommended are stored into a database table that includes an item's ID, its name, cuisine, service cost (Pazzani & Billsus, 2007). There are two types of values – structured and un-structured. Structured generally refers to data that is predefined (Pazzani & Billsus, 2007) such as metadata. An example would be an item's release date. Un-structured data refers to data that is much more disorganized (Pazzani & Billsus, 2007) such as the description of an item. For example, the recommender system could pick up certain key words out of a paragraph that it has previously recognized. Having structured data is

generally much more effective when constructing a recommender system. Many domains that exist today use a combination of some semi-structured data in which there are restricted values as well as free-text fields (Pazzani & Billsus, 2007).

In a broad sense, the user profile of a content-based recommender system refers to a function that for any item, predicts the likelihood that the user will be interested in said item (Pazzani & Billsus, 2007). Generally, this takes into account the descriptions of the types of items that interest the user and the history of the user's interactions with the recommender system (Pazzani & Billsus, 2007). Some recommender systems have menus that explicitly ask the user what types of items he or she like, thus narrowing down the recommendations and making the system more accurate. This is called user customization and it is often found in most prominent recommender systems. A common method of implementing this is by having the user check a number of boxes that will allow the user to select from a known value of attributes (Pazzani & Billsus, 2007). Examples could include types of restaurants, the names of favorite sports teams, or favourite sections of a news site (Pazzani & Billsus, 2007). Another way would be to include a form that allows the user to type in certain key words that occur in un-structured data such as descriptions (Pazzani & Billsus, 2007).

Within many content-based systems are rule-based recommenders (Pazzani & Billsus, 2007). In a broad sense, these systems essentially implement specific rules to the recommender function. They look at more than just similar metadata attributes and un-structured data. For example, a rule might be that the recommender system should recommend complimentary items to the item being viewed. Another might be to recommend sequels to a book or movie that a user may have



purchased earlier (Pazzani & Billsus, 2007).

There is also the prospect of creating a content-based recommender system that uses strictly structured data as a means to recommending new items. This might be using certain attributes such as metadata to recommending new items. This approach generally uses both item-intrinsic and item-extrinsic metadata (Bogers & van den Bosch). Item-intrinsic refers to something that is inherent to the specified item – something that wouldn't change if the item existed in a different context (Greenberg, McPherson, & Mayernik 2010). These include things like the date created or the type of material. Item-extrinsic refers to the parts of an item that are external to its essence (Greenberg et al., 2010). Examples might be where a certain item was purchased or where an item was accessed. Generally, item-intrinsic fields are preferred because they can be used as stand-alone sources for recommending content whereas item-extrinsic is usually much more unique (Bogers & van den Bosch). This type of system however can be limited with the user only being recommended items that are similar to the ones he or she is looking at rather than new types of items that may interest him or her.

There are a number of strengths to using the content-based recommender system. Generally, the process is easy to explain and convey to users (Ahn & Shi, 2008). The functions aren't necessarily too complex to understand, which makes users much more content to input their preferences willingly (Ahn & Shi, 2008). The user profile aspect is another strong point. As the user utilizes the system more and more, his or her profile becomes more and more personalized to his or her individualized taste. The content-based system is serviceable and practical, meaning that not many resources are required to construct nor is there much of a computation process

(Ahn & Shi, 2008). Non-personalized systems are also possible with the content-based system because the coverage is so wide, therefore there is no constraint except for the item-analysis technique (Ahn & Shi, 2008). Overall, the primary benefit of a content-based recommender system is the fact that the user is assured he or she will have similar-like recommended items.

There are of course a number of limitations to the content-based system. Firstly, they require effort from the user regarding the user profile – in the sense that they are often required to input preferences whether it be checking off boxes in a menu or simply indicating whether or not they like an item (Pazzani & Billsus, 2007). Users' interests can also change, which can be problematic when the user profile has been trained to only recommend certain types of items to a user. These recommender systems also don't determine the order in which to present the items (Pazzani & Billsus, 2007). The modalities of certain for which the content-based system can deal with are also limited, only a shallow analysis of certain kinds of content can be supplied (Ahn & Shi, 2008). Recommended items can also be over-specialized in the sense that the user is only recommended items that are similar, unless a rule is inputted into the algorithm (Ahn & Shi, 2008).

### **Collaborative Filtering:**

Another common approach to recommender systems is using the collaborative filtering method. This system, used commonly on e-commerce websites, uses inputs about a user's interests and compares them to other users to generate a list of recommended items (Linden, Smith, & York, 2003) generally through segmentation. This approach is usually highly effective when forecasting user precedence in terms of the choice of certain items (Rafsanjani, Salim, Aghdam,

& Fard, 2013). This method flourished in the middle of the 1990s with the emergence of a number of high-powered companies like Netflix, Amazon, and Elsevier (Rafsanjani et al., 2013). Collaborative filtering uses a number of different types of inputs of implicit and explicit information (Linden et al., 2003). It is considered one of the most popular approaches for recommender system design and uses large amounts of data that is collected from past user behavior and predicts which items users will like (Wei et al., 2017). It relies heavily on the relationship between users as well as items and encodes the data into a rating feedback matrix (Wei et al., 2017).

Collaborative filtering focuses on usage patterns as well as the adding and rating of items (Bogers & ven den Bosch). Collaborative filtering usually attempts to utilize a “word-of-mouth” technique toward the recommendation method where certain items are recommended according to the similarity in interests among customers and how they have categorized these objects (Shardanand, 1995). The way most Collaborative Filtering algorithms usually start however is by finding the set of customers whose explicit and implicit feedback overlap with a user’s similar explicit and implicit feedback (Resnick et al., 1994). The system then aggregates items from this set of customers, eliminates the ones that the user has already purchased, and then recommends the remaining items to the user (Linden et al., 2003). Collaborative Filtering algorithms often utilize patterns that demonstrate users’ interaction and feedback toward items (Rafsanjani et al., 2013). Many traditional collaborative filtering models represent a user as an  $n$ -dimensional vector of items, where  $n$  refers to the number of distinct items (Linden et al., 2003). In this case, the algorithm generates recommendations based on a small number of users who are most similar to the user using the platform. The system can then pick and choose certain recommendations

from similar users' through a number of methods, one of which is to rank each item based upon how many similar customers have purchased it (Linden et al., 2003).

Collaborative Filtering has a wide array of strengths. Firstly, it can take an item's quality or lack thereof into account when recommending certain items (Rafsanjani et al., 2013). For example, in a content-based recommender system, a poor-quality action movie could fall into the same category as an all-time classic such as Die Hard. In collaborative filtering though, item quality is taken much more into account by taking the preferences of similar customers and ranking them. Another pro to collaborative filtering is the fact that they are useful within spheres where analysis of content can be difficult to retrieve such as music or film suggestions (Burke, 2002).

One of the issues with collaborative filtering is the fact it can suffer from a "complete cold start" problem – where there are no rating records available for items, and an "incomplete cold start" problem – where only a small number of rating records are available within the database (Wei et al., 2017). In order for collaborative filtering to be effective, there needs to be a substantial number of ratings on items within the system. Another issue can be the concept of "Gray Sheep", which refers to the hardship that is had by certain users who do not belong to a specific user segment (Claypool et al., 1999). Scalability can also pose an issue for these systems – as segments, users, and items grow, more systems will be required to manage the overwhelming demand which can be costly and difficult (Rafsanjani et al., 2013).

## **Cluster Models**

Cluster models can be considered as somewhat of a sub-set of collaborative filtering. Cluster models divide the user base into different segments and treat the task of recommendation as a classification problem (Linden et al., 2003). In this case, the algorithm's goal is to assign a user to a specific segment that contains the most similar customers (Linden et al., 2003). The system then uses a combination of purchases and ratings of customers within a specified segment to generate a list of recommendations (Linden et al., 2003). These segments are generally created through a cluster or through machine learning algorithms (Linden et al., 2003). Clustering algorithms typically start with an initial set of segments, which often contain one randomly selected customer each, then repeatedly matches customers to existing segments (Linden et al., 2003). One of the strengths of cluster models is the fact that they have strong online scalability capabilities and performance compared to more traditional models of collaborative filtering (Linden et al., 2003). However, the recommendation quality is generally poor due to the fact that there is a finite number of user segments (Linden et al., 2003).

## **Search-based Models:**

Search based methods treat the task of recommendation as a "search for related items" issue (Balabanovic & Shoham, 1997). The algorithm uses a search query to find other popular items by the same author, artist, or director – or similar keywords (Linden et al., 2003). An example might be if someone who buys an Indiana Jones movie, the system might recommend other action-adventure movies, or it might recommend movies from the same director. If the user has few purchases and gives minimal ratings, this recommendation system works well (Linden et al., 2003). However, for users who have made hundreds of purchases, it is much harder to base a

query on all these items (Linden et al., 2003). Overall, the recommendation quality here is relatively poor with the issue being the fact that recommendations are either too general or too narrow (Linden et al., 2003).

### **Hybrid Recommender Systems:**

Hybrid systems are a combination of a number of different recommender systems, usually a combination between content-based and collaborative filtering techniques (Rafsanjani et al., 2013). The idea here is to take advantage of the strengths of multiple systems and to minimize on the weaknesses therefore producing a more efficient result (Wei et al., 2017). There are a number of combination approaches to hybrid systems:

- **Mixed:** Various types of recommender systems are presented simultaneously.
- **Weighted:** Utilization of explicit feedback to produce recommendation approaches.
- **Feature Combination:** Certain characteristics which are related to various recommendation data resources are brought together into a single algorithm.
- **Cascade:** System refines the recommendations from another recommendation system.
- **Feature augmentation:** The output from one technique is used as input features to another recommender system.
- **Meta-level:** Model that is developed through one recommender is used as input toward another.
- **Switching:** The recommendation approach switches based upon the situation. (Burke, 2002).

## **Evaluation:**

After carefully evaluating the pros and cons to each recommender system, a content-based recommender system was chosen for TCA. The system uses item-intrinsic metadata attributes as a basis for recommending similar artifacts. Due to the fact that we don't have user profiles implemented yet; this aspect was not included. The general idea for our recommender system is so that archivists, historians, and archeologists can find similar artifacts to study so that new pieces of history and relationships can be discovered.

## **TCA's Content-Based Recommender System:**

A content-based system was chosen because at this stage in the project, we only wanted to recommend similar artifacts to the artifact that is being observed. The content-based system was also much easier to develop conceptually, compared to the others, due to the fact that we only needed to include 3 sets of intrinsic data and develop an algorithm to calculate a relevance score. Since this prototype was entirely hypothetical and we are still in the prototype stage, we decided not to include user profiles just yet.

## **Metadata:**

For the prototype, 14 artifacts were randomly chosen from a larger collection located in the Oriental Institute at the University of Chicago. In the context of cultural research, there are currently no standards regarding metadata that cater to the digitization, storage and management, and dynamic creation of virtual exhibitions of 3D objects (Patel et al., 2005). Within the 14 artifacts selected, three categories of intrinsic metadata were chosen: the time period, the country of origin, and the material – which are generally accepted as presentation metadata attributes within the cultural context (Patel et al, 2005). These three attributes were chosen because out of

the artifacts that we chose; they were the three intrinsic categories that had values inputted for each artifact. Other attributes available were either extrinsic, which provide little value for what we are trying to achieve or had unknown values for certain artifacts. Each unique attribute was given a numeric value. These values can be seen in Figure 1. The way each value was determined was by looking at the relative similarity between each attribute. For the Time Period values, we associated 0 with the earliest time period, the Bronze Age. We associated 8 with the latest time period, WWII. For the Country of Origin values, we associated each value with relative distance between each country. For the material values, we associated 0 with copper, 1 with bronze, 2 with silver, and 3 with gold. Each attribute was given the same weight. This may change in future iterations as the user may be given a choice on which attribute he or she may want to put more emphasis on.

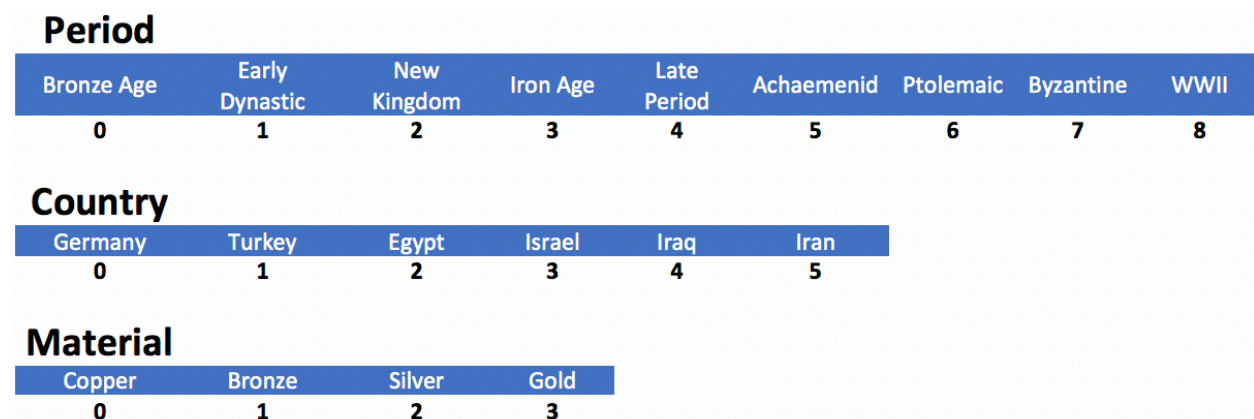


Figure 1



**Algorithm:**

To figure out comparisons based upon aspects of similarity (which artifacts were most similar to another), an equation was used that is shown in Figure 2 along with what each variable represents. The algorithm gives a relevance percentage associated with each artifact. The higher the percentage, the more similar an artifact is to the artifact being compared. The algorithm works by adding the inverse of the absolute difference of the Time Period Value of the Selected Artifact and the Time Period Value of the Artifact that is being compared, the inverse of the absolute difference between the Country of Origin Value of the selected artifact and the Country of Origin Value of the artifact that is being compared, and the inverse of the absolute difference between the Material Value of the selected artifact and the Material Value of the selected artifact that is being compared. The purpose of getting the inverse of each of these differences is so that we can add as many artifacts and attributes to the equation, and the equated value will not exceed 1. We find the absolute differences since the equation would not work if there were any negative numbers involved. We then divide the sum of these three values by 3 to find the relevance percentage. The table in Figure 3 shows the rankings of 14 selected artifacts in terms of potential similarity to a coin from the Byzantine Period. Each attribute is weighted the same, however this would ideally change in future iterations as user preferences are implemented into the algorithm.

$$\text{Relevance Percentage} = \frac{\frac{1}{\text{ABS}(x_1 - x_2 + 1)} + \frac{1}{\text{ABS}(y_1 - y_2 + 1)} + \frac{1}{\text{ABS}(z_1 - z_2 + 1)}}{3}$$

$x_1$  = Time Period Value of selected artifact

$x_2$  = Time Period Value of artifact that is being compared

$y_1$  = Country of Origin Value of selected artifact

$y_2$  = Country of Origin Value of artifact that is being compared

$z_1$  = Material Value of selected artifact

$z_2$  = Material Value of artifact that is being compared

Figure 2

Object Observed	Period	Country	Material	Period Score	Country Score	Material Score	Overall Score	Match
Coin - Byzantine	7	1	1					
Lamp - Byzantine	7	2	1	1.00	0.50	1.00	2.50	83%
Coin - Byzantine, (2)	7	1	2	1.00	1.00	0.50	2.50	83%
Sistrum - Ptolemaic	6	2	1	0.50	0.50	1.00	2.00	67%
Ring - WWII	8	0	1	0.50	0.50	1.00	2.00	67%
Statue, Deity - Late Period	4	2	1	0.25	0.50	1.00	1.75	58%
Axe - Bronze Age	0	1	0	0.13	1.00	0.50	1.63	54%
Buckle - Achaemenid	5	5	1	0.33	0.20	1.00	1.53	51%
Pin - Iron Age	3	5	1	0.20	0.20	1.00	1.40	47%
Foundation Peg - Early Dynastic	1	4	1	0.14	0.25	1.00	1.39	46%
Bead - Early Dynastic	1	4	1	0.14	0.25	1.00	1.39	46%
Mirror - New Kingdom	2	2	0	0.17	0.50	0.50	1.17	39%
Roundel - Early Dynastic	1	4	2	0.14	0.25	0.50	0.89	30%
Gaming Piece - Bronze Age	0	3	3	0.13	0.33	0.33	0.79	26%
Pendant - Bronze Age	0	3	3	0.13	0.33	0.33	0.79	26%

Figure 3

## **Discussion and Future Work:**

This algorithm is relatively basic in the sense that it only takes into account 3 metadata attributes. One of the biggest limitations has to do with the country attribute scoring system. Currently the scoring system and values associated with each country only work in this prototype because there are only 8 countries. As more countries are introduced, this scoring system will become obsolete since the distance between countries is not linear in comparison to the time period scoring system which is. One solution to this is to change the country scoring system to geographical coordinates. The algorithm would then instead take into account the exact geographical distance when comparing artifacts, making for a more accurate distance score. Introducing collaborative filtering should also be taken into account. One of the major themes that has been discussed in this project overall, is collaboration among researchers. This is not to say that the content-based system should be completely taken out, but rather create a hybrid system that uses traits from both. Grouping users into segments will help with relevant recommendations in an immense way. There is also the possibility of allowing for users to be able to interact with other users through their profiles to further promote this collaboration among researchers. Instead of simply recommending similar artifacts, the goal here should be to recommend artifacts that the user will want to see. This means taking into account many more datasets and possibly implementing a cluster model as well. Currently, this system has all attributes weighted equal, however as more artifacts are introduced, the prospect of allowing for users to have different attributes weighted more is a must. This could be done by implementing a menu within the user interface that gives the user all sorts of ways to customize his or her experience. Introducing a teaching salesman (Stolze and Ströbel, 2004) is also an option.

Something that helps the user select an artifact based on his or her needs. This interaction is divided into three parts: 1) preference discovery based on questions about needs, 2) preference optimization that allows for users to tweak the model: they can easily access preferences for specific features, and 3) the preference debugging phase, in addition to the top relevance score recommendations, two more choices are shown (Peinter, Viappiani, & Yorke-Smith, 2008). The first step however, should be to user test this type of recommender system among cultural researchers, archivists, and archeologists. This type of recommender system is relatively new in the field of cultural history, so getting opinions and data will help provide better guidance on where the pain points and opportunities are regarding recommendations among artifacts.

## **Conclusion:**

There are endless possibilities to implementing emerging technologies in the field of cultural research. There seems to be a tangible benefit in using Augmented Reality to help facilitate the research of artifacts and ephemera. However, there needs to be more qualitative and quantitative research done before cultural historians decide to move away from the status quo. More user studies will need to be conducted in order to learn more about pain points and opportunities that are faced everyday by these historians. Usability tests will also need to be conducted now that there is a functioning prototype.

In the case of recommender systems, a user study dedicated to this certainly needs to be conducted since there is very limited data on how these systems can benefit cultural research.

This is where the opportunity lies and where the TCA team can make a big impact in the fields of not only cultural research, but also digital media as a whole.

## References:

- (2013, October 2). ScopifyROM: New App Lands at the ROM. Retrieved from <https://www.rom.on.ca/en/about-us/newsroom/press-releases/scopifyrom-new-app-lands-at-the-rom>
- (n.d.). Jurascope, 2007. Retrieved from <https://artcom.de/en/project/jurascope/>
- (n.d.). Search Our Collections: The Oriental Institute of the University of Chicago. Retrieved from <https://oi-idb.uchicago.edu/results.php?tab=MC&q=IsOnDisplay:Yes+AND+SrchMaterials:Metal&sort=score+desc&start=250&rows=50&wt=json&facet=true&facet.mincount=1&&facet.sort=count>
- Anderle, D. (1978). The Artists Files At The New York Public Library. *ARLIS/NA Newsletter*, 6(6), 101–102. doi: 10.1086/arlisanews.6.6.27946050
- Anghelescu, H. G. B. (2001). A Bit of History in the Library Attic. *Collection Management*, 25(4), 61–75. doi: 10.1300/j105v25n04\_07
- Balabanović, M., & Shoham, Y. (1997). Fab: content-based, collaborative recommendation. *Communications of the ACM*, 40(3), 66–72. doi: 10.1145/245108.245124
- Boehmer, J., Jung, Y., & Wash, R. (2015). e-Commerce Recommender Systems. *The International Encyclopedia of Digital Communication and Society*, 9999(9999), 1-8.
- Bogers, T., & van den Bosch, A. (2009). Collaborative and Content-based Filtering for Item Recommendation on Social Bookmarking Websites Toine. *Proceedings of the ACM RecSys'09 Workshop on Recommender Systems & the Social Web*, 9–16.
- Burke, R. (2002). Hybrid Web Recommender Systems. *The Adaptive Web Lecture Notes in Computer Science*, 377–408. doi: 10.1007/978-3-540-72079-9\_12
- Chu, J. H., Clifton, P., Harley, D., Pavao, J., & Mazalek, A. (2015). Mapping Place. *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction - TEI 14*. doi: 10.1145/2677199.2680559
- Claypool et al., M. (1999). Worcester.
- Deldjoo, Y., Elahi, M., Cremonesi, P., Garzotto, F., Piazzolla, P., & Quadrana, M. (2016). Content-Based Video Recommendation System Based on Stylistic Visual Features. *Journal on Data Semantics*, 5(2), 99–113. doi: 10.1007/s13740-016-0060-9

- Dieck, M. C. T., & Jung, T. (2015). A theoretical model of mobile augmented reality acceptance in urban heritage tourism. *Current Issues in Tourism*, 21(2), 154–174. doi: 10.1080/13683500.2015.1070801
- Fowler, H. W. (1949). The concise Oxford dictionary of current English Oxford Dictionary. *The concise Oxford dictionary of current English Oxford Dictionary*. Oxford: Clarendon.
- Glushko, R. J., McPherson, K., Greenberg, R., & Mayernik, M. (2013). *The Discipline of Organizing*. Cambridge (Massachusetts): MIT Press.
- He, Z., Wu, L., & Li, X. (2018). When art meets tech: The role of augmented reality in enhancing museum experiences and purchase intentions. *Tourism Management*, 68(Complete), 127-139. doi:10.1016/j.tourman.2018.03.003
- Horn, M. S., & Jacob, R. J. K. (2007). Designing tangible programming languages for classroom use. *Proceedings of the 1st International Conference on Tangible and Embedded Interaction - TEI 07*. doi: 10.1145/1226969.1227003
- Kalinda, D. (2019). *Tca Formative Study. TCA Formative Study*.
- Kress, G. R. (2010). *Multimodality: a social semiotic approach to contemporary communication*. London: Routledge.
- Kwan, J., Chu, J. H., Harley, D., McBride, M., & Mazalek, A. (2016). Grasping Cultural Context through Multisensory Interactions. *Proceedings of the TEI 16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction - TEI 16*. doi: 10.1145/2839462.2856537
- Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1), 76–80. doi: 10.1109/mic.2003.1167344
- Mussell, J. (2012). The Passing Of Print. *Media History*, 18(1), 77–92. doi: 10.1080/13688804.2011.637666
- Patel, M., White, M., Mourkoussis, N., Walczak, K., Wojciechowski, R., & Chmielewski, J. (2005). Metadata requirements for digital museum environments. *International Journal on Digital Libraries*, 5(3), 179–192. doi: 10.1007/s00799-004-0104-x
- Pazzani, M. J., & Billsus, D. (2007). Content-Based Recommendation Systems. *The Adaptive Web Lecture Notes in Computer Science*, 325–341. doi: 10.1007/978-3-540-72079-9\_10
- Peintner, B., Viappiani, P., & Yorke-Smith, N. (2008). Preferences in Interactive Systems: Technical Challenges and Case Studies. *AI Magazine*, 29(4), 13. doi: 10.1609/aimag.v29i4.2197

- Rafsanjani, A. H. N., Salim, N., Aghdam, A., & Fard, K. (2013). Recommendation Systems: A review. *International Journal of Computation*, 3(5), 47–52.
- Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens. *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work - CSCW 94*. doi: 10.1145/192844.192905
- Shardanand, U., & Maes, P. (1995). Social information filtering. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI 95*, 210–217. doi: 10.1145/223904.223931
- Stolze, M., & Ströbel, M. (2004). Retrieved from [https://pdfs.semanticscholar.org/2810/1a7434027cb4e238ef286f16bffd9e757e1.pdf?\\_ga=2.116519238.1753928892.1567188159-579555047.1566508053](https://pdfs.semanticscholar.org/2810/1a7434027cb4e238ef286f16bffd9e757e1.pdf?_ga=2.116519238.1753928892.1567188159-579555047.1566508053)
- Wei, J., He, J., Chen, K., Zhou, Y., & Tang, Z. (2017). Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications*, 69, 29–39. doi: 10.1016/j.eswa.2016.09.040