Journal of eScience Librarianship
putting the pieces together: theory and practice

Automatic Expansion of Metadata Standards for Historic Costume Collections

Caleb McIrvin, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
Chreston Miller, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, chmille3@vt.edu (iD)
Dina Smith-Glaviana, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
Wen Nie Ng, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Abstract

Objective: This project focuses on Artificial Intelligence (AI) supported enhancement of descriptive metadata for fashion collections (otherwise known as costume or dress and textile collections) through expanding costume-specific controlled terms. The authors use Natural Language Processing (NLP) techniques along with a human-in-the-loop process to support selection of descriptive terms for inclusion in the controlled terms of a metadata schema. This project seeks to expand upon existing domain-specific schema, Costume Core, by enhancing the schema with a comprehensive set of descriptors. This enhancement will allow for more accurate and detailed descriptions of costume artifacts. This article describes this process and the outcomes of AI approaches for providing this metadata expansion, who this process is for, ethical considerations, and lessons learned.

Citation: McIrvin, Caleb, Chreston Miller, Dina Smith-Glaviana, and Wen Nie Ng. 2024. "Automatic Expansion of Metadata Standards for Historic Costume Collections." *Journal of eScience Librarianship* 13 (1): e845. https://doi.org/10.7191/jeslib.845.

Data Availability: The data presented in this study are available on request from the corresponding author.

Abstract continued

Methods: We approached our problem with an investigation into using word embeddings to aid in supporting the suggesting of new metadata terms. Several word embedding models were applied with the more descriptive one chosen for final use in a human-in-the-loop selection process. This selection process provided domain experts to identify which terms chosen by the model are of relevant value. We then compare what was chosen by the domain experts and what the model produced to get an idea as to the value the model provides in the process of metadata expansion.

Results: The metadata expansion process was a success. An AI supported process aided domain experts in choosing relevant terms to include in their metadata schema. Therefore, the results were a methodology for using identified AI models for the problem, an interactive system to aid the domain experts (software system), and how to evaluate the results.

Conclusion: The application of AI technologies (word embeddings) provided a successful pipeline for supporting domain experts to expand the metadata schema with additional descriptors. Enhancing the metadata schema with additional descriptors improves its usability for fashion collection managers and allows for a more precise description of the artifacts. As a result, many new terms that were expertly chosen were added to the metadata schema.

Introduction

We postulate that problems in cataloging efforts in the historical costuming domain can be mitigated through implementation of a standardized metadata schema. Existing metadata schemas that utilize controlled descriptive terminology for fashion artifacts, such as historic costume or dress, and items related to the process and product of dressing the body (Eicher and Evenson 2014), which include clothing, textiles, and accessories, are often constrained by the insufficient number of description fields and a limited vocabulary set. By expanding the number of terms available using Natural Language Processing methods, we can develop high-quality, consistent metadata enabling better data sharing across collections and increasing cataloging accuracy, resulting in improved dress record searchability. Confirming generated descriptor choices via a human-in-the-loop approach allowed us to alleviate ethical concerns surrounding the quality of the descriptors chosen for our updated metadata schema. This approach also allowed us to ensure our descriptor selections for the updated schema drew from an ethnically diverse array of sources while avoiding misleading or culturally insensitive terms. Additionally, we emphasized adding terms that would allow for inclusive language, for example enhancing the use of colloquial terms such as "robe," which could refer either to "bathrobe" or traditional Chinese robes, by adding more precise terms.

Project details

This project was a collaboration between the Virginia Tech Fashion Merchandising and Design (FMD) department and Virginia Tech University Libraries. An expert in dress and students from the FMD department participated in selecting terms. The expert supervised the student assistants and was the final decision-making authority on choosing the descriptors used to expand the metadata schema. The Data and Informatics Consultant within the University Libraries worked with an undergraduate Computer Science student to develop the Natural Language Processing (NLP) approach and provide analysis of the resulting chosen descriptors. Our dress domain expert collaborated with our digital collections specialists within the University Libraries to support the process of finalizing the expanded metadata schema. The collection used for this work was the Oris Glisson Historic Costume and Textile Collection. The collection was a suitable collection since the dress domain expert had previously cataloged the items using Costume Core (Kirkland 2018), the metadata schema our project strives to expand. To aid in this expansion, we used a pretrained NLP model to generate word embeddings. These word embeddings can be used to identify descriptors that have conceptually similar meanings but also can identify descriptors that are a little conceptually "further" away. This also allows us to introduce descriptor diversity and explore terms that were previously not present in the schema. To help with sifting through the identified potential descriptors, we used a lightweight server to host an application we developed to aid in sifting through NLP suggested descriptors. The main challenge we encountered during our pipeline was how to best share the NLP results with the students and the dress domain expert for selection and confirmation. We additionally emphasized ethical consciousness when generating the terms, ensuring the terms we generated and confirmed for addition to the schema would be valuable for alleviating the confusion of costume collection users..

Background

An accurate understanding of historical cultural eras allows historians / fashion experts to better make judgments about social values of the period. To achieve this understanding, it is important to consider the historical costuming conventions of the day. Recognizing the importance of costuming artifacts to a proper cultural understanding, many universities, libraries, and museums have amassed large collections of historical and contemporary dress items. However, these pieces are frequently poorly described, with little to no interaction between collections on how to standardize piece descriptions. Although an increased emphasis on the advantages associated with collection digitization (Zeng 1999) has started to appear in recent years, there is a dearth of research in this area regarding standardization of piece description.

Metadata expansion efforts can be found across a variety of fields, though relatively few use NLP to improve development speed. Specifically in the fashion domain, several pushes have been made towards a unified ontology through metadata expansion. The Kent State University collection that utilized Dublin Core was one such effort (Zeng 1999). Additionally, Costume Core (Kirkland 2018), upon which our metadata schema is built, serves as an effective groundwork for metadata expansions. Both of these schemas, however,

suffer from a lack of granularity due to fewer metadata levels or controlled terms than required for accurate cataloging. In (Valentino 2017), another such linked metadata schema is presented. Ryerson University has also made efforts to more clearly display their fashion collection (Eichenlaub, Morgan, and Masak-Mida 2014) using Dublin Core. A crowdsourcing effort based on Costume Core uses survey data as a promising tool to combat the lack of generality present in many of the above schemas (Cai et al. 2012), but their work is specific to Chinese-style costumes.

Numerous efforts have been made towards ontology development as well, both in the fashion domain and elsewhere. A work in the fashion domain (Bollacker, Díaz-Rodríguez, and Li 2016) claims that ontologies taking only garment attributes into account provide insufficient information and seeks to build a subjective influence network to incorporate more data into the ontology. In Novalija and Leban 2013, work is done to construct an ontology of designer garments, connecting pieces based on Wikipedia link structures.

The primary benefits of a more comprehensive metadata schema include improvement of searchability and discoverability. Streamlining the expansion of metadata schema can be efficiently achieved by employing NLP techniques. Outfit2Vec (Jaradat, Dokoohaki, and Matskin 2020) uses clothing metadata to build machine learning models that can better recommend garments to consumers. Tahsin in (Tahsin et al. 2016) uses NLP to extract geographic metadata from text corpuses to increase location specificity. One approach towards solving this problem, taken by (Cai et al. 2012), is crowdsourcing—researchers use NLP techniques coupled with input from 100 students with regards to metadata element importance. However, the techniques used in the above paper did not generate new descriptors, only assisted in confirming previously selected categorizations.

A standardized set of descriptors needs to be developed in order to allow visitors to digital collections to quickly search for a particular type of garment. Such a set of descriptors is Costume Core, as mentioned above. Unfortunately, the scope of Costume Core is limited by its size – many potentially useful descriptors and several valuable categories are left out, restricting the utility of the project. In our work, we seek to expand upon the Costume Core vocabulary by using NLP techniques to efficiently identify new descriptors previously not included in the schema, expanding the size of the Costume Core vocabulary to enhance digital collection cataloging and search capabilities.

Methodology

Our process to identify high-quality costume descriptors consisted of multiple steps. Firstly, we generated hundreds of potential descriptors using word embeddings, an NLP technique, from the initial Costume Core schema. Afterwards, we used our Model Output Confirmative Helper Application (MOCHA) to facilitate the review process by our trained fashion students. Finally, our domain expert reviewed all selections, trimming the choices down to ensure quality.

To obtain the initial descriptors used in similarity generation, we adapted a popular set of descriptors found in the Costume Core vocabulary commonly used in fashion metadata description tasks (Kirkland 2018). While the keywords contained in the vocabulary were acceptable in many cases, some categories could be removed, as new, meaningful descriptors were unlikely to be generated for the category. One such example is the "Socio-economic class" category, which, in the original Costume Core vocabulary, contains the descriptors "middle class," "upper class," and "working class." As models are unlikely to create useful descriptors for this category, this category and similar others were removed from our analysis. In addition, slight manual lemmatizations (such as changing "coatdresses" to "coatdress") were made to generate more accurate predictions.

Data preprocessing

Additional data pre-processing was necessary to convert keywords to a format usable by our selected models. As the keywords were initially stored in an Excel file, we needed to convert this to a file more conducive to model format. To accomplish this, we removed characters our models wouldn't recognize, such as "$" and "!," from the file using a regular expression. Additionally, we performed some minor manual tweaking of the initial keyword selections to maximize the number of potential new descriptors output by our model.

Model selection

Similar descriptors can be generated efficiently using a cosine similarity method, a technique that uses distance between vector representations of tokens, also known as word embeddings, as a measure of similarity between words. Gensim's Word2Vec, a Python library developed to work with word embeddings, provides functionality to quickly and easily generate the most similar words for a model from a vocabulary. For accurate comparison testing, we tested three separate gensim Word2Vec models, specifically the Google News, Mpnet, and Sentence-t5 models. While these models were not specifically fine-tuned on fashion literature, they were still capable of accurately modeling the relationships between descriptors, as seen in Figure 2. After initial data analysis and feedback from our reviewers, we elected to narrow down our focus to emphasize solely the Google News model.

Initial Costume Core Network Data Visualization

In order to gain an understanding of the different connections between Costume Core keywords as captured by the models, we create visualizations of relationships between keywords. To create these visualizations, we first load in the Costume Core keywords, organized by category, as well as our three separate models. For our model, we iterate over model keywords twice, calculating the cosine similarity values between keywords and storing the values greater than a set percentage in a tabular format (Figure 1).

	keyword_1	keyword_2	weight
0	blouse	dress	0.865597
1	blouse	shirt	0.890284
2	blouse	skirt	0.875139
3	blouse	sweater	0.879506
4	blouse	underwear	0.854485
...	...	...	...
1593	stripe	paint	0.851266
1594	stripe	ribbon	0.860757
1595	stripe	seam	0.851272
1596	stripe	uniform	0.866175
1597	stripe	plaid	0.856900

Figure 1: Cosine similarity weights between the Costume Core keywords (keyword_1) and their most similar tokens (keyword_2) as predicted by the model (weight).

We then use the Python NetworkX (Hagberg, Swart, and Schult 2008) library to export a graph as a format readable by the Orange (Demsˇar et al. 2013) software. In addition, we create separate files to specify Orange visualization format by providing additional keyword category information. To assess the quality of the model's representations, we graphed our initial descriptors by category, connecting them by the strength of their cosine similarity weights. As seen in Figure 2, the model performs reasonably well at clustering similar Costume Core keywords, as evidenced by the tightly clustered "color" and "material" categories.
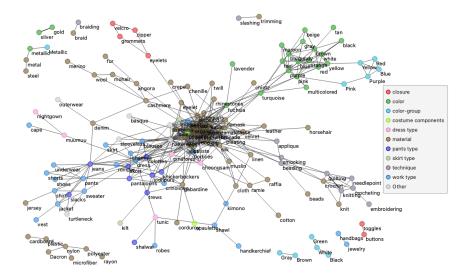

Figure 2: Google News representation of Costume Core keywords. Colors represent categories of descriptors, e.g. material, neckline, technique, etc.

Descriptor generation

Once we had our descriptors and models loaded into the correct format, we ran the descriptors through our Word2Vec models to generate new potential descriptors for later analysis. For each valid Costume Core descriptor, we generated the top 25 most similar potential descriptors. After the models finished generating similar potential descriptors, we saved the results to separate .csv files for storage later—these similar potential descriptors contained important information for analysis.

Supporting Human in the Loop term selection

Before the potential descriptors generated by the model could be released, we needed a way to confirm the descriptors were actually valid descriptors for historical costuming metadata. We determined that the most effective way of confirming these descriptors was a human-in-the-loop approach, in which fashion metadata experts would check over the model selected descriptors, selecting the most accurate / relevant descriptors. This approach is partially motivated by the fact that their selections would allow us to calculate several statistics measuring the accuracy of the models, definitively demonstrating the effectiveness of NLP models in generating new descriptors in the fashion metadata domain. Our process is displayed in more detail in Figure 3.
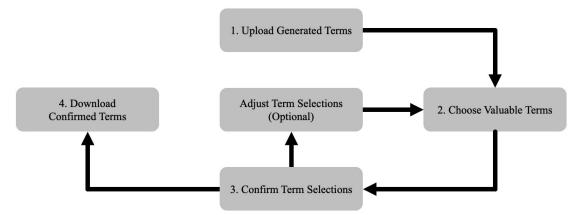

Figure 3: Descriptor confirmation loop using a Model Output Confirmative Helper Application (MOCHA).

To expedite the process and allow our domain experts to easily and efficiently process the generated descriptors, we created a Model Output Confirmative Helper Application (MOCHA) to present the model-generated descriptors. To operate the web application, users load in model-generated words, at which point, they can visually select a subset of descriptors to classify as confirmed descriptors. These model-generated words came in the form of the top 25 descriptors similar to each token in the Costume Core vocabulary. Multiple usability enhancements allow users to navigate quickly between pages of descriptors, in case labeling needs to be broken up into multiple sessions. Functionality for clearing selected descriptors is added to allow for greater flexibility. After all descriptors have been confirmed / rejected, or after a labeling session has ended, users can download the descriptors they've confirmed for analysis.
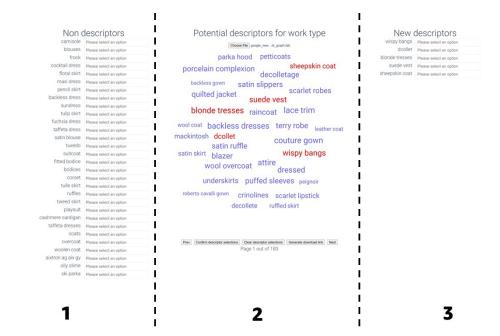
Figure 4: MOCHA application. Column 1 stores descriptors not selected by our reviewers, column 2 stores descriptors currently being processed, and column 3 stores descriptors confirmed by our reviewers.

To confirm our selections, we first had a group of three trained fashion students confirm model-generated descriptors using the web application, as seen above. To use the web application, descriptors were loaded in as textual data, where they appeared as a word cloud in column 2. After clicking on terms to select beneficial ones, they would appear in column 3 and could be downloaded for analysis. Terms which remained unselected were placed in column 1, where they would be discarded. After our students had finalized reviewing potential descriptors, our domain expert edited / revised the students' selections. After finalizing revision, we converted the collected confirmed descriptors to a form more suitable for analysis, combining the descriptors generated by the domain experts into a single file. Due to input from our domain expert / students on the relevance of the generated descriptors produced by the three models, we decided to proceed with the Google News model, as the descriptors generated were found to be more relevant to the domain-specific task. As a result, analyses presented in the following section were obtained from descriptors generated by the Google News Word2Vec model.

Results

As mentioned above, the domain experts processed the top 25 similar words for each word in the Costume Core vocabulary. We also created files of the top 20, 15, 10, and 5 most similar words, as generated by the models. Below are the plotted graphs of cosine similarity score (x-axis) versus the percentage of words having that cosine similarity score (y-axis), for both model-generated and reviewer-confirmed tokens. As expected, the two graphs in Figure 5 show that the top 5 most similar words generated by the models have higher cosine similarity values on average than the top 25 most similar words.
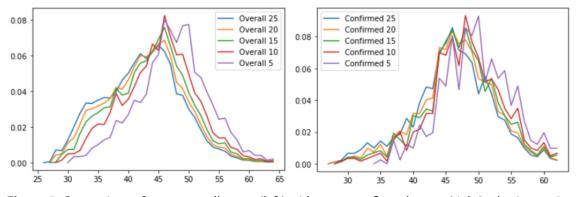
Figure 5: Comparison of top n overall terms (left) with top n confirmed terms (right), plotting cosine similarity score against the percentage of descriptors with the same score.

However, the measure of the model's efficacy in predicting descriptors is displayed in the gap between the cosine similarity scores of the confirmed descriptors and the overall generated descriptors. If the model's predictions are accurate, we would expect words with higher cosine similarity scores to have a larger chance of being confirmed by our domain experts.
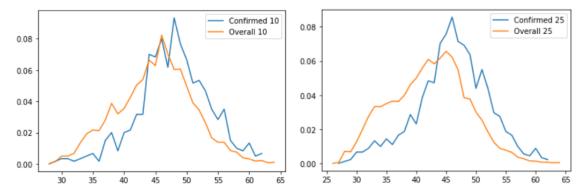

Figure 5: Comparison of top n overall terms (left) with top n confirmed terms (right), plotting cosine similarity score against the percentage of descriptors with the same score.

As seen from Figure 6, there is a clear distinction between the original, model-generated descriptors and the descriptors that were actually confirmed. To further demonstrate this difference, we show the relative averages of confirmed and overall descriptors in Table 1. Consistently, the confirmed descriptors had a higher cosine similarity score on average than the overall model-generated descriptors.

In Table 1, we see that consistently, across multiple different groups of generated vs confirmed terms, the confirmed terms had higher cosine similarity scores on average. This gap between generated and confirmed terms leads us to believe that our model was effective at generating high-quality descriptors, as descriptors which the model deemed more likely to be included in the schema (higher cosine similarity score) were indeed selected more frequently on average, indicating that the model performed well here.

Table 1: Relative averages of confirmed and overall descriptors.

	Top 25	Top 20	Top 15	Top 10	Top 5
Confirmed CS Score	0.6063	0.6145	0.6244	0.6329	0.6575
Overall CS Score	0.5688	0.5777	0.5901	0.6071	0.6370
Hit rate	14.2%	15.6%	17.8%	21.3%	27.5%
Additional statistics on percentages of term cosine similarity scores - hit rate = % of generated descriptors that were confirmed.					

Ethical considerations

The main objective of this work was to minimize the confusion felt by many costume collection users by providing an expanded set of metadata descriptors for use in cataloging efforts. An important priority for us was to ensure that our generated descriptors were diverse enough to encompass a wide array of periods and cultures, in order to avoid discriminatory exclusions of garment types. In pursuit of this goal, we used models trained on a wide variety of different sources to generate our descriptors, ensuring that these models would have been exposed to data from many different areas, hopefully alleviating many of these potential concerns. In addition, our human-in-the-loop approach allowed us to exclude potentially negative or harmful terms from being added to our schema by providing an additional layer of protection.

A point that may be considered is the impact of the human-in-the-loop approach to term biases—did the fact that we had one final reviewer potentially cause an opportunity for bias to be created? To combat this, our initial term selections were made by separate reviewers, so that our fashion metadata expert only made final confirmation decisions on descriptors judged to be valuable from a variety of sources. Additionally, our fashion metadata expert has extensive experience with different terminologies and is knowledgeable on best practices in the field, two characteristics which help to ensure the quality of our schema. These measures, taken on both the generation and filtration ends of the process, serve to minimize the risk of potentially harmful terms being added to the schema. However, despite these precautions, it may still be valuable to consolidate opinions from a variety of stakeholders, such as users, curators, and collection managers.

Who is affected by this project?

Our efforts to use NLP to select accurately controlled vocabularies can benefit any professional or institution that manages fashion collections, artifacts including private archives and museums owned by fashion companies such as Michael Kors and Armani Silos (Franceschini 2019), regional historical societies/ museums, and university fashion study collections (Green and Reddy-Best 2022) as it would allow them to select from controlled vocabularies that precisely describe and catalog artifacts. This project can also

benefit digital librarians and other personnel collaborating with fashion domain professionals to create online digital libraries. Because the use of NLP has led to the addition of accurate/sufficient metadata elements to Costume Core, which provides a means for structuring data (Kirkland 2018), digital librarians can more easily map Costume Core vocabularies to those of pre-existing schemas such as Dublin Core when preparing to export metadata to online portals and aggregators.

Several university fashion collections have committed to using the Costume Core metadata schema to support two inter-institutional projects that aim to contribute to initiatives to standardize metadata across the historic dress and fashion domain (Kirkland 2018). In addition, standardizing the metadata has implications for online users of digital fashion collections. Without standardized metadata, online users may experience failed searches which limit the reach and accessibility of online fashion digital collections. Thus, the benefits of NLP may extend to online users as it contributes to an initiative to standardize metadata within the fashion domain.

Lessons learned and future work

While our process was fairly straightforward, there were a few issues we encountered along the way that, if not properly addressed, could have become stumbling blocks. One such area was our choice of model used to generate the terms. Two models we originally attempted to use were deemed unsatisfactory for our use case due to the low quality of terms generated. However, trying a variety of models allowed us to select a model, Google News, with excellent representations of our space. The model outputs and processing code are currently being prepared for dissemination, but the categorization tool code (MOCHA) is available.

Another issue we encountered was that of sharing our web application. Bundling up the tool and sending the files via a messaging service seemed likely to cause version control issues as well as potentially being difficult to set up for non-technical users. As a result, we found it efficient to set our application up on AWS Lightsail, a virtual server designed for running lighter-weight applications like MOCHA. This provided an easily accessible platform for our fashion students and domain experts to use while allowing us to perform minor updates easily, without needing to resend large files after every update.

As for future work, we would like to create a visual thesaurus tool with Costume Core metadata (old and new) to help catalogers choose the most accurate term(s). Also, because (Kirkland et al. 2023) found that users searched for holdings on historical collection websites using retail or lay terminology, future work will / may include reviewing fashion lay and retail terms, comparing them with other established controlled vocabularies, including the International Council of Museums (ICOM) Vocabulary of Basic Terms for Cataloguing Costume and Getty Art and Architecture Thesaurus, and adding them to Costume Core.

Data Availability

The data presented in this study are available on request from the corresponding author.

Acknowledgements

Competing Interests

The authors declare that they have no competing interests.

References

Bollacker, Kurt, Natalia Díaz-Rodríguez, and Xian Li. 2016. "Beyond Clothing Ontologies: Modeling Fashion with Subjective Influence Networks." Paper presented at *Knowledge Discovery and Data Mining: Machine learning meets fashion: Data, algorithms and analytics for the fashion industry*, San Francisco, August 13-17. https://www.researchgate.net/publication/304762196_Beyond_Clothing_Ontologies_Modeling_Fashion_with_Subjective_Influence_Networks.

Cai, Yundong, Yin-Leng Theng, Qimeng Cai, Zhi Ling, Yangbo Ou, and Gladys Theng. 2012. "Crowdsourcing Metadata Schema Generation for Chinese-Style Costume Digital Library." In *The Outreach of Digital Libraries: A Globalized Resource Network*, edited by Hsin-Hsi Chen and Gobinda Chowdhury, 97–105. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-34752-8_13.

Demsˇar, Janez, Tomazˇ Curk, Alesˇ Erjavec, Janez Demsar, Tomaz Curk, Ales Erjave, Crt Gorup, et al. 2013. "Orange: Data Mining Toolbox in Python." In *Journal of Machine Learning Research* 14 (2013): 2349–2353. https://jmlr.org/papers/volume14/demsar13a/demsar13a.pdf.

Eichenlaub, Naomi, Marina Morgan, and Ingrid Masak-Mida. 2014. "Undressing Fashion Metadata: Ryerson University Fashion Research Collection." Paper presented at *International Conference on Dublin Core and Metadata Applications*, October, 191–195. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=4864f92bb4b5781847a5d3f2e3691d0a68c7e290.

Eicher, Joanne B., and Sandra Lee Evenson. 2014. *The Visible Self: Global Perspectives on Dress, Culture and Society*. USA: Bloomsbury Publishing. https://www.isbns.net/isbn/9781609018702.

Franceschini, Marta. 2019. "Navigating Fashion: On the Role of Digital Fashion Archives in the Preservation, Classification and Dissemination of Fashion Heritage." *Critical Studies in Fashion & Beauty* 10 (1): 69–90. https://doi.org/10.1386/csfb.10.1.69_1.

Green, Denise Nicole, and Kelly L. Reddy-Best. 2022. "Curatorial Reflections in North American University Fashion Collections: Challenging the Canon." *Critical Studies in Fashion & Beauty* 13 (1): 7–20. https://doi.org/10.1386/csfb_00035_2.

Hagberg, Aric, Pieter J. Swart, and Daniel A. Schult. 2008. "Exploring Network Structure, Dynamics, and Function Using NetworkX." LA-UR-08-5495. Los Alamos National Laboratory (LANL), Los Alamos, NM (United States). https://www.osti.gov/biblio/960616.

Jaradat, Shatha, Nima Dokoohaki, and Mihhail Matskin. 2020. "Outfit2Vec: Incorporating Clothing Hierarchical MetaData into Outfits' Recommendation." In *Fashion Recommender Systems*, edited by Nima Dokoohaki, 87–107. Lecture Notes in Social Networks. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-55218-3_5.

Kirkland, Arden. 2018. "Costume Core: Metadata for Historic Clothing." *Visual Resources Association Bulletin* 45 (2): 6. https://online.vraweb.org/index.php/vrab/article/view/36.

Kirkland, Arden, Monica Sklar, Clare Sauro, Leon Wiebers, Sara Idacavage, and Julia Mun. 2023. "'I'm Not Searching the Right Words': User Experience Searching Historic Clothing Collection Websites." *The International Journal of the Inclusive Museum* 16 (1): 119–146. https://doi.org/10.18848/1835-2014/CGP/v16i01/119-146.

Novalija, Inna, and Gregor Leban. 2013. "Applying NLP for Building Domain Ontology: Fashion Collection." Paper presented at *Conference on Data Mining and Data Warehouses (SiKDD)*. https://ailab.ijs.si/dunja/SiKDD2013/Papers/Novalija-FashionCollection.pdf.

Tahsin, Tasnia, Davy Weissenbacher, Robert Rivera, Rachel Beard, Mari Firago, Garrick Wallstrom, Matthew Scotch, and Graciela Gonzalez. 2016. "A High-Precision Rule-Based Extraction System for Expanding Geospatial Metadata in GenBank Records." *Journal of the American Medical Informatics Association* 23 (5): 934–941. https://doi.org/10.1093/jamia/ocv172.

Valentino, Maura. 2017. "Linked Data Metadata for Digital Clothing Collections." *Journal of Web Librarianship* 11 (3–4): 231–240. https://doi.org/10.1080/19322909.2017.1359135.

Zeng, Marcia Lei. 1999. "Metadata Elements for Object Description and Representation: A Case Report from a Digitized Historical Fashion Collection Project." *Journal of the American Society for Information Science* 50 (13): 1193–1208. https://doi.org/10.1002/(SICI)1097-4571(1999)50:13<1193::AID-ASI5>3.0.CO;2-C.