# Fashion Terminology Today Describe your heritage collections with an eye on the future

Ykje Wildenborg MoMu – Fashion Museum of the Province of Antwerp, Belgium Europeana Fashion, Modemuze

# Abstract:

This article was written for 'non-techy people', or people with a basic knowledge of information technology, interested in preparing their fashion heritage metadata for publication online.

Publishing fashion heritage on the web brings about the undisputed need for a shared vocabulary, especially when merged. This is not only a question of multilingualism. Between collections and even within collections different words have been used to describe, for example, the same types of objects, materials or techniques. In professional language: the data often is "unclean".

Linked Data is the name of a development in information technology that could prove useful for fashion collecting institutions. It means that the descriptions of collections, in a computer readable format, have a structure that is extremely easy for the device to read. As alien as it may sound, Linked Data practices are already used by the data departments of larger museums, companies and governmental institutions around the world. It eliminates the need for translation or actual changing of the content of databases. It only concerns 'labeling' of terms in databases with an identifier.

With this in mind, MoMu, the fashion museum of Antwerp, Belgium, is carrying out a terminology project in Flanders and the Netherlands, in order to motivate institutions to accomplish the task of labeling their terms. This article concludes with some of the experiences of this adventure, but firstly elucidates the context of the situation.

#### Content:

Introduction / Some context / Visual Thesaurus for Fashion and Costume / Conclusion / References

#### Introduction

Upon an internet search for the term "ball gown" the results invariably yield numerous shiny modern examples of this garment, but may also reveal striking heritage objects kept by museums. In addition to the individual websites of these establishments, there is a growing number of portals that bring a plurality of heritage holdings together. 2014 marked the launch of Europeana Fashion, a website unifying the fashion and costume collections of 30 European institutions. In Sydney, The Australian Dress Register (australiandressregister.org) was set up, inviting collection-keepers to add Australian clothes and their stories, whereas Modemuze (modemuze.nl) integrates the collections of twelve Dutch museums. Most likely, there are other initiatives around the globe, existing or in the making, that concentrate on accumulating fashion collections online.

#### Some context

During the 150 years in which costumes have been collected, the institutions collecting them did not constantly possess the money and the specialist knowledge to correctly describe and document this form of heritage. It is painfully redundant to mention its frequently subordinated treatment in museums and the mistrusted legitimacy of its study.

With the 'sudden' publication of these records online, the data pollution as a result of years and years of this process becomes apparent. For some, this is a reason to show only a selection of their objects; others take this opportunity to show what they have and put everything out there, stressing the need for action. The author has perceived errors that have

slipped in – among others – designers called "Emilio Puzzi" and "Vivian Westwood" or collection items named "T-shits". These mistakes do not help the cataloger indexing an object nor do they help the searching end user.

# Aggregators and controlled vocabularies

Fashion heritage is being made available through a diversity of websites. They can easily be divided in two categories. The first concerns collection websites or thematic portals, based on the collections of an individual organization. The second corresponds to websites that bring heterogeneous sets of metadata together, made available by different institutions. They are also called *aggregators*, since they bring together or *aggregate* metadata. A few monolingual examples of the aggregators for fashion heritage are the Dutch portal Modemuze and the Australian Dress Register.

A multilingual aggregator is Europeana Fashion, a heritage portal created between 2012 and 2015 with a still increasing amount of records. During this project, which was set up independently from Europeana itself, a ten-language thesaurus was created for the fields (or *facets*) 'Object type', 'Materials', 'Techniques' and 'Colors'. It greatly improves search results for the data queried through the search box or user friendly filters. Despite the practical and effective outcomes of this approach, this is not a long term solution. All the words used for describing a collection in any of the facets mentioned, have to be "tagged" with a term from the thesaurus during the upload and publication process – an extensive yet unsustainable task.

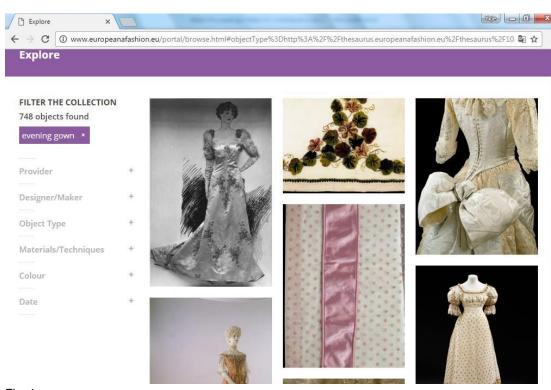


Fig. 1: Europeana Fashion is a European aggregator for fashion heritage, including both private and public collections.

So far, it is uncommon to see museums and institutions share the same vocabulary. However, this is the situation we need to pursue if we want to "sit at a computer, and search across the collections of historic clothing throughout the world on one site" – an appealing vision sketched in 2015's fall issue of *Dress* (Kirkland et al. 2015, 120). To make this scenario become a reality, we should start looking into the use of controlled, machine readable vocabularies. This is possible, because many vocabularies were recently made available 'open source' – Getty's Art & Architecture Thesaurus (AAT) has been free to use for everyone since March 2014. The good thing of this method is that everyone can keep their own terminology. The next subchapters will discuss the how and why of this approach in more detail.

#### Linked Data?

Using a controlled vocabulary can prevent making mistakes, or, registering numerous terms for one type of object. A controlled vocabulary's main goals are to "gather variant terms and synonyms for concepts and link concepts in a logical order or sort them into categories" (Harpring, 2010, 12). As a piece of software, it can aid search engines in gathering the records a user is looking for. For naturally, computers do not *think* and cannot establish links, where humans can: we have to tell them when object names or materials are synonymous. A 'combinaison' and a 'combination' for a computer system, don't have much in common. Of course one can give a list to their web developer, listing 'combination equals combinaison' and other synonyms, but now a less ephemeral solution exists. It corresponds to taking charge ourselves, making use of a controlled open vocabulary while cataloguing – an idea part of a development called "Linked Data".



Fig. 2: When searching for 'combinaison' in the Dutch heritage portal Modemuze, the result will not include similar garments described as 'combination'.

Linked Data concerns "a set of best practices for the publication of structured data on the web" (Hooland and Verborgh 2014, 3). This piece of information technology, introduced by the renowned Tim Berners-Lee as early as 2006, is now a hot topic for anyone interested in databases meant for online publication (see w3.org/DesignIssues/LinkedData.html). As we can expect, this is an important subject on the agendas of thematic and national aggregators such as Europeana (europeana.eu) or DPLA (The Digital Public Library of America, see https://dp.la/).

It is unnecessary to immerse oneself in the theory of Linked Data in order to be able to work with it. The important thing to comprehend – for the subject of this paper – is the fact that terms, in the 'controlled' text fields of our databases, should be tagged with an identifier. This ID corresponds to a 'concept' - not a word, but a thing - and gives computer systems access to its preferred name, alternative terms, translations and a description. This information is published somewhere on the web; that is why this ID looks like a sort of URL called a Uniform Resource Identifier (URI). It is the home of a 'dress', 'jacket', or 'skirt'; they are records in a controlled, but open (i.e. free to use) vocabulary.

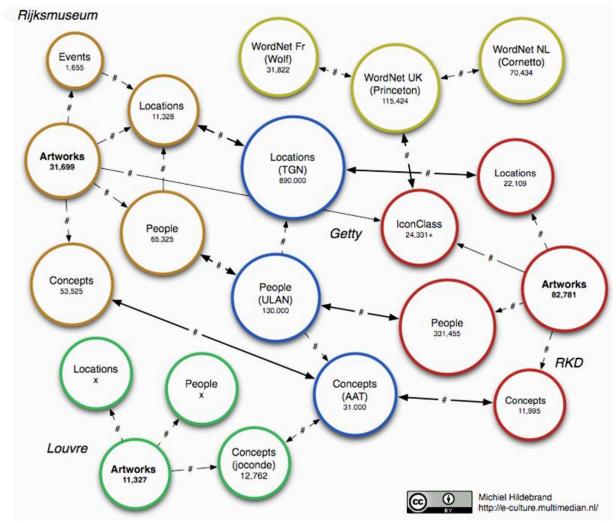


Fig. 3: This graph shows how collections of the Rijksmuseum, the Louvre and the Dutch Institute for Art History can be linked, through the use of controlled vocabularies and lists of names. Made by Michiel Hildebrand for the project 'The MultimediaN E-Culture demonstrator'.

Adding their identifiers to your terms or names does not mean that you need to change anything in your own lexicon. However, before you start the 'linking process', it is best to start cleaning up the keywords used for one of your metadata fields such as object type, or materials. OpenRefine (openrefine.org), a software sometimes coined "Excel-on-steroids" (ibid., 6), enables people to take care of both jobs. Downloadable without cost, it automatically finds possibly equivalent terms in your own terminology and, for instance, the AAT. A lot of helpful information is found on the website of linked data evangelists Seth van Hooland and Ruben Verborgh: freeyourmetadata.org, or in their publication Linked data for Libraries, Archives and Museums (Hooland and Verborgh, 2014). Should it be too technical, and no specialist assistance is near, one can consider copy-pasting the URIs into one's database by hand, as was done until fairly recently at the Rijksmuseum in Amsterdam (Hildebrand et al., 2009, 2).

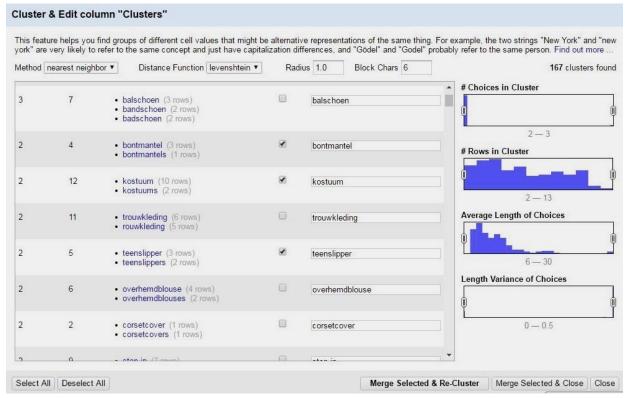


Fig. 4:
The clustering function in OpenRefine helps one to find similar terms inside a column in a set of data.

Once the data is published, it is easy to find what you are searching. A fashion researcher will for instance look for 'mitaine' – a word used in French and in Dutch for fingerless mitts – in a search box of a database. The search engine will then find the word 'mitaine' in the thesaurus concept corresponding to http://vocab.getty.edu/aat/300210013, and will return all items that carry the same URI, whichever object type (word) they've been labeled with.

### **Visual Thesaurus for Fashion and Costume**

To accelerate the process just outlined, the Fashion Museum of Antwerp is on a mission for ten museums in the Dutch speaking area in Western Europe. The heritage portal Modemuze originally inspired the project in 2015: it could use a shared vocabulary for better findability; this project is managed by Mila Ernst. The curators of the museums that launched this portal in April of 2015 had always been speaking of a general list of terms for them to use, but without the Linked Data dimension.

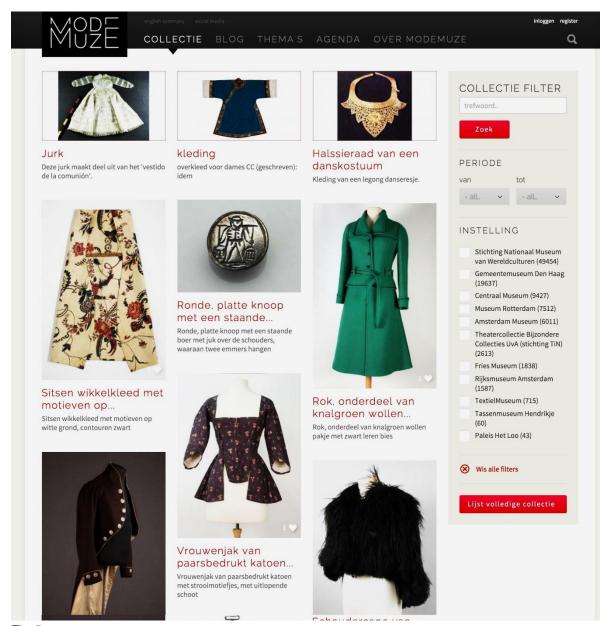


Fig. 5: The interface of Modemuze.nl.

A Dutch-Belgian collaboration was set up, based in Antwerp – its fashion museum, one of Europeana Fashion's initiators, could build on the experiences of the European project. Trying to limit unrealistic ambitions, the museum launched a pilot project to motivate the institutions to add URIs to the Object Type field of all their databases. Funding was found at the Flemish Government, for an international cultural heritage project (Nov 2015-Aug 2016).

# Our approach

What was needed to motivate the curators as well as the record-keepers to add identifiers to their metadata? First: make sure the shared vocabulary is "complete" – because it was not. Further in this paper you will find a description of this difficult task. Second: teach them why they should add links to their records. Therefore, in May 2016, a conference was held in Antwerp during which over 40 curators and catalogers found out about Linked Data and the practice of cleaning and linking. Third: show them that it actually works. We plan to realize this in 2017 through the integration of the thesaurus in the platform Modemuze and its presentation on the front end of the website, for the end user. Drawings made by fashion designer Dávid Ring will complete the visual aspect of the terminology baptized 'Visuele Thesaurus voor Mode en Kostuums'.



Fig. 6: The visuals Dávid Ring sketched for the Europeana Fashion thesaurus are being completed in the course of the project Visual Thesaurus for Fashion and Costume. They were published on Wikimedia Commons, free to use.

A core project group of five people meets every trimester and eight other enthusiasts are available for questions if needed – with backgrounds in user experience, IT, cataloging and thesauri management. The project is meant to function as a test for the Dutch-Belgian museum sector, as well as for the international heritage community. The Dutch Network for Digital Heritage, a governmental initiative, follows and supports the project, interested in several of its aspects. The next paragraph outlines the goals achieved so far.

# Creating a new terminology, based on the AAT

The beginning of the project mainly focused on juxtaposing the Art & Architecture Thesaurus (AAT) with the object names used by any of the ten museums mentioned. AAT's 1,599 terms for objects of fashion and costume (ID: 300209261) proved insufficient to describe the items in the fashion and costume collections concerned, but are extensive where armor or liturgical clothing are considered. 'Mary Janes' only come up as a type of cannabis instead of footwear, whereas 'applications' are solely defined as pieces of software waiting to be installed. Moreover, the scope notes and the translations seem to sometimes contain mistakes.

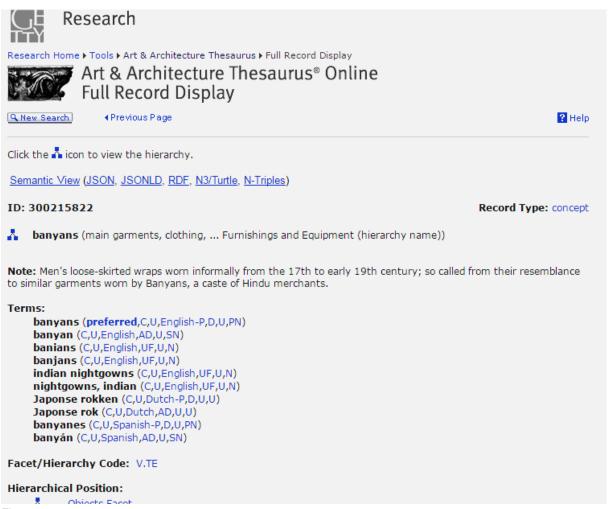


Fig. 7:

The concept 'Banyan' in the AAT.

However, the AAT was chosen as a foundation for our own thesaurus. It is increasingly being applied in Belgium and the Netherlands (three out of seven museums in the Modemuze consortium use it for their holdings). Furthermore, it is freely available as Linked Data, contains a clear hierarchy, is almost comprehensive and open to receive new terms. Given the fact that not the words but only the URIs need to be integrated into one's database, much of the worries, "to be obliged to copy AAT-terms" can be forgotten.

The analysis highlighted concepts currently absent in the AAT. For those missing concepts, new ones were created including their broader term, scope note and URI, in close collaboration with the curators. As is advised in the decision of the Advisory committee of the Dutch Network for Digital Heritage our thesaurus is published as one complete terminology in machine readable format - combining both AAT concepts and new ones. Adding new terms to the Californian AAT is a process, which can span a period of half a year for a large group of new concepts, whereas their acceptation is not guaranteed. Another part of the second phase of the project (Sept 2016-Aug 2017) corresponds to relaunching the activities for the fields of (textile) materials and techniques, planning to propose new terms for that facet of the AAT as well.

#### The process

At first, I (the author) collected overviews of both inventory numbers and the object names of ten fashion collections: in Belgium Modemuseum Hasselt, MoMu – Fashion Museum of Antwerp, Koninklijk Instituut voor Kunstpatrimonium (KIK, Brussels); in The Netherlands Amsterdam Museum, Rijksmuseum Amsterdam, Centraal Museum (Utrecht), Paleis Het Loo (Apeldoorn), Fries Museum (Leeuwarden), Gemeentemuseum Den Haag, Museum Rotterdam. Since the formats of these documents differed from one to another, they were all

manipulated in Excel to look alike. I divided the sheets into three columns: term, institution, number of objects tagged. Subsequently, I listed the 13,360 terms in one big spreadsheet, including the Dutch preferred and alternative labels of the AAT. Any part of the AAT can be retrieved from http://vocab.getty.edu/, the 'SPARQL endpoint' of the institution, using a 'JSON query'. I received help during this process through the Google group 'Getty Vocabularies as Linked Open Data'.

I decided to only select terms used for objects that "dress or decorate the body". So: no snuff boxes, stuffed animals or curtains, even though they showed up in many costume collections. Furthermore, 'combined terms', such as the many women's skirt, women's dress, women's shoe, women's fan and an enumeration in which you can replace 'women's' by 'mourning', 'evening' or 'summer' were removed. They should not be included in thesauri. In these cases, cataloguing, two terms should be used: a more general one and a specific one, i.e. 'women's clothing' and 'trousers' (Riesthuis 1992, 57-64). The terms also needed to be single values: no object types such as 'skirt and blouse' or 'pantyhose, with package'. Last were removed terms containing question marks, brackets or spelling mistakes.

In the long list of over remaining 9000 terms I wanted to find the terms that were already part of the AAT. Using the "Clustering" function in OpenRefine, I was able to automatically find the words that are comparable, such as 'corset' and 'korsetten' (AAT). Finally, once I found all the equivalents in the spreadsheet, and aligned them, I was able to tell which terms used by the museums appear in the AAT, by removing duplicates in my first column. 548 did, and 1034 did not – a number that seemed rather high.

A	A	В	С	D	E	G		J
						Japonnen bedoeld voor formele of		
						semiformele gelegenheden gedurende de		
54	avondjapon	avondjapon	- 7	Rijksmuseum	71	avond bijvoorbeeld theaterbezoek.	http://vocab.getty.edu/aat/300243843	300243843
						Japonnen bedoeld voor formele of		
						semiformele gelegenheden gedurende de		
55	avondjurk	avondjurk	3	AAT-Alt		avond bijvoorbeeld theaterbezoek.	http://vocab.getty.edu/aat/300243843	300243843
						Japonnen bedoeld voor formele of		
						semiformele gelegenheden gedurende de		
56	avondjurken	avondjurk	2	AAT-Alt		avond bijvoorbeeld theaterbezoek.	http://vocab.getty.edu/aat/300243843	30024384
						Japonnen bedoeld voor formele of		
						semiformele gelegenheden gedurende de		
157	avondjurken	avondjurk	2	Gemeentemuseum Den Haag	512	avond bijvoorbeeld theaterbezoek.	http://vocab.getty.edu/aat/300243843	30024384
						Japonnen bedoeld voor formele of		
			_			semiformele gelegenheden gedurende de		
58	avondjurk	avondjurk	3	Modemuseum Hasselt	257	avond bijvoorbeeld theaterbezoek.	http://vocab.getty.edu/aat/300243843	30024384
						Japonnen bedoeld voor formele of		
				B		semiformele gelegenheden gedurende de		
59	avondjurk	avondjurk		PaleisHetLoo		avond bijvoorbeeld theaterbezoek.	http://vocab.getty.edu/aat/300243843	30024384
<u> </u>	avondkleding	avondkleding		Amsterdam Museum	39			
81	avondkleding	avondkleding		Gemeentemuseum Den Haag	1446			
62	avondkleding	avondkleding		Modemuseum Hasselt	1			
63	avondkleding	avondkleding		Rijksmuseum	25			
64	avondkleding	avondkleding	_	RKD-AAT kandidaatterm				
65	avondmantels	avondmantels		Gemeentemuseum Den Haag	15			
66	avondmantel	avondmantels	1	Rijksmuseum	9			
			_			Nauwsluitend ondergoed gedragen ter		
37	b.h.'s	b.h.'s	2	AAT-Alt		ondersteuning van de buste.	http://vocab.getty.edu/aat/300210538	30021053
	l		_			Nauwsluitend ondergoed gedragen ter		
	b.h.'s	b.h.'s		Gemeentemuseum Den Haag		ondersteuning van de buste.	http://vocab.getty.edu/aat/300210538	30021053
39	babydoll	babydol		MoMu	1	Pyjama's bestaande uit een tot de heupen r		30021673
70		a babydoll (pyjam		AAT-Alt	_	Pyjama's bestaande uit een tot de heupen r		30021673
71	babydoll	babydoll (pyjam		Centraal Museum		Pyjama's bestaande uit een tot de heupen r		30021673
72		babydoll (pyjam:		Gemeentemuseum Den Haag		Pyjama's bestaande uit een tot de heupen r		30021673
73	babydoll	babydoll (pyjam-		Modemuseum Hasselt		Pyjama's bestaande uit een tot de heupen r		30021673
74	babydoll	babydoll (pyjam:	5	Museum Rotterdam	2	Pyjama's bestaande uit een tot de heupen r	http://vocab.getty.edu/aat/300216736	30021673
						Overkleding voor zuigelingen en baby's die		
						het lichaam in één gesloten zak omhullen al		
						dan niet met capuchon en mouwen;		
75	babyslaapzak	babyslaapzak	2	AAT-Alt		ééndelig of met sluitingen in de taille.	http://vocab.getty.edu/aat/300210598	300210598
						Overkleding voor zuigelingen en baby's die		

Fig. 8: Thesaurus worksheet; the selection shows terms for which the AAT is lacking a URI. The first number says how many museums use the term out of the ten, the second how many items were described with this term inside a collection.

### Address the curators' knowledge

The overview of the 1034 terms the AAT seemed to be missing was used during three editing sessions in the late spring and summer of 2016, with the curators of our museums (and interested fashion specialists). The project was kindly being supported with the knowledge of Gieneke Arnolli, Frieda Sorber, Sjouk Hoitsma, Wim Mertens, Jacoba de Jonge, Karolien de Clippel, Judith van Amelsvoort, Marit Eisses, Birthe Weijkamp, Mila Ernst, Jacco Hooikammer, Hanneke van Zuthem, Leonie Sterenborg, Anne-Karlijn van Kesteren, Lisa Whittle, Inge Specht, Ninke Bloemberg and Trudie Rosa de Carvalho. The questions

asked during these sessions were: "can you find it in the AAT anyway?", "do we need this concept?" and "what is it"? Throughout these editorial meetings, the list of 1,038 items, comprised of terms used in their own institutions, was dealt with. It was saved as a Google Doc. Based on their replies, I was able to decide whether a new concept should be created for the term, a new alternative label, or if the term should be discarded.



Fig. 9: No concept in the AAT referred to the 'kniptassen' (bags closed by a clasp), needed for describing some of these 19<sup>th</sup>-century bead knitted items. Collection of the Fries Museum in Leeuwarden. Photo: Erik en Petra Hesmerg.

After drafting a selective list of needed new concepts (150) and new alternative labels (45), these were appointed to each of the curators and myself. As homework, everyone needed to find three sources per new concept, write scope notes, and include a source for those too. Such are the editing instructions of the Dutch AAT office, through which we will propose our new concepts. Over a hundred-fifty books were consulted to find the terms or their description, as well as the Vocabulary of Basic Terms for Cataloguing Costume, which was set up by the ICOM Costume Committee.

(See: collectionstrust.org.uk/assets/thesaurus\_icombts/vbt00e.htm.)

Concepts have a plural preferred term, to be in line with the ISO standard for thesauri (ISO2788:1986). The process of collecting and comparing terminologies covered 1.5 months of full time work. Preparing the new terms aided by the curators, and keeping track of changes to be made in the AAT export, covered another 2 months. At the moment this article is written, the AAT has been imported into the Austrian thesaurus software called Poolparty. Our adaptations and additions are being incorporated, to prepare the approaching publication of the thesaurus as Linked Data and its integration in the Modemuze portal.



Fig. 10: A presentation in May 2016 introducing Linked Data to Dutch and Belgian curators and record-keepers. Photo: Monica Ho/MoMu, Fashion Museum of Antwerp

# A few complications

Some challenges surfaced during our project, both technical and terminology-related; few will be addressed here. One of the main goals was to find out whether a term already existed in the AAT. This process was helped by OpenRefine, but still needed a lot of manual double checking on the AAT website. The upside: one gets to know one's terminology. It helps if there is technical back-up available or if the one in charge is at ease with manipulating data in an Excel spreadsheet (sort, distribute text to columns, import a CSV format). Collaboration between a record-keeper and a curator is recommended.

When creating or adapting a vocabulary, other questions rise, regarding content. For instance, certain words mean something totally different in another language. In the AAT, the Dutch 'spencer' or 'slipover' (a pullover without sleeves), is now a synonym for the English and French 'spencer': an early 19<sup>th</sup> century high waisted women's jacket. A sweater vest actually does not yet exist in the terminology: the project will introduce a new concept, and will have the Dutch 'spencer' removed from the concept mentioned. Apart from having technical help on hand to work with the various programs and software, having some assistance in researching glossaries and secondary literature is not an excessive luxury.

#### Conclusion

At present, databases are filled with a textual mount of information. Its content can be queried using words: combinations of letters. The results of these queries can be relatively good, but do not compare to the results of a query that is supported by Linked Data. This article addressed the importance of using URIs to tag terms inside databases. Hopefully it has put forward the benefits of linking metadata to shared controlled vocabularies for institutions collecting fashion heritage. The goal of the article was also to provide the reader with a sense of the organizational and technical aspects of how this process can be tackled. The experiences of the project Visual Thesaurus for Fashion & Costume, ran by MoMu, the Fashion Museum of Antwerp, served as examples.

The terminology-related example challenge of this project addressed in the previous subsection is the tip of the iceberg, which could be discovered further and possibly dismantled on an international level. The author and community working on the initiative in The Netherlands and Belgium are happy to share the experiences so far and remain at the disposal of anyone interested in the subject.

#### References

Bloemberg, Ninke, Mila Ernst, and Sjouk Hoitsma. 2015. "Museum Fashion and Costume Collections Online: Two Collaborations." In *Proceedings of the 2014 Annual Meeting of the ICOM Costume Committee*. Nafplion: PFF, 45-50.

Hildebrand, Michiel, Ossenbruggen, Jacco van, Hardman, Lynda, Jacobs, Geertje. 2009. "Supporting subject matter annotation using heterogeneous thesauri. A user study in Web data reuse." *International Journal of Human Computer Studies* 67, 888-903.

Harpring, Patricia. 2010. *Introduction to Controlled Vocabularies. Terminology for Art, Architecture and Other Cultural Works.* Los Angeles: Getty.

Hooland, Seth van, and Ruben Verborgh. 2014. *Linked Data for Libraries, Archives and Museums. How to clean, link and publish your metadata.* London: Facet.

Hooland, Seth van, Ruben Verborgh, Max de Wilde, Johannes Hercher, Erik Mannens and Rik van de Walle. 2013. "Evaluating the success of vocabulary reconciliation for cultural heritage collections." In *Journal of the American Society for Information Science and Technology* 64, 464-479.

Kirkland, Arden, Kathi Martin, Marlise Schoeny, Kiki Smithand Gayle Strege. 2015. "Sharing Historic Costume Collections Online." In *Dress* 41, 107-127.

Riesthuis, Gerrit and Gerard Wijnands (eds.). 1992. *Thesaurusbouw: handboek voor opleiding en praktijk.* The Hague: Nederlands Bibliotheek en Lektuur Centrum.