

Golden Agents. A web of linked biographical data for the Dutch Golden Age

Judith Brouwer & Harm Nijboer

Huygens ING, Oudezijds Achterburgwal 185, 1012 DK Amsterdam, The Netherlands

Email: judith.brouwer@huygens.knaw.nl, harm.nijboer@huygens.knaw.nl

Abstract

Being home to famous painters, prolific printmakers, and internationally operating printers and publishers seventeenth century Amsterdam was a 'creative city' *avant la lettre*. As such it is a topic of continuous interest to a broad range of scholars. The production of art, books, literature and other creative products in this period is covered by many electronic resources like collection databases of museums and libraries, dedicated documentation systems, and research databases. Often biographical records are at the heart of these systems, but typically all these resources use their own subset of biographical data. In the Golden Agents program we will connect these resources in a linked data framework. This will result in a sustainable infrastructure to study relations and interactions between producers and consumers of creative goods in the Dutch Golden Age. In this paper we will discuss our strategies and experiences in connecting various sets of biographical data. Finally we will highlight the research potential of the aggregated data by the case of the Dutch poet, writer and diplomat Constantijn Huygens (1596-1687).

Keywords: Linked Open Data, Dutch Golden Age, creative industries, biographical data, prosopography, network analysis.

1. Introduction

By the mid seventeenth century Amsterdam was the cultural hotspot of the Western world. The city was not only home to Rembrandt, the most innovative painter of his time, and many other painters that did or did not belong to his circle, it also housed a large number of publishing houses, printmakers, writers and thinkers. The Blaeu and Janssonius publishing houses issued beautiful maps of almost every corner of the then known world (Van Netten, 2014). In a different – and often more hidden – segment of the market there was an explosion of publications by Cartesian, Spinozist and otherwise heterodox thinkers, that paved the way for the rise of Enlightenment philosophy (Israel, 2001). Meanwhile the Amsterdam public theatre and its associated translating industry evolved into a hub for the diffusion of 'modern' Spanish and French drama across northwestern Europe (Blom, et al. 2016).

The success of Amsterdam as a creative city was clearly not the result of a single art or trade. Nevertheless there is a long tradition of researching painting, literature, book production and other arts and crafts as more or less autonomous phenomena, as if they happened in real life perfectly within the borders of art history, the history of literature, book history or another specialist field of study. Recently, Rasterhoff (2017) called for a more integrative approach. In her innovative study on painters and publishers in the Dutch Republic she investigated the two trades in comparison by

making intensive use of existing datasets. Rasterhoff's book does not only clearly demonstrate the potential of cross-disciplinary data driven research on the production of goods like art, books, drama and literature, it also shows that for more in-depth analysis there is an urgent need to connect data sets and to advance their interoperability.

In the Golden Agents program we will address this need by creating a Linked Open Data framework that will allow for combined access to various distributed resources concerning the production and consumption of creative goods in the Dutch Golden Age. Identifying persons and the harmonization of biographical data is of course a key factor in achieving this objective.

2. A web of biographical data

The Golden Agents program will proceed in a now well established tradition of using Semantic Web techniques and Linked Data principles to connect distributed data sets in the fields of cultural heritage (cf. Hyvonen 2012; Jones & Seikel 2016) and historical research data (e.g. De Boer et al. 2014). These techniques are built upon the idea that http web addresses (URIs) can be used to denote things, persons, properties and relations under the (ideal) condition that those URIs refer to an actual place on the Web where a formal description can be found of the denotatum. We can use these URIs to make subject-predicate-object statements following the RDF syntax. We can for instance state that the

Rembrandt described in the ECARTICO database is the same as the Rembrandt in Wikidata by using the following triple of URIs:

```
<http://www.vondel.humanities.uva.nl/ecartico/persons/6292>
<http://www.w3.org/TR/2004/REC-owl-semantic-20040210/#owl_sameAs>
<http://www.wikidata.org/entity/Q5598>
```

The predicate in this example refers to a formal definition of ‘being the same as’ in the OWL Web Ontology Language. Although there is a lot of discussion and confusion about the precise semantics of such ‘same as’ statements (cf. Halpin et al., 2010), it provides a straightforward and easy to implement schema to harmonize lots of entities across different data sets. And this is already done at a large scale, also with regard to resources that are of interest to the study of the Dutch Golden Age. This can be illustrated by looking at an important resource for studying Dutch culture in the Golden Age: the Short Title Catalogue Netherlands (STCN).¹

The STCN is a retrospective bibliography of books printed in the Netherlands prior to 1800, which is hosted by the National Library of the Netherlands. Since the STCN uses the same authority files for author names as the National Library, the authority files of the STCN are automatically clustered by the Virtual International Authority File (VIAF)². And in the RDF version of the STCN, launched in 2015 (cf. Boot, et al. 2016), every local authority file is linked to a VIAF URI using an ‘owl:sameAs’ predicate.

In the same manner resources like Wikidata³ and ECARTICO⁴ link their representations of people to VIAF identifiers as well as to other identifiers like those provided by RKDartists⁵, Getty’s Union List of Artist Names (ULAN)⁶, the Dutch Biography Portal⁷ and the Digital Library of Dutch Literature (DBNL)⁸. Figure 1 gives an overview of the current state of affairs. Notwithstanding that many connections still have to be made, it clearly demonstrates how interconnected these resources already are.

1 <http://picarta.nl/xslt/DB=3.11/>
2 <https://viaf.org/>
3 <https://www.wikidata.org/>
4 <http://www.vondel.humanities.uva.nl/ecartico/>
5 <https://rkd.nl/nl/explore/artists>
6 <http://www.getty.edu/research/tools/vocabularies/ulan/>
7 <http://www.biografischportaal.nl/>
8 <http://www.dbnl.org/>

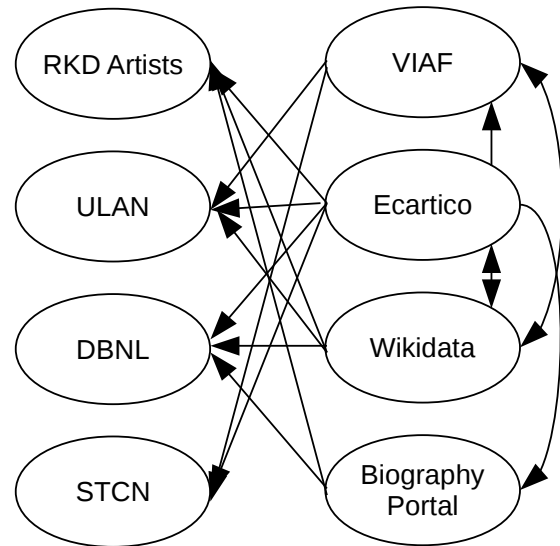


Figure 1. Linkage between various sets of biographical data (current state of affairs)

3. Adding prosopographical depth with ECARTICO

One should be aware that many of the resources mentioned so far are basically indexes. Person entities described by VIAF and the DBNL are only provided because there are texts written or published by or about these people. As a consequence the basic objective of these representations of people is to function as key words by which objects and documents can be searched and retrieved. *Mutatis mutandis* the same is true for the lists of artists published by the RKD (Netherlands Institute for Art History) and the Getty Research Institute – although they tend to provide more biographical depth – and the Dutch Biography Portal which actually is an index to a series of biographical dictionaries.

The status of Wikidata in this respect is somewhat ambiguous. On the one hand it serves as an index to various document-oriented resources of the Wikimedia family (e.g. Wikipedia, Wikimedia Commons and Wikisource); on the other hand it has the ambition of being a global knowledge base. But the global perspective will not always comply with researchers’ requirements to data. Being a community driven project, without an editorial board (and editorial directives), it has inevitably issues with regard to consistency in data modeling. Voß (2016) pointed out that the strategy of Wikidata is to store statements rather than facts and as a result of this strategy contradicting statements are not forbidden – on purpose. Moreover, community

directives are rather vague in defining which entities should and which entities should not be covered by Wikidata, except that an entity should meet the criterion of 'notability'.⁹ Although the Wikidata community tends to be a bit more lenient in this respect than its Wikipedia counterpart, the criterion does exist and ultimately it does pose a limit on what can be included in Wikidata (cf. Nielsen, et al. 2017; Roued-Cunliffe 2017).

There is a general tendency of biographical resources to be biased towards famous people (cf. Fokkens et al., 2017). Famous painters like Rembrandt and Vermeer are of course well documented, but this is definitely not the case for lesser painters like Casparus Hoomis, Jacques Luls, Jan Jeuriaenz Swijnscoep, and many others who supplied the market with an abundance of relatively cheap paintings (Jager 2016). These 'daubers' might raise little interest among art historians who exclusively focus on high art, but it is evident that identifying these lesser painters is a necessary step in creating a view on the market for paintings as whole.

Mapping the market for creative goods, both high and low, is the primary objective of ECARTICO, a comprehensive collection of structured biographical data concerning painters, engravers, printers, booksellers, gold- and silversmiths and others involved in the 'creative industries' of the Low Countries from circa 1475 to circa 1725. The data is (mostly manually) taken from both secondary (literature) and primary (archival) sources and is updated on an almost daily basis. The data is editorially reviewed by the authors of this paper and Marten Jan Bok (University of Amsterdam).

Unlike the resources mentioned before ECARTICO has the explicit objective to get a view on these industries as a whole. This means that ECARTICO is inclusive with regard to lesser artists and practitioners of lesser arts (e.g. silversmiths and illuminators of maps and prints). ECARTICO does also include many creative entrepreneurs whose existence is only testified by archival sources and who have not been described in the existing literature, yet. Furthermore, it does provide an extensive mapping of direct and indirect relations and interactions between these people. As a consequence, ECARTICO does not only provide representations of people (and URIs) for those actively involved in creative industries, but also for their direct relatives (parents, spouses, children), customers and other relevant contacts.

ECARTICO is available online since december 2011 under a liberal license. At the moment of writing ECARTICO contains structural biographical data on over 36 000 persons



Figure 2. Adriaen Hanneman, Portrait of Constantijn Huygens (1596-1687) with his five children (1640). Oil on canvas. Mauritshuis, The Hague

with over 21 000 relations defined between these persons. So far, nearly 5 000 persons have been connected to external resources by over 16 800 (and counting) links. As such it constitutes a prosopographical layer that can be used in combination with resources like the STCN, RKDartists&, Wikidata, and others to study the social history of creative industries in relation to its output.

4. Aggregating data

ECARTICO and (subsets of) other resources mentioned above will be aggregated in a dedicated Golden Agents instance of the Timbuctoo Linked Data store.¹⁰ In the run of the Golden Agents program these data will be supplemented by archival data, data from museum collections and data sets by individual researchers. For reasons of sustainability, stability and scalability we have chosen to aggregate data instead of relying on the availability of multiple endpoints. A first stack of the data aggregation is expected to be available by the end of 2018.

9 <https://www.wikidata.org/wiki/Wikidata:Notability>

10 <https://timbuctoo.huygens.knaw.nl/>

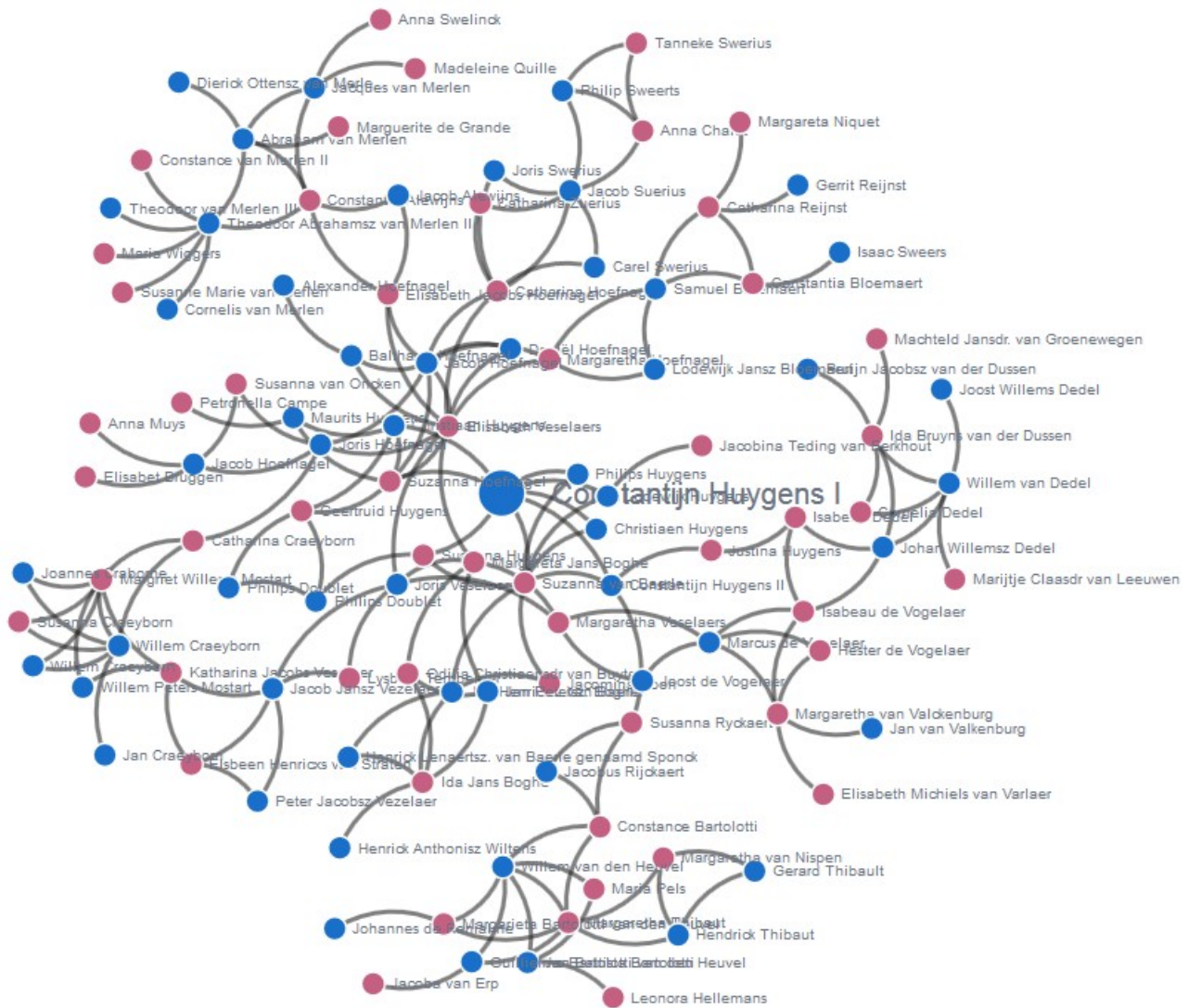


Figure 3. Six degree kinship network of Constantijn Huygens. Source: ECARTICO.

The research potential of this aggregate of biographical, archival and object related resources can be illustrated by looking at the networks around the Dutch poet, writer, diplomat and patron of the arts Constantijn Huygens (1596-1687).

5. Use case: the six degrees of Huygens

Looking at the painting (Figure 2) by Adriaen Hanneman (c. 1604-1671), we see Constantijn Huygens as pater familias amidst his five children (clockwise) Suzanne, Constantijn, Lodewijk, Philips and Christiaan (the internationally better known physicist Huygens). Huygens' wife Suzanne van Baerle (1599-1637) is missing; she died a few years earlier, shortly after giving birth to their youngest daughter. The

painting represents only a very tiny part of Huygens' network. It is a well known fact that the interconnected group of people around Huygens was considerable. This is demonstrated in itself by the circa 72000 outgoing letters of his correspondence (Huygens, 2003).

Being one of the most versatile persons of his time, it is not surprising that Huygens' network is just as universal. In his autobiographical *Mijn leven verteld aan mijn kinderen* ('My life as told to my children', 1678) Huygens entitles princes, courtiers, scientists and other prominent persons as his friends (Huygens, 2003; Liedtke, 2001). In early modern times, friendship was – far more than today – a strategic relationship, carefully nourished by mutual favours and it

was closely associated with (distant) kinship (cf. Kooijmans, 2016). Various important links in Huygens' network of kin and friends were established via his mother Susanna Hoefnagel (1561-1633).¹¹ He was well aware of the significance of her ancestry. In *Mijn jeugd* Huygens (2008) describes her family as 'a family of muses'.

With ECARTICO we can test whether this statement by Huygens holds any truth, since every person's page in ECARTICO gives users access to a tool to generate and visualize egocentric networks around the person in question. This tool uses a native social networks API and the widely used D3.js¹² visualization library. The networks API sets default to kinship networks (marriages and parent-child relations), but parameters can be adjusted to include any other relational type (e.g. master-pupil relations) that is modeled in the data set.

Figure 3 shows the six degree kinship network of Huygens based on the data modeled in ECARTICO.¹³ It demonstrates that Huygens was indeed closely related to the Hoefnagel and Van Merlen families of artists. And when we would expand this network with another two degrees even more artists (e.g. Gerbrand van den Eeckhout and Pieter Isaacs) and writers (e.g. Pieter Cornelisz Hooft and Jacob Cats) would appear.

Six and eight degree kinship networks result in more or less intelligible figures. If we would extend the six degree network around Huygens to all relations modeled in ECARTICO we would end up with a 'hairball' of 5 415 persons and 7 602 relations. Of these persons 1 790 persons are linked to external resources with in total 7 618 external links. And these links in their turn provide access to thousands of descriptions of books, writings and objects of art.

6. Final remarks

The aggregation of resources in the Golden Agents framework will result in complex graphs of people and creative works as the example of Constantijn Huygens shows. These graphs might not provide direct answers to researchers. But since the data behind these graphs can be processed by computers, researchers can call on assistive technology to search for meaningful patterns and relations. In the Golden Agents program a multi-agent platform will be developed to assist researchers in querying the data. But given the current advances in fields like artificial

intelligence and computer vision we might already think of a whole range of other assistive technologies that can be plugged in the Golden Agents framework in time.

7. References

7.1 Resources

- Digital Library of Dutch Literature (DBNL)
<http://www.dbnl.org/>
- Dutch Biography Portal
<http://www.biografischportaal.nl/>
- ECARTICO
<http://www.vondel.humanities.uva.nl/ecartico/>
- Getty Union List of Artist Names (ULAN)
<http://www.getty.edu/research/tools/vocabularies/ulan/>
- RKDartists&
<https://rkd.nl/nl/explore/artists>
- Short-Title Catalogue, Netherlands (STCN)
<http://picarta.nl/xslt/DB=3.11/>
- VIAF
<https://viaf.org/>
- Wikidata
<https://www.wikidata.org/>

7.2 Literature

- Blom, F. R. E., Jautze, K., Alvarez Frances, L. (2016) Spaans theater in de Amsterdamse Schouwburg (1638-1672). Kwantitatieve en kwalitatieve analyse van de creatieve industrie van het vertalen. *De Zeventiende Eeuw*, 32 (1), pp.12-39.
- Boot, P., Dietz, F., Stronks, E., & Zwart, W. (2016). Young Agents. Jonge auteurs op de vroegmoderne boekenmarkt. *Nederlandse Letterkunde*, 21, pp. 1-34.
- De Boer V., Van Rossum M., Leinenga J., Hoekstra R. (2014). Dutch Ships and Sailors Linked Data. In: Mika P. et al. (eds) *The Semantic Web – ISWC 2014. ISWC 2014. Lecture Notes in Computer Science*, vol 8796. Cham: Springer, pp. 229-245.
- Fokkens, A., Ter Braake, S., Ockeloen, N., Vossen, P., Legêne, S., Guus Schreiber, G., De Boer, V. (2017). BiographyNet: Extracting Relations Between People and Events. In: Á. Z. Bernád, C. Gruber, M. Kaiser eds., *Europa baut auf Biographien: Aspekte, Bausteine, Normen und Standards für eine europäische Biographik*. Wien: New Academic Press, pp. 193-224.
- Halpin H., Hayes P.J., McCusker J.P., McGuinness D.L., Thompson H.S. (2010) When owl:sameAs Isn't the Same: An Analysis of Identity in Linked Data. In: Patel-Schneider P.F. et al. (Eds.) *The Semantic Web – ISWC 2010. ISWC 2010. Lecture Notes in Computer Science*, vol 6496. Berlin/Heidelberg: Springer.
- Huygens, C. (1898), *De gedichten van Constantijn Huygens*. 9 vols., vol. 8. Ed. J.A. Worp. Groningen: Wolters.

11 <http://www.vondel.humanities.uva.nl/ecartico/persons/10002>

12 <https://d3js.org/>

13 <http://www.vondel.humanities.uva.nl/ecartico/networks/index.php?ego=3958&level=6>

- Huygens, C. (2003). *Mijn leven verteld aan mijn kinderen in twee boeken*. Ed. F.R.E. Blom. Amsterdam: Prometheus/Bert Bakker.
- Huygens, C. (2008). *Mijn jeugd*. Ed. C. Heesakkers. Leiden: DBNL.
- Hyvönen, E. (2012). Publishing and Using Cultural Heritage Linked Data on the Semantic Web. *Synthesis Lectures on the Semantic Web: Theory and Technology 2* (1), pp. 1-159.
- Israel, J. (2001). *Radical enlightenment: philosophy and the making of modernity 1650-1750*. Oxford [etc.]: Oxford University Press.
- Jager, A. (2016). *'Galey-schilders' en 'dosijnwerck'. De productie, distributie en consumptie van goedkope historiestukken in zeventiende-eeuws Amsterdam*. PhD thesis University of Amsterdam.
- Jones, E., Seikel, M. (2016). *Linked data for cultural heritage*. Chicago: ALA Editions.
- Kooijmans, L. (2016). *Vriendschap en de kunst van het overleven in de zeventiende en achttiende eeuw*. 2nd. rev. Ed. Amsterdam: Uitgeverij Bert Bakker.
- Liedtke, W.A. (2001). *Vermeer and the Delft School*. New York: Metropolitan Museum of Art, New Haven [etc.]: Yale University press.
- Netten, D. van. (2014). *Koopman in kennis: de uitgever Willem Jansz Blaeu in de geleerde wereld (1571-1638)*. Zutphen: Walburg Pers.
- Nielsen F.Å., Mietchen D., Willighagen E. (2017) Scholia, Scientometrics and Wikidata. In: Blomqvist E., Hose K., Paulheim H., Ławrynowicz A., Ciravegna F., Hartig O. (eds) *The Semantic Web: ESWC 2017 Satellite Events. ESWC 2017. Lecture Notes in Computer Science*, vol 10577. Cham: Springer, pp. 237-259.
- Rasterhoff, C. (2017). *The fabric of creativity in the Dutch Republic. Painting and publishing as cultural industries, 1580-1800*. Amsterdam: Amsterdam University Press, Chicago: University of Chicago Press (PhD thesis Utrecht University 2012).
- Roued-Cunliffe, H. (2017) Visualising Historical Networks: Family Trees and Wikipedia. *Akademisk Kvarter / Academic Quarter* 15, pp. 40-53.
- Voß, J. (2016). Classification of Knowledge Organization Systems with Wikidata. In: Mayr, P., Douglas Tudhope, D., Golub, K., Wartena, C., De Luca, E. W. eds., *Proceedings of the 15th European Networked Knowledge Organization Systems Workshop (NKOS 2016), co-located with the 20th International Conference on Theory and Practice of Digital Libraries 2016 (TPDL 2016), Hannover, Germany, September 9, 2016*, pp. 15-22.
- is a collaboration between Huygens ING, the Meertens Institute, the University of Amsterdam, Utrecht University, VU University Amsterdam, the Rijksmuseum, KB National Library of the Netherlands, the Amsterdam City Archives, the RKD Netherlands Institute for Art History and Lab1100.

8. Acknowledgments

The Golden Agents program is sponsored by The Netherlands Organisation for Scientific Research NWO and