Alex Mayhew
University of Western Ontario, London, ON, Canada

Phylomemetics as a Framework for Bibliographic Synthesis (Paper)

Abstract:
The description of texts based on their title, author, publisher, and with appeals to authority records, has been a cornerstone of bibliographic description as practiced in library catalogues. There are also many parallel or alternative descriptive projects from other fields. Examples of these include fairy tale motifs, literature tropes, character archetypes, story patterns, subject headings, semantic features, citations, allusions, musical influence, theater family trees, and more. Each of these projects represent understandings of texts that are valued by people and could enhance the library catalogue. What is needed is a framework for synthesizing a diversity of descriptions into the catalogue.

1.   Introduction

Bibliographic description in library cataloguing has settled on a selection of textual features for the creation of bibliographic records. These textual features are described in rigorous detail in systems like the Anglo-American Cataloging Rules Second Edition (AACR2) and Resource Description and Access (RDA) and include elements such as: Title, Statement of Responsibility, Edition Statement, Numbering of Serials, Publication Statement, Copyright Date, and unique identifiers like ISBNs (RDA, 2019). Due to works existing in many languages as well as many alternate spellings of names, traditional bibliographic description also relies on authorized versions of elements like titles and names to collect together texts (Authorities, 2019).

Beyond traditional bibliographic description for library catalogues there are many diverse projects aiming to describe texts focusing on. Two examples are citations and descriptions of recurring motifs in fairy tales. These projects are pursued because people value many diverse descriptions of texts. Some of these projects distinguish more granular descriptions of texts, breaking the text down into many discrete components. Other projects focus on creating connections between texts, situating each work in a broader context.

The library catalogue could serve as the nexus where many of these diverse descriptions are brought together and are made available to all catalogue users. Simply adding the information from many projects may allow interesting results to emerge, but we may be able to provide a scaffold or framework to allow these different descriptions to do more than simply sit side-by-side.

2.   Phylomemetics

There are many parallel bibliographic description projects, a selection of which will be discussed below. Incorporating any one of them into the library catalogue, for example tagging relevant works with fairytale motifs or adding citation links, would be a significant endeavor.

Noticing the commonalities of these various parallel bibliographic description projects may allow us to develop standards that would ease the integration of them into the catalogue, if desired.

Without delving into the particulars of any specific project just yet, we can notice that there are at least two types of projects: Ones that create more granular descriptions and ones that create connections between texts.

We can make an analogy with evolutionary biology to see how these description types are actually complementary. Biology may seem an odd choice as a source of inspiration for cataloguing, however, the classification of species (taxonomy) in particular has a number of analogous concerns to cataloguing. Cataloguing is concerned with naming and organizing texts, while taxonomy is the "scientific discipline concerned with naming and classifying the diverse forms of life" (Urry et al, 2017).

The modern conception of evolution is composed of two important elements: genes and descent with modification. Credit for these ideas is usually given to Gregor Mendel and Charles Darwin respectively. This modern version of the theory of evolution is called the Neo-Darwinian Synthesis, and its study is called phylogenetics. Phylogenetics reveals that the change in the composition of genes is the modification that happens during descent.

The question the theory of evolution seeks to answer is: What explains the diversity of life? We can ask an analogous question: How can we characterize the diversity of texts?

Historically, two schools developed to explain the diversity of life. Mendel deduced not only the existence of discrete hereditary units (genes) but also that the units were present in pairs in the pea plant, as well as other organisms (Britannica, 2022). We can identify analogous discrete units in texts such as motifs, themes, archetypes, patterns, tropes, and clichés, or more generally: memes. Like genes, memes can exist as many different levels of abstraction, sometimes appearing in a single sentence, or sometimes spread out across an entire book. Through mutation or chance genes or memes could end up appearing in any organism or text.

Darwinians noted that empirical evidence indicated that variation was continuous in most organisms, not discrete as Mendelism seemed to predict, species change fluidly with time, sometimes called adaptive radiation (Britannica, 2022). Individuals of a species would be expected to be some variation of their parents, with no discrete jumps. By analogy texts could be thought of as emerging from a particular cultural milieu or be situated in a particular genre or academic tradition. In this way texts can be seen as carrying the imprint of what came before, but also bound to that legacy, though not permanently, and not in a clearly definable way.

Like biological species, texts can be understood as both being made up of decomposable components in a Mendealian fashion, and being the product of influences inherited from previous texts in a Darwinian fashion. We can understand most parallel bibliographic description projects as either Mendelian or Darwinian. This suggests the possibility of a synthesis of bibliographic descriptions and the library catalogue, as potentially the most complete repository of texts, is well suited to being the site of such an undertaking. By analogy to phylogenetics the study of texts from this perspective has been named phylomemetics (Howe & Windram, 2011).

3.   Mendelian Projects

The idea of applying biological models to cultural phenomena has gathered momentum following Dawkins's articulation of the meme in *The Selfish Gene*. Susan Blackmore re-stated the definition of memes as: "whatever is copied from one person to another person, whether habits, skills, songs, stories, or any other kind of information. Memes, like genes, are replicators in the sense as defined by Dawkins." In other words, if genes are the units of biological inheritance, memes are the units of cultural inheritance. Memes can exist at various levels of abstraction and can be difficult to pick out. The goal of Mendelian description projects is identifying various memes.

Brief descriptions of a number of Mendelian bibliographic description projects follows:

1. Fairy Tale Motifs

The Aarne–Thompson–Uther Index (ATU Index) is a catalogue of folktale types used in folklore studies (Uther, 2004). The ATU Index describes hundreds of story patterns. Vladimir Propp's *Morphology of the Folktale* is similar, but focused on Russian folklore (Propp, 1968).

2. Literary Criticism

There is a long history of describing and analyzing the components of stories. Character archetypes or stock characters abound, such as 'The Gladiator' (Steenberg, 2020) or 'The Crone' (Becvar, 2005). There is also work done on the types of plot structure, from Joseph Campbell's monomyth (Campbell, 1949) to Christopher Booker's *Seven Basic Plots* (Booker, 2004). Finally, there is the identification of symbolism, such as the apple as a symbol of discord in Greek and Roman literature (Littlewood, 1968) and the white horse as a symbol of life and death (Despenser, 2004).

3. Science Fiction and Fantasy Conceits

Like literary criticism generally, there is considerable effort to classify the conceits of various genres, science fiction and fantasy has been particularly fruitful. Clare Winger Harris's 1931 list of science fiction story types was foundational (Jones, 2020). The 1952 *Fantasy Classification System* by Alastair Cameron "does a credible job of providing a classification scheme for all of science fiction" (Bray, 2020). The *Historical Dictionary of Science Fiction* is an offshoot of a project begun by the Oxford English Dictionary, now edited by Jesse Sheidlower, lists the definitions and origins of many terms (Sheidlower. 2022).

4. TvTropes

"TV Tropes is a wiki documenting, in a fairly informal manner, the various conventions of fiction" (TvTropes, 2021). It collects and describes tropes and links them to works across television, film, and literature. Tropes "can be a plot trick, a setup, a narrative structure, a character type, a linguistic idiom" (TvTropes, 2021). The tropes available include many of the types and examples seen above.

5. Structured Scientific Concepts

The process of science itself has also seen its various elements isolated and described. Structured models of scientific concepts are used as a basis for organizing, accessing, and using learning materials (Smith et al, 2002). There are efforts to quantify the rise and fall of scientific

fields (Singh, et al, 2021). Finally, an ontology of research methods and experiments has been created (Soldatova & King, 2006).

6. Music Basic Concepts

In the domain of music Szostak and Smiraglia are creating the Basic Concepts Classification (Szostak & Smiraglia, 2019). There are also efforts to use social tagging for music information retrieval (Lamere, 2008). Finally, there is work being done to describe how musical styles propagate through time (Jan, 1999).

7. Subject Headings

It is also valuable to mention the Library of Congress Subject Headings (LCSH) which are a collection of terms for describing the content of a wide variety of texts. Each term sits in relation to other terms, being either broader, narrower, or a more amorphous related term. Usually a relatively small number of terms is assigned to each text representing its dominant memetic content. The Library of Congress also maintains the Subject Authority Headings service, which includes the LCSH as well as terms for genre and form for library and archival materials, as well as Children's Subject Headings (Library of Congress, 2015).

4.   Darwinian Projects

Noticing the influence of one text on another is as old as the first re-telling of a story. Allusions, quotations, and references have been enriching works for longer than is worth mentioning. Many projects exist that attempt to map these lines of influence between texts, sometimes going as far as to build family trees. We can label these as Darwinian description projects.

1. Citations

The classic example of a relationship of influence in scholarship is the citation. It is not always clear what exactly is being inherited, just that there is a connection between the texts. Citation analysis has been used as the basis for discovering lines of descent propagating through multiple generations of articles (Asubiaro, 2021).

2. Literary Criticism

Beyond the allusions, quotations and references mentioned above, literary criticism has a variety of subfields. One looks for implied missing documents like the hypothesized Q source for the Bible (Scholer, 1990). There is also stemmatics, which creates visualizations of the connections between texts. For example, Maas created a computer generated relationship diagram of fifty-two versions of Carakasaṃhitā Vimānasthāna, a diagram that bares a noticeable resemblance to a family tree (Maas, 2009). There are also semantic web projects, such as linking the works related to Jane Austen (Jane-athon, 2015).

3. Genealogy of Plays

Theater is full of works taking inspiration from those that have gone before. Sometimes this connection is due to being an explicit and intentional retelling, like *West Side Story* being an adaptation of *Romeo and Julliet*. Less clear connections are the subject of critical inquiry. For example, Williams maps out the lines of influence of *A Midsummer Night's Dream* (Williams, 1997).

4. Scientific Topic Evolution

While video game tech trees may give the impression that the connections between various scientific insights and fields is well understood (Ghys, 2012), mapping these connections is very much an active field of investigation. While Asimov's Chronology of Science and Discovery was formative, it simply lists events deemed important chronologically (Asimov, 1989). The question of scientific topic evolution has begun to be addressed more directly (Li et al, 2021). This involves collecting as searching for correlations in large archives of scientific papers (Li, 2021)(Jeantet, 2021) or machine learning (Bentley et al, 2021).

5. Music Evolution

Music often has strong connections to past works. The evolution of electronic music has been reconstructed in one project (Youngblood et al, 2021). Other projects look at the evolution of music across genres (Savage, 2020) and across cultures (Savage et al, 2020).

5. Synthesis

By understanding texts as sitting in evolutionary relationships to each other we can create standards which allow parallel bibliographic description projects to be compatible with library catalogue and each other. By analogy with phylogenetics we can outline a phylomemetic synthesis.

Domains	Mendelian Projects - memes	Darwinian Projects - descent
Storytelling/ Fairy Tales/ Theater	Literary Criticism: Character Archetypes, Plot Structure Science Fiction and Fantasy Conceits Fairy Tale Motifs TvTropes	Literary Criticism: Stemmatics Genealogy of Plays
Academics	Ontology of Research Methods Subject Headings	Citations
Science	Structured Scientific Concepts Ontology of Research Methods	Scientific Topic Evolution Asimov's timeline of science
Music	Music Basic Concepts	Music Evolution

Table 1: Mendelian and Darwinian bibliographic description projects across several domains

Neither this list of domains nor this list projects is exhaustive. A possible additional domain is cookbooks. From a Mendelian perspective each recipe would consist of granular components like ingredients and cooking methods, or even the symbolism or perception of healthiness of the foods. From a Darwinian perspective each recipe also experiences modification over time.

Returning to the notion of a synthesis, we are left with the question of what bringing all these projects together would look like and what functionality doing so might enable.

Stemmatics illustrates the sort of complex family tree visualizations that would be possible. These visualizations could in principle be dynamically generated from any Darwinian project's data.

The real power of the phylomemetic synthesis is rooted in its analogy with biology. If two species share a gene we can infer that their common ancestor likely shared that gene as well (Urry et al, 2017). Similarly, if two texts share a meme we can infer that their common ancestor likely shared that meme as well. This could be useful for guiding textual criticism. But it could also allow the catalogue to suggest some components for the record of an as yet uncatalogued text.

Additionally, if two species share a considerable amount of genetic material, we can infer that they are somewhat closely related, even if that connection has not yet been otherwise established. And so it could go with texts.

6.   Conclusion

Phylomemetics is a high-level abstract framework for the unification of many different projects and efforts into the library catalogue. While the general idea is still at a somewhat early stage, phylomemetics represents a way of making space for many alternative bibliographic description projects, inviting them into the library catalogue to enrich us all.

Several possibilities for what that enrichment may look like can be ventured, even at this early stage. For example, Phylomemetics provides a framework under which metadata generated from outside the library cataloguing community could be integrated into the catalogue. Perhaps in the future this could serve as a model for incorporating user participation in the catalogue in a productive manner.

Another possibility is that of visualizations. These enriched catalogue records could be used to generate family trees of works, or cladograms of particular memes. These visualizations could provide users with new methods of browsing the collection.

Much work still needs to be done before these possibilities become actualities, but we have the opportunity now to invite in these parallel cataloging projects.

References

Authorities. (2019) Library of Congress Authorities. *Library of Congress*. https://authorities.loc.gov/

Asimov, I. (1989). *Asimov's chronology of science and discovery*.

Asubiaro, T., (2021). *Exploiting Semantic Similarity Between Citation Contexts For Direct Citation Weighting And Residual Citation*. University of Western Ontario Electronic Thesis and Dissertation Repository. https://ir.lib.uwo.ca/etd/8008

Becvar, D. S. (2005). *Tracking the archetype of the wise woman/crone*. ReVision, 28(1), 20-24.https://link.gale.com/apps/doc/A136649716/AONE?u=lond95336&sid=bookmark-AONE&xid=a130b9e8

Bentley, R. A., Borycz, J., Carrignon, S., Ruck, D. J., & O'Brien, M. J. (2021). Machine learning for rediscovering revolutionary ideas of the past. *Adaptive Behavior,*. https://journals.sagepub.com/doi/10.1177/1059712320983045

Blackmore, S. (2017). *About Memes.* https://www.susanblackmore.uk/memetics/about-memes/

Booker, C. (2004). *The seven basic plots: Why we tell stories*. A&C Black.

Britannica. (2022). *The work of Mendel.*
https://www.britannica.com/science/genetics/The-work-of-Mendel

Britannica. (2022). *Evolution: scientific theory.*
https://www.britannica.com/science/evolution-scientific-theory

Campbell, J. (1949). *The Hero with a Thousand Faces*.

Dawkins, R. (1976). *The selfish gene*.

Despenser, S. (2004). Life and death: the white horse. *Psychodynamic Practice*, 10(4), 522-528.
https://doi.org/10.1080/14753630412331313749

Bray, J. (2020). Fantasy Classification System. *Fancyclopedia 3*.
http://fancyclopedia.org/Fantasy-classification-system

Howe, C. J., & Windram, H. F. (2011). Phylomemetics—Evolutionary Analysis beyond the
Gene. *PLOS Biology*, *9*(5), 5. https://doi.org/10.1371/journal.pbio.1001069

Jan, S. (1999). The Selfish Meme: Particularity, Replication, and Evolution in Musical Style.
International Journal of Musicology. Vol. 8 (1999), pp. 9-76 (68 pages).
https://www.jstor.org/stable/24621004?seq=1

Jane-athon. (2015). *Jane-athon*. http://rballs.info/topics/p/jane/janeathon.html

Jeantet, I. (2021). *Hierarchical and temporal analysis of scientific corpora as tools for the
history of science* (Doctoral dissertation, Université Rennes 1).
https://tel.archives-ouvertes.fr/tel-03108773

Jones, J. (2020). Every Possible Kind of Science Fiction Story: An Exhaustive List Created by
Pioneering 1920s SciFi Writer Clare Winger Harris (1931). *Open Culture.*
http://www.openculture.com/2020/08/every-possible-kind-of-science-fiction-story-1931.
html

Lamere, P. (2008). Social tagging and music information retrieval. *Journal of New Music
Research*, *37*(2), 101–114.

Li, K. (2021). *Exploring Topic Evolution in Large Scientific Archives with Pivot Graphs*
(Doctoral dissertation, Sorbonne Université).
https://hal.archives-ouvertes.fr/tel-03297258/document

Li, K., Naacke, H., & Amann, B. (2021). An Analytic Graph Data Model and Query Language
for Exploring the Evolution of Science. *Big Data Research*, 26, 100247.
https://www.sciencedirect.com/science/article/abs/pii/S2214579621000642

Library of Congress. (2015). *Searching Subject Authority Headings.*
https://authorities.loc.gov/help/subj-auth.htm

Littlewood, A. R. (1968). The symbolism of the apple in Greek and Roman literature. *Harvard
Studies in Classical Philology*, 72, 147-181.

Maas, P. A.. (2009). Computer Aided Stemmatics — The Case of Fifty-Two Text Versions of
Carakasaṃhitā Vimānasthāna 8.67-157. *Wiener Zeitschrift Für Die Kunde Südasiens /
Vienna Journal of South Asian Studies*, *52/53*, 63–119. Retrieved from
http://www.jstor.org/stable/24008034

Propp, V. (1968). *Morphology of the Folktale*. University of Texas Press.

RDA. (2019). Introduction to RDA. *International Federation of Library Associations and Institutions*. https://access.rdatoolkit.org/Guidance/Index?externalId=en-US_ala-d3b460ae-1818-3459-8c9c-9dd5fda8edf8

Ghys, T. (2012). Technology trees: Freedom and determinism in historical strategy games. *Game Studies*, 12(1), 143-52. http://gamestudies.org/1201/articles/tuur_ghys

Savage, P. E. (2020). Measuring the cultural evolution of music: Cross-cultural and cross-genre case studies. https://doi.org/10.31234/osf.io/mxrkw

Savage, P. E., Chiba, G., Currie, T. E., Suzuki, H., & Atkinson, Q. (2020). Sequence alignment of folk song melodies reveals cross-cultural mechanisms of musical evolution. https://doi.org/10.31234/osf.io/5rj6y

Scholer, D. M. (1990). Q bibliography. In Society of Biblical Literature Seminar papers (Vol. 126, No. 29, pp. 11-11). Scholars Press, Atlanta 1965–2003, ISSN 0160-7588

Sheidlower. J. (2022). *Historical Dictionary of Science Fiction.* https://sfdictionary.com/

Singh, C., Barme, E., Ward, R., Tupikina, L., & Santolini, M. (2021). Quantifying the rise and fall of scientific fields. https://arxiv.org/abs/2107.03749

Smith, T. R., Zeng, M. L., & Knowledge, A. D. E. P. T. (2002). Structured models of scientific concepts as a basis for organizing, accessing, and using learning materials. https://www.researchgate.net/publication/277294974

Soldatova, L. N., & King, R. D. (2006). An ontology of scientific experiments. *Journal of the Royal Society Interface*, 3(11), 795-803. https://doi.org/10.1098/rsif.2006.0134

Steenberg, L. (2020). *Are You Not Entertained?: Mapping the Gladiator Across Visual Media*. Bloomsbury Publishing.

Szostak, R., & Smiraglia, R. (2019). Classifying Music within the Basic Concepts Classification. *Proceedings of the Annual Conference of CAIS / Actes Du congrès Annuel De l'ACSI.* https://doi.org/10.29173/cais1064

TvTropes. (2022). *TVTropes*. https://tvtropes.org/pmwiki/pmwiki.php/Wiki/TVTropes

Uther, H.-J., & Folklore Fellows. (2004). *The types of international folktales: A classification and bibliography, based on the system of Antti Aarne and Stith Thompson.* Helsinki: Suomalainen Tiedeakatemia, Academia Scientiarum Fennica.

Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Reece, J. B. (2017). *Campbell Biology 11th Edition*. Pearson Education, Incorporated.

Williams, G. J. (1997). *Our moonlight revels: A Midsummer Night's Dream in the theatre*. Iowa City: University of Iowa Press.

Youngblood, M., Baraghith, K., & Savage, P. E. (2021). Phylogenetic reconstruction of the cultural evolution of electronic music via dynamic community detection (1975–1999). *Evolution and Human Behavior*, 42(6), 573-582. https://arxiv.org/pdf/2011.02460.pdf