Implementing a Digital Library for Afrikaans Poetry I: Theoretical Foundations

Karen de Wet and Theo J.D. Bothma

Introduction

In this article a discussion is offered of a project that started out as a multimedia project aimed at making Afrikaans poetry accessible, before it evolved into the realm of a digital library being built in an XML environment. The complementary natures of the hypertext environment and the system theory approach to literature, are described in order to clarify the theoretical foundations and context of the project. The possibilities offered by envisaging and structuring the project as a digital library are discussed—this is elaborated in Part II where an indication is given of the advantages and challenges involved in situating the digital library in an XML-environment.

1. History

The idea for this project originated from a contemplation on the complementary natures of the hypertext environment and the system theory approach to literature.

The theoretical origins lie in Even-Zohar's formulation of the literary system theory, as well as George Landow's views on literary theory and hypertext that he put into practice in e.g. *The Dickens Web*[1]. This gave rise to the idea of creating an environment (at first aimed at a text archive, later identified as the makings of a digital library) where it will be possible to

[1] Built by Landow, Launhardt & Kahn, the original Intermedia *Dickens Web* has been published in Storyspace, as is the case with a similar project on Tennyson (*In Memoriam Web*). Both are now also integrated into the HTML published *Victorian Web* (http://landow.stg.brown.edu/victorian/victov.html).

match system theoretically identified relations between defined entities within a specific literary system, to the linkable functionalities made possible by hyperlinks in a hypermedia environment.

Products such as *The Dickens Web* use a hypertext environment to offer access to databases containing primary literary works as well as relevant surrounding and contextualising texts. The Intermedia version of *The Dickens Web* comprised of different 'set(s) of links joining documents on a similar topic' to which entry points were made available via 'two main overviews, one focusing on Dickens as an author and the other focusing on the central literary work, *Great Expectations*' (Landow & Kahn 1992:150). One can also refer to the work done by Chadwyck-Healy. This partnership originally created electronic collections of English and American poetry: *The English Poetry Full-Text Database and The American Poetry Full-Text Database.* They are at present integrated into the *Literature Online*-collection[2]. Collections such as these can, however, be greatly enhanced regarding scope, accessibility, usability and continuous updating—when it is presented as a full digital library, rather than only another combination of functionalities such as interface, search tool and databases. On the 10th of January 2001 the *Literature Online for Schools* site was officially launched, 'dedicated to supporting the teaching and study of English Literature' and offering 'hundreds of study unit pages dedicated to particular authors and topics [bringing] together texts, critical articles, biographies and web links and are surrounded by a mass of other useful material: 90,000 poems, 800 plays, 1000 prose works' (Motion 2001).

In a similar manner a digital library of poetry could be defined as a comprehensive collection of all poems belonging to a specific language, genre, period (or any other specification)—or a combination of these. What is of importance here is to consider how such a library would be able to surpass a *Dickens Web* or digital poetry collections (even when it is situated in an active hypermedia environment). The premise is that existing collections can be greatly enhanced when envisaged and incorporated into a digital library—especially because of the possibilities arising for adding various additional information items to the database containing primary texts. Such a library could for instance offer integrated access to databases containing critical analyses of poems, biographical as well as bibliographical data on the poets,

[2] Described as 'bring(ing) together the contents of ... databases into one, fully searchable resource' (http://www.chadwyck.com/products/viewproduct.asp?key=784).

explanatory notes aimed at supporting the reading process in an educational environment, close readings of the poems, lexias on literary history, etc. Such an extended library could be a most useful tool—for both teaching and research.

After an informal survey amongst teachers and lecturers of Afrikaans poetry in South Africa (part of it conducted during three formal workshops on the reading and teaching of poetry[3]), it was established that the need does exist for such a more comprehensive resource for teaching at both school and university level; as well as for use in a research environment. It was also immediately evident that an undertaking of this nature falls outside the scope of a single individual and that the project would only be able to succeed with the participation of experts working as part of a team. The expertise of the team working on the project includes the fields of hypermedia technology, XML-programming, information architecture, literary theory and knowledge management.

The endeavour is thus to create a digital library by combining the following four aspects:

1. a theoretical approach to literature;
2. the primary literary text itself (for instance a poem), including making available the web of systemic relations with other entities in the identified literary system that surround them;
3. all relevant aspects that can be captured and that will eventually comprise the ongoing capturing of literary history in the making; and
4. the possibilities offered by the technological developments in the fields of multimedia, hypertext, and mark-up languages.

Therefore, it can be said that this project builds on the endeavour to combine a system theory approach to literature, with the non-linear, interactive and searchable possibilities offered by multimedia and hypertext technology, all being made possible and accessible in an XML-environment.

2. Theoretical Foundations

The system-theoretical approach to literature and literary historiography, and

[3] Conducted during the *Lenteseminaar vir Afrikaansonderwysers* at the University of Pretoria: September 1996; as well as 11 & 12 September 1998; (see De Wet 1996; 1998).

specifically Even-Zohar's poly-system theory, is already quite established. In previous research De Wet (2000) argued that when literature (or a specified part thereof, e.g. a part defined according to language or genre, as is the case in this project, being focused on Afrikaans poetry) is regarded as a literary system, it means that everything / anybody that bears relation to it / influences it, can be identified as entities within this open, flexible system. The relations that exist between these entities (e.g. poet, text, oeuvre, publisher, reader, interpretations, critics, etc.) can be described by regarding them as utterances in a kind of literary discourse.

These relations function within the identified literary system as forces, influencing not only the entities, as well as their positions and actions within the system, but inevitably also the whole of the identified system, its nature, its temporary and shifting boundaries, the whole systemic functioning thereof. The moment the researcher is focused on these relations, it means that whatever is observed there, is actually disclosing the state of the literary system itself. And if this can be observed, it opens the way to identifying the different literary discourses generated by different entities that assume a (sometimes only temporary) pivotal function within the specific literary system. The term 'pivotal function' is used to indicate the ability of an entity to generate diverging and/or converging forces that influence the relations (and therefore entities) of that specific literary system. Obviously (generated) forces within a (literary) system will be influencing at least *some* of the entities, and definitely (aspects of) the whole of the system. It follows, therefore, that the literary researcher will be interested in such forces and their manifestations, as well as their origins, because if these forces/relations between literary entities can be described (according to, e.g. origin and influence) and recorded, it opens the way to a convenient (and one dare say less subjective, less periodised, less genre-oriented, less canonised) possibility of a type of literary history that essentially consists of a polyphony of voices.

There is thus practical advantage to be gained from subscribing to the view of a system-theoretical approach to literature and literary historiography, because it enables the researcher to identify forces (i.e. literary discourses) generated by pivotal entities within the identified literary system, and therefore opens up the possibility to record, on a regular basis, such discourses which can eventually result in a literary history offering a contemporary view.

Even-Zohar already pointed out in 1979 that the systemic approach is a more meaningful and more scientific principle of arrangement than a positivistic collection of data. His words were:

> the positivistic collection of data, taken *bona fide* on empiricist grounds and analysed on the basis of their material substance, has been replaced by a functional approach based on the analysis of relations.

At the University of Leuven in April 1996 (during a conference celebrating its predecessor twenty-five years earlier that initiated the system theory school) Even-Zohar told his audience of scientists involved in system theory research: 'The theory exists, it has been described, now apply it so that it can become means for investigation'. This utterance can be seen as part of the incentive for this researching and building of a digital library for Afrikaans poetry.

A system theory approach to literature implies that the focus falls on the dynamic, interactive and procedural relations existing between system entities. It is clear, Landow (1997:4) argues, that

> scholarly articles situate themselves within a field of relations, most of which print medium keeps out of sight and relatively difficult to follow, because the referenced (or linked) material lie spatially distant from the references to them. Electronic hypertext, in contrast, makes individual references easy to follow and the entire field of interconnections obvious and easy to navigate ... the article would now be woven more tightly into its context than would be a printed counterpart.

George Landow (1992) was the first researcher to point out the obvious parallels between the technology of hypertext and its possibilities, and the notions that thinkers such as Barthes, Derrida and Foucault have on networked texts. Both the view of literature as consisting of networked texts, and the hypertext enabling non-linear webs of text through hyperlinking, support the system theory approach to literature where the focus falls on the dynamic, interactive and procedural relations existing between system entities. As Katz and Kahn (1966:18) pointed out: 'System theory is basically concerned with problems of relationships, of structure, and of interdependence rather than with constant attributes of objects'.

This is the broad picture—with endless possibilities and advantages. It means that, for example, if a poem gets published today and a review appears tomorrow, all the related information can be hyperlinked, which will further

literary research (as well as the availability thereof) immensely. On an educational level it means if the teacher is discussing certain themes in the poetic oeuvre of a certain poet, this valuable research information can be hyperlinked to whatever key terms in the poems are applicable, and thus be made available—as part of the web of the literary system, and *ipso facto* captured as part of the digital library.

3. The Digital Library

Views on what a digital library is, what its structure and functions should be, and how it can be defined, vary immensely—as can be seen from the different contexts where the term is used. Even a resource from which one would expect a more focused and in a sense 'scientific' approach to the use of the term, such as IFLANET's Electronic Collections on Resources and Projects regarding Digital Libraries[4] does not distinguish between the various endeavours, goals or even 'formats' of projects that use the term 'digital library'.

Although there is clearly a research need here, this issue will not be addressed further in the context of the present article. Suffice to say that we are using the term 'digital library' for this project, on the following grounds:

- it offers searchable and expandable databases;
- it functions in the digital environment;
- it offers options to create virtual documents;
- it collects and archives records in digital format; and
- it enables the capturing and managing of information that would otherwise be lost.

It also should be clear why the project migrated to the environment of a digital library:

- not only because of its resource richness,
- and because of the management of the underlying databases and the information contained in it,

[4] (http://www.ifla.org/II/diglib.htm).

- but also because of its ability to accommodate contributions from content providers that can be provided for within a virtual community,

- and ultimately because of the possibilities to also capture, process and make information available—in the knowledge management sense of the word.

In *The Archaeology of Knowledge* Foucault (1976:23) points out that the 'frontiers of a book are never clear-cut [because] it is caught up in a system of references to other books, other texts, other sentences: it is a node within a network ... a network of references'.

What is being done in the building of this project's digital library, therefore, is to start with a poem (as a basic entity within the literary system) and then create a web of interpretative possibilities, intertextual references, as well as inter-systemic relations surrounding it. This 'web' comes into being by documenting and/or capturing information on both the entity as well as the relations it is displaying within the literary system (metadata is used), as well as by establishing hyperlinks to relevant chunks of information maintained in the various databases of the digital library (using information architecture that builds on entity relationships and slice design).

This part of the product can therefore be used to aid teachers of literature, because the role of the interpreting mediator can be fulfilled by the availability of clickable contextualizing information.

The use and range of the product is, however, not limited to an educational and interpretative environment. The digital library can also be used by poetry lovers who would like to gain additional knowledge on poems, poets, oeuvres, intertexts—or who would, for example, like to see how the poet looked like at that stage of his/her life when writing the poem, or hear a poem read by its author.

A third intended user group of our digital library project is literary scholars and researchers, who can use the advanced search features of the digital library made possible by the metadata and the XML environment in which the project is being built—an environment specifically decided upon because of the advantages it offers regarding on-the-fly linking, advanced searching, and possibilities for the creation of virtual documents.

Thus, by combining (1) a theoretical approach to literature, and (2) a technologically empowering format, not only an educational tool for aiding

the interpretation and enjoyment of poetry is being established, but also a tool for capturing knowledge and managing it—up to the point of being able to create a virtual literary historiography. This possibility arises because of the content and modelling of the databases built into the project from where data is summoned whenever a link is activated. Typically these databases need not only offer information needed for the contextualised and interactive interpretation of a poem, but can also be used to establish the basis of a literary historiography that does not need to 'be written', as it is constantly in the process of being written. The implications and possibilities regarding the practical application of knowledge management strategies here, provide for very rich applications within the project. A practical example of one of the virtual aspects that can be created by the digital library, is the automated creation-on-demand of a timeline, or the creation of other virtual documents, e.g. virtual anthologies.

The concept, construction and possibilities of a digital library is therefore used in this project to make explicit, tacit and useful the foundations of several theoretic constructs. For instance: the whole point of taking the system theory view of literature into account, is to show that it is possible to match this system theory approach to the physical structure of a digital library.

A digital library offers different searchable databases of information[5] on the identified entities of the literary system—the most basic of these being:

- a database of biographical information on authors;
- a database on oeuvres (created virtually);
- a text archive containing:
    - the primary texts (the poems);
    - as well as links to information that can contextualise the poem and/or aid with the interpretation thereof;
    - and links making explicit (references to) intertexts;
- databases containing secondary texts of scholarly activity on any aspects of the identified literary system—e.g. reviews, close readings, interpretations, comparative readings, etc.; and
- databases containing dictionaries and glossaries.

[5] These databases can contain material in multimedia format—it is not only text that is referred to. For instance, biographical information on an author can also contain photographs, voice clips, or video clips.

Conclusion

We may conclude and point out that what we are aiming for in this project of building a digital library, is:

- to capture data (in multimedia format) that represent aspects that can be viewed as entities of a literary system;
- to mark it up (using XML, Dublin Core, and TEI headers);
- to hyperlink (using both soft and hard linking methods) it to other entities; and
- to activate it within an identified information architecture (relying on entity relationships and slice design, using the RMM-model[6].

Through this process the theorised relations between entities in the literary system, are being made visible and interactively usable and searchable. We are thus using the possibilities of non-linear, multi-linear and multi-sequential text within a hypertext environment to make (1) visible and (2) readily accessible, an entire field of previously inaccessible intertextual relations and interpretations—and: with the additional benefit that all of this becomes woven into the context of the literary system. It may be true that contextually relevant information and knowledge of intertexts (and/or intertextual references) that our digital library project is able to make available, can be known (and at least partially) to scholars or experienced and sophisticated readers; but the visibility and usability thereof via what the digital library offers, proves to be invaluable to the effort of both capturing this knowledge as well as making it available upon request. And specifically upon request: this is something we deem very important in reference to the many different user groups we aim to be able to serve with the digital library, as well as in reference to the familiar problem of information overload: the links to supportive, explanatory, contextualising, and intertextual information is made visible and usable to the user, but it remains a choice to access it.

Department of Information Science
University of Pretoria

[6] Relationship Management Model (Isakowitz et al. 1995).

References

Bothma, TJD, K De Wet, P Bothma, & H Pretorius 2000. Implementing a Digital Library for Afrikaans Poetry in XML. *Proceedings of the 24th Online Information Meeting London 5 - 7 December 2000*, pp. 193 - 200.

Bothma, TJD, K de Wet, & P Bothma 2002. Implementing a Digital Library for Afrikaans Poetry II: Technical Issues. *Alternation* 9:2:150-166.

De Wet, K 1996. Poësieplan: St. 6-8. Unpublished Paper and Workshop. *Lenteseminaar vir Afrikaansonderwysers*, Department of Afrikaans, University of Pretoria. 5 September 1996.

De Wet, K 1998. 'Speel die spel met woorde': Voorstelle vir 'n Poësie Lesreeks (Junior Grade, Hoërskool). Unpublished Paper and Workshop. *Lenteseminaar vir Afrikaansonderwysers*, Department of Afrikaans, University of Pretoria. 11 September 1998.

De Wet, K 1998. Werkwinkel vir Voorgeskrewe Gedigte, 1998. Unpublished Paper and Workshop. *Lenteseminaar vir Afrikaansonderwysers*, Department of Afrikaans, University of Pretoria. 12 September 1998.

De Wet, K 2000. Dialogues Generated by Pivotal Figures in Literary Systems: A Systemic Approach to the Study of Literature. In D'haen, T (ed): *Methods for the Study of Literature as Cultural Memory, Volume 6 of the Proceedings of the XVth Congress of the International Comparative Literature Association*, Leiden. Amsterdam/Atlanta: Rodopi.

De Wet, K 2000. Creating a Digital Library: Lessons Learned from a Multimedia Poetry Project. Unpublished Paper. *ProLissa 2000*, Pretoria, South Africa. 27 October 2000.

Even-Zohar, I 1979. Polysystem Theory. *Poetics Today* 1,102: 287-310.

Foucault, M 1976. *The Archaeology of Knowledge and the Discourse of Language*. Sheridan Smith, AM (trans). New York: Harper & Row.

International Federation of Library Associations (IFLA). [Online] http://www.ifla. org/II/diglib.htm.

Isakowitz, T *et al.* 1995. RMM: A Methodology for Structured Hypermedia Design. *Communications of the ACM* 38,8: 34-44.

Katz, D & RL Kahn 1966. *The Social Psychology of Organisations*. New York: John Wiley.

Landow, GP 1992. *Hypertext: The Convergence of Contemporary Critical Theory and Technology*. Baltimore: John Hopkins University Press.

Landow, GP 1997. *Hypertext 2.0: The Convergence of Contemporary Critical Theory and Technology*. Baltimore: John Hopkins University Press.

Landow, GP & J Lanestedt (comp & ed) *In Memoriam Web*. Watergate: Eastgate Hypertext Systems.

Landow, GP (comp & ed) *The Dickens Web*. Watergate: Eastgate Hypertext Systems.

Landow, GP & P Kahn 1992. Where's the Hypertext: The Dickens Web as a System-independent Hypertext. *Proceedings of the ACM Conference on Hypertext; November 30 - December 4, 1992, Milan Italy*. [Online] http://dev.acm.org/pubs/articles/proceedings/hypertext/168466/p149landow/p149-landow.pdf

Literature Online. [Online] http://www.chadwyck.com/products/viewproduct.asp? key=784.

Literature Online for Schools. [Online] http://www.lion-schools.co.uk/.

Motion, A 2001. *Speech at Launch at BETT 2001*, January 10. [Online] http://www. chadwyck.co.uk/andrew%20motion.rtf.

The American Poetry Full-Text Database. [Online] http://www.chadwyck.com/ products/viewproduct.asp?key=784.

The English Poetry Full-Text Database. [Online] http://www.chadwyck.com/products/viewproduct.asp?key=784.

The Victorian Web. [Online] http://landow.stg.brown.edu/victorian/misc/credits. html.