

The Naïve Ontologist and “Intelligent” Story Telling

Mícheál Mac an Airchinnigh

School of Computer Science and Statistics, University of Dublin, Trinity College Dublin,
Ireland
mmacanai@cs.tcd.ie

Abstract.

Semantic Search has become, that is to say been re-discovered to be, the key theme for Intelligent Systems in 2012. No humans need apply. Semantic search, as Tim Berners-Lee predicted in his seminal little book on the World Wide Web, is the ultimate goal. After the Ultimate Goal has been obtained, and all the machines are happily communicating, what is in it for us, humans? To be sure, for us, it is to be more human! Humans tell stories, to themselves, and to other humans. Many of these stories have been, in the first instance, of an oral nature. Then they were written down. Now they appear electronically. The full spectrum of the technology of our age is used to reinvent the same old stories (7 in total). Precision in identity is sought through the formal ontology. A new story concerning Vela Peeva is revealed.

Keywords: game, keyimage, magic realism, ontology, semantic search, storytelling

1. Google’s Semantic Search

Out of the blue recently came the announcement that Google was doing “*Semantic Search*” [1], reported on the Wall Street Journal [2]. We already knew that Google’s digitization of the books was not only for the purpose of having electronic versions of same so that they might be “offered to the public” [3], but also that the electronic text might be searchable. Moreover, it is not the word that is sought, but the word (as used) in context. Another way of expressing this, is to say that context gives meaning to the word. Meaning is what we seek. Hence the “Seeking of meaning” is another way of talking about “Searching for semantics”. But what exactly does this “Semantic Search” entail? Ultimately, for Google, it might mean that the books “can talk to each other” on many different levels and at times arrive at a “Eureka!” moment of understanding which will be “output” to the Google AI Engine? Some idea of the colossal magnitude of the project is available on Wikipedia [3].

The reference to Wikipedia, cited above [1], presents an enormous amount of information on all aspects of “Semantic Search”. Of particular note, relevant to the “Island of Ireland”, is the “Semantic Search Engine”, SophiaSearch [4], developed in Belfast [5]. It eschews the idea of “(formal) ontology” and has just introduced the “Digital Librarian” [6]. But Wikipedia itself is a foundation for its own “inner Semantic Search.” Specifically, there is now a new initiative:

2 Mícheál Mac an Airchinnigh

“Wikidata” which “aims to create a free knowledge base about the world that can be read and edited by humans and machines alike. It will provide data in all the languages of the Wikimedia projects, and allow for the central access to data in a similar vein as Wikimedia Commons does for multimedia files. Wikidata is proposed as a new Wikimedia hosted and maintained project.” [7, 8].

We have also become accustomed to rely on the use of a formal ontology [9] such as the CIDOC Conceptual Reference Model [10] supported by (electronic) tools such as Protégé [11].

In the context of the WWW, we know from Tim Berners-Lee, who had a dream and in the second part of that dream “Machines become capable of analyzing all the data on the Web—the content, links, and transactions between people and computers. A “Semantic Web,” which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy, and our daily lives will be handled by machines, leaving humans to provide the inspiration and intuition [12, 13].

1.1 Нека си представим?

Let us imagine that the Berners-Lee view of the “Semantic Web” has come about? Let us imagine that Google’s “Semantic Search” is available to us? What shall we (as individual) do? In this paper we propose a return (ricorso) to the human normality of individuality of the person. In other words, given all of this “newly found” technology, how shall we be (more) human? In what sense shall we be free?

Storytelling is an ancient art form, originally oral, and now electronic. The key to good storytelling lies in the use of

- (i) stock phrases (“Once upon a time”),
- (ii) scene setting (“They wandered into the Dark Wood”),
- (iii) anticipation (“They heard footsteps behind them”,
- (iv) ambiguity (“What Big Eyes you have, Granny”),

and so on. A story can be re-told to the same people. Each re-telling adds to the understanding and pleasure of the listener.

In our times, it is customary that a story be accompanied by pictures, images that invoke some key aspect of the scene, or of a character, or of an object. Pictures have always been used to “tell stories” since, at least, Stone Age Man.

][How can I find information relating to this claim?][

A search is required. But to do the search I normally need to use some “key words.” Ideally, a key word may be formally attached to (or be derived from) a lemma (which is defined to be the *canonical form*, *dictionary form*, or *citation form* of a set of [words](#) ([headword](#))) [14]. One needs to adopt a specific

lemmatization [15] and harmonise it with the constructed ontology that I have been using for years, the CIDOC CRM 5.0.2 [16], processed by Protégé (now at version 4.2) [11].

For the record (of the intelligent story) we may wish to exhibit (and subsequently ontologize) the plan of the house of a character, say, Todor Stoinmenov Peshterski [17], who once lived in Velingrad (see Fig.1). In the figure shown, there is a photograph of what seems to be an elderly couple from the village of Draginovo. One assumes that the house shown, built in 1890, belongs to the man on the left, by the name of Yavor(?) Karabelov. Ontologization of this picture will provide us with a straightforward negation, that the house shown is not that of Peshterski.

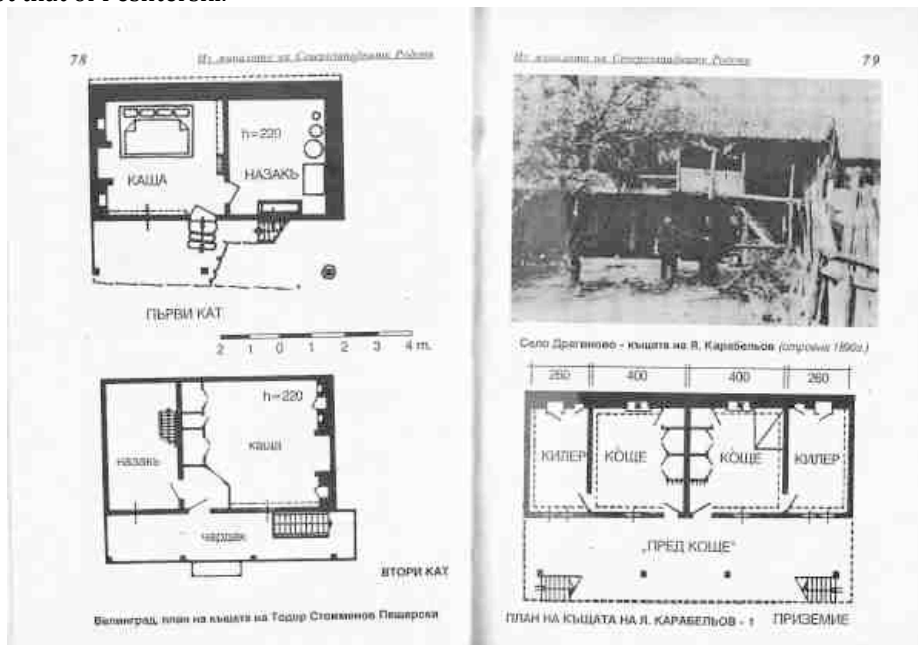


Figure 1: Велинград, план на къщата на Тодор Стоименов Пещерски

There have been, and there are, many “electronic” storytelling software platforms, such as LE-Story [18], SAGE [19] and KidPad [20]. Undoubtedly, there will be many more in future. I am not concerned with such software platforms. My interest is in the nature and structure of the story to be told. It is said that the plays of Shakespeare cover all of the stories that one might want to tell. That is to say, Shakespeare has covered all the storytelling categories. What are they? Not remembering the original source, one seeks to learn in the usual way in 2012 [21].

There are only 7 stories :

1. Tragedy: Hero with a fatal flaw meets tragic end. Macbeth [22] or Madame Bovary.

4 Mícheál Mac an Airchinnigh

2. Comedy: Not necessary laugh-out-loud, but always with a happy ending, typically of romantic fulfilment, as in Jane Austen.
3. Overcoming the Monster: As in Frankenstein or 'Jaws'. Its psychological appeal is obvious and eternal.
4. Voyage and Return: Booker argues that stories as diverse as Alice in Wonderland and H G Wells' The Time Machine and Coleridge's The Rime of the Ancient Mariner follow the same archetypal structure of personal development through leaving, then returning home.
5. Quest: Whether the quest is for a holy grail, a whale, or a kidnapped child it is the plot that links a lot of the most popular fiction. The quest plot links Lords of the Rings with Moby Dick and a thousand others in between.
6. Rags to Riches: The riches in question can be literal or metaphoric. See Cinderella, David Copperfield, Pygmalion.
7. Rebirth: The 'rebirth' plot - where a central character suddenly finds a new reason for living - can be seen in A Christmas Carol, It's a Wonderful Life, Crime and Punishment and Peer Gynt.

Naturally, there must be Bulgarian language categories to cover the same ground. What are they? "Google translate" provides good possibilities: трагедия, комедия, преодоляването на чудовището, пътуването и връщане, търсене, дрипи до богатство, прераждане. Now we need to give corresponding examples, but from Bulgarian culture. For example, what is the Bulgarian "equivalent" to Macbeth? Maybe Хан Аспарух [23]? As an aside, it is noteworthy that the Bulgarian Wikipedia often marks the stress syllable! Here is not the place to provide all these detailed inter-cultural correspondents. Google's Semantic Search engine will do that for us, if it is really "intelligent".

We need to have an intelligent storyteller, one who fits into a category, for our times. The rebirth category (7 above) suggests... renewal... Будител! One who awakens, enlightens, inspires, ... In principle, the Buditel is a(n) enlightener, a(n) (intelligent) story teller, one who looks to the future, remembering the past that can never be changed! Let us see how to play the role? We will need characters for the story. We need to enhance our continuing theme of Bulgarian Art/Culture. We will pick from the dead, and from the living. We will begin with a man (Тома ПЕТРОВ) who benefitted from the Bulgarian Enlightenment and continue with a woman Илона ЗАХАРИЕВА, a painter and former curator of Kyustendil Art Gallery (Кюстендил), the centre of the Maistora Collection.

1.2 Encoding the painter/painting "Тома ПЕТРОВ 1908-1972"

I am looking at a self-portrait of Toma Petrov (Тома ПЕТРОВ) as I write [24]. He is wearing a hat and smoking a cigarette. He wears a large overcoat and around his neck is a beautiful green scarf. The background to the self-portrait is mottled bright red, giving a purplish effect. One wonders what image springs to mind for the reader? At first the image was deliberately omitted in the belief that

the reader would be able to find that self-portrait online and verify that it is indeed Toma Petrov! Having failed to find such an image online, I have included it here (in black and white):

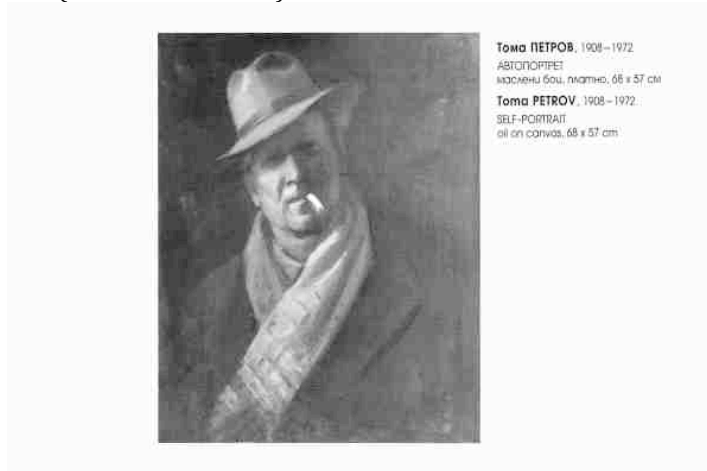


Figure 2 Тома ПЕТРОВ 1908–1972

Based on past experience of using Protégé and having checked the Bulgarian Wikipedia, it seems reasonable to encode Toma Petrov in the form:

E21 Person: ПЕТРОВ_Тома_1908-1972

E70 Thing: <http://www.flickr.com/photos/mihalorel/7086955705/>

E78 Collection: Sofia Art Gallery

E53 Place: СОФИЙСКА ГРАДСКА ХУДОЖЕСТВЕНА ГАЛЕРИЯ

where we are still using the Conceptual Reference Model (CRM) Ontological framework, with the aid of Stanford’s Protégé 4.2. Note that there is a reasonably good representation of the self portrait recorded as E70 and located on Flickr.

There are 43 persons identified as Toma Petrov on the Bulgarian Wikipedia [25]. Of those, 8 persons are listed but have no corresponding entries. On the corresponding page of the English Wikipedia [26] there are around 60 persons identified by the patronymic Petrov, none of whom is the artist in question. One obvious interpretation/hypothesis is that our artist is unknown, or of no importance, in the realm of Bulgarian Art?

For the woman, the (former) Curator and Painter, we choose her own visual identification

6 Mícheál Mac an Airchinnigh



Figure 3 Ilona Zaharieva

and encode (it/her) ontologically as

E21_Person: ЗАХАРИЕВА_Илона_р.1950

E22_Man-Made_Object: Автопортрет_ЗАХАРИЕВА_Илона_2000

E35_Title: АВТОБИОГРАФИЧНА КВАДРОКАЛУЦИЯ, 2000

E78 Collection: Sofia Art Gallery

E53 Place: СОФИЙСКА ГРАДСКА ХУДОЖЕСТВЕНА ГАЛЕРИЯ

This keyimage, shown above, is specifically chosen for the “Intelligent Story”, precisely because the author has met the person portrayed: Илона Захариева, former director of the Art Gallery in Kyustendil (Кюстендил) and as mentioned already, Kyustendil is chosen specifically because it uniquely associated with Vladimir Dimitrov — Maistora. In addition she was born in the same year as Mihal Orela.

1.3 Groups of Painters, Schools of Painting

One obvious easily accessible starting point for the general collective of public knowledge about Bulgarian artists is in Wikipedia [27]. Another might be the great variety of books on the subject, such as an historical scholarly account of Bulgarian Art in the context of a certain period such as 1904-1912 [28] or contained in certain archives such as that of the Bulgarian Academy of Sciences (BAS) [29]. A third popular source is the catalogues and books associated with either permanent exhibitions or special exhibitions. One can ontologise one’s favourite collection of painters in the form:

E74_Group: Bulgarian_Painters [same as]

E74_Group: Български_художници

Again, for the purpose of the (Intelligent) Story Telling one has recourse to the personal (re-)sources. The *early period* of “Тома ПЕТРОВ 1908-1972,” is well summed up in Milena Georgieva’s book (Sofia, 2008), “South Slav

Dialogues in Modernism.” Therein is mention of “The Exhibition of Serbian Artists of *Lada* Union at Trupko’s Gallery in Sofia (1911)” p.165.

One will be aware of the significance of Lada (Лада) “as vehicle” and even more so with Union (“Съюз”). However, although one might speculate about the possibility of “*Lada* Union” as a car-sponsoring organization, a more plausible (?) meaning is to be derived from Lada as Goddess [30]. Given the many possible interpretations, we shall decide to enter Лада (Lada) into the ontology with preferred reference to the Bulgarian Wikipedia entry [31].

E21_Person: Лада

Now we have ontologically stated that “our Lada” is a (female) person and not the car! Of course, there must be many more assertions to ground “our Lada”’s (presumed original) existence and reality, since she is a real character in an Intelligent Story. We will introduce Lada into our Intelligent Story later below.

The names of people are always of prime interest in storytelling. In Georgieva’s book there is mention of the place, the Art Gallery, where “Trupko” [Тръпко] hosted the Exhibition. Fortune would have it that Google’s “semantic search” [in Bulgarian] returns a unique Wikipedia page for Тръпко Василев [32]. The page is unique. There is currently no electronic equivalent in any other natural language? From the bibliographic information given, we also know that Тръпко was born in Личища [33].

E21_Person: Тръпко [same as] **E21_Person:** Тръпко Василев [same as] ...

E53 Place: Личища (originally part of ...)(now in Greece)(close to Ohrid)...

But here it is enough to get started with the Игра (game/play)! It is a happy fact that there is just one Bulgarian word for the “Play.” We can use it ambiguously to describe the story we are telling, as a play in, lets us say, 3 Acts? Since the paper limit is about 10 pages or so... We can set down the Play structure in the frame of the “Digital re-Discovery of Culture (DrDC) Game of Inquiry!” [34].

2. Let us imagine that the (Game/Play) defines the Culture?

We begin, as we must, with the identification of the player, or storyteller, or reader, or whomever. Here is the storyteller identified by:

OpenID: <https://profiles.google.com/Mihal.Orel/about> [8, p51]

2.1 Prelude

Lada is so ancient that there could not be an image of her that would survive into modern times? Could she be the equivalent of the figure Eve? Might not her Spirit re-appear again and again? Why did Ilona Zaharieva become, not only an artist in her own right, but take on the role of Director of the Art Gallery in Kyustendil?

2.2 The Play's the Thing: Game as Research Method

The idea of using Игpa (play/game) to explore the issues of Culture was presented in an early paper on the "Digital re-discovery of Culture" and the Game of Inquiry/Identity [34]. There are 5 basic elements: the Backstory text, the trinity of Web pages, the visuality of Keyimages including video, the End Goal, and the Keywords [35]. At the time, from experience, it was felt that the game needed focus and direction. The keyword was intended as a strong direct hint to the player. Let us sketch a basic game template?

[1] *The Backstory (after Menard [36])*

When I was a young boy, around the age of 12, I visited Velingrad with my parents. I remember well the old house of Todor Peshterski, a very close friend of my father. I remember going into the storeroom secretly while the adults were busy, drinking some rakia and/or ayran. There was lots of old stuff, especially things that I knew to be Ottoman. But in one corner there was a picture of Peshterski, all dressed up in strange coloured clothes like the haiduk revolutionary, Panayot Hitov [37]. It was a coloured painting about my size. I went closer to take a look. I was mesmerized. I do not know how long I stared at the portrait. My mother and Maika Peshterski had come to find me. They said that I looked just like a statue, staring at the painting. That was the day I knew I would become a painter.

[2] *Web Pages (3)*

- 1) http://bg.wikipedia.org/wiki/Вела_Пеева [20120419]
- 2) http://bg.wikipedia.org/wiki/Панайот_Хитов [20120419]
- 3) <http://bg.wikipedia.org/wiki/Пещерски> [20120419]



[3] *Keyimage(s)*

Every image is semantically (over)loaded, being layered with many meanings, each of which is revealed according to the cultural background of the beholder and the circumstance of the time and place of such beholding. The image is decoded by the Eye of the Beholder depending upon the circumstances at the time. For this game, the following keyimage is chosen:

[4] *Goal: When and where did Toma Petrov meet Vela Peeva?*

[5] *Keywords*: Vela Peeva, Chepino, Ladzhene, Kamenitsa, Kleptuza

The play/game is won/finished when the player achieves internal personal psychological closure. There may be more than one ending/solution to the play/game depending on the culture of the player.

2.3 Postlude

There is no evidence to suggest that Lada participated in the painting frenzy in Lascaux [38]. It is good to remember that we have been playing with words, with images, with history. Characters have been brought together, some real, some fictional. Since there are many images involved, one needs to be able to connect them with the corresponding Ontology. Currently, the author uses KeyNote (the equivalent of PowerPoint) and Flickr for this purpose. In the Protégé ontology, the comment section is used to host the links to the images concerned. Finally, it must be understood that we have been playing a game in the strict sense. Much of the fun is obtained by finding the deliberate contradictions and impossibilities. It is Category 7: Rebirth, rediscovery, re-enlightenment. Ultimately, one needs to know how Velinograd came to be!

3 Acknowledgements

A special thanx are due to Михаил Орела who provided much of the Bulgarian material used in the Intelligent Story Telling.

References

1. Wikipedia editors *Semantic search*, http://en.wikipedia.org/wiki/Semantic_search. 2012.
2. Efrati, A. *Google Gives Search a Refresh*. 2012 [cited 2012 March 29]; Available from: <http://online.wsj.com/article/SB10001424052702304459804577281842851136290.html>.
3. Wikipedia Editors *Google Books Library Project*. Wikipedia, 2012.
4. SophiaSearch. 2012; Available from: <http://www.sophiasearch.com/>.
5. Wikipedia Editors *Sophia Search Limited*. 2012.
6. SyncNI *Sophia Search Discovers US Market*. 2012.
7. Wikimedia. *Wikidata*. 2012 [cited 2012 April 13]; Available from: <http://meta.wikimedia.org/wiki/Wikidata>.
8. Perez, S. *Wikipedia's Next Big Thing: Wikidata, A Machine-Readable, User-Editable Database Funded By Google, Paul Allen And Others*. 2012 [cited

10 **Mícheál Mac an Airchinnigh**

- 2012 April 13]; Available from: <http://techcrunch.com/2012/03/30/wikipedias-next-big-thing-wikidata-a-machine-readable-user-editable-database-funded-by-google-paul-allen-and-others/>.
9. Wikipedia editors *Ontology (information science)*. 2012. **2011**.
10. *The CIDOC Conceptual Reference Model*. [cited 2011 Feb 27]; Available from: http://www.cidoc-crm.org/official_release_cidoc.html.
11. Stanford Center for Biomedical Informatics Research. *Protégé 4.2*. 2012 [cited 2012 April 5]; Available from: <http://protege.stanford.edu/>.
12. Berners-Lee, T. and M. Fischetti, *Weaving the Web : the original design and ultimate destiny of the World Wide Web by its inventor*. 1st pbk. ed2000, New York: HarperCollins Publishers. ix, 246 p.
13. W3C. *Semantic Web*. [cited 2012 April 13]; Available from: <http://www.w3.org/standards/semanticweb/>.
14. Wikipedia editors *Lemma (morphology)*. 2010.
15. Wikipedia Editors *Lemmatisation*. 2012.
16. *Erlangen CRM / OWL 2010-10-01 based on CIDOC CRM 5.0.2 January 2010*. 2010 [cited 2011 Jan 28]; Available from: <http://erlangen-crm.org/current-version>.
17. Арнаудов, А., *Из миналото на северозападните родопи*, 1995, Пазарджик: ИК „Беллопринт“.
18. Yu-Lin Chu, Chun-Hung Lu, and W.-L. Hsu *LE-Story: An Intelligent Storytelling Environment for Chinese Learning*. 2009.
19. MIT. *SAGE (Storyteller Agent Generation Environment)*. 2012 [cited 2012 April 4]; Available from: <http://xenia.media.mit.edu/~marinau/Sage/>.
20. University of Maryland. *KidPad*. 1998 [cited 2012 April 4]; Available from: <http://www.cs.umd.edu/hcil/kiddesign/kidpad.shtml>.
21. Haig, M. *The Seven Stories That Rule the World*. 2008 [cited 2012 April 16]; Available from: <http://us.penguin.com/static/html/blogs/seven-stories-rule-world-matt-haig>.
22. Wikipedia Editors *Макбет*. 2012.
23. Wikipedia editors *Asparuk of Bulgaria*. 2012.
24. Petrov, T., *SELF-PORTRAIT, 1908-1972, "SELF-PORTRAIT, The visible image and the hidden meaning"*, Sofia Art Gallery, 2008: Sofia.
25. Уикипедия. *Петров*. 2012 [cited 2012 March 26]; Available from: [FinalPaper.doc](#).
26. Wikipedia. *Petrov (surname)*. [cited 2012 March 26]; Available from: [http://en.wikipedia.org/wiki/Petrov_\(surname\)](http://en.wikipedia.org/wiki/Petrov_(surname)).
27. Уикипедия. *Категория:Български художници*. 2012 [cited 2012 Аврил 14]; Available from: [FinalPaper.doc](#).
28. Georgieva, M., *South Slav Dialogues in Modernism 2008*, Sofia: Bulgarian Academy of Sciences.
29. BAS. *Музей "Стара София" - Художествена колекция*. [cited 2012 2012-03-18]; Available from: <http://oldsofiaart.cl.bas.bg/>.
30. Wikipedia Editors *Lada (goddess)*. 2012.
31. Уикипедия *Лада*. 2012.

32. Wikipedia Editors *Трънко Василев*. 2012.
33. Wikipedia Editors *Личица*. 2012.
34. Mac an Airchinnigh, M., K. Sotirova, and Y. Tonta, *Digital Re-discovery of Culture Game of Inquiry & The Physicality of Soul, 2006*. Review of the National Center for Digitization., 2006: p. 19-37.
35. Mac an Airchinnigh, M., *Digital re-discovery of Culture and the Game of Inquiry/Identity*.
36. Wikipedia editors *Luis. Pierre Menard, Author of the Quixote. 1939*. 2012.
37. Wikimedia Commons. *File:Panayot Hitov Photo.jpg*. [cited 2012 April 18]; Available from: http://commons.wikimedia.org/wiki/File:Panayot_Hitov_Photo.jpg.
38. Wikipedia Editors *Lascaux*. 2012.