

Opening the can: public interaction with ready-made contents

Shlomo Dubnov¹, Philippe Codognet²

¹ Music Department,
University of California in San Diego
La Jolla, CA, USA
sdubnov@ucsd.edu

² CNRS / UPMC/ University of Tokyo
JFLI, Information Technology Center,
2-11-16 Yayoi, Bunkyo-ku
113-8658 Tokyo, Japan
codognet@jfli.ite.u-tokyo.ac.jp

Abstract. We present a system for public interaction during presentations of non-interactive readymade (canned) materials such as films or pre-composed performances. Using concepts of user back-channeling, moderator and shared canvas, the system allows a significant part of an art piece that is actually happening “in the heads” of the audience to be shared in parallel or in sequential manner with the main show. We also briefly present the historical background of the blind manipulation of symbols (canned contents) and argue that a new cognitive paradigm is at work with user interaction in computer-based entertainment and digital art.

1 Introduction

Most of the contents on the web are not interactive. Videos, music, images and other media are abundantly produced by professionals and amateurs and put on the web as readymade contents that express their creators' own voices. Even though most of these contents are available for reuse under various kinds of fair use licenses, only a small fraction of it indeed finds its way into other productions. User interaction with such contents is usually limited to search or recommendation of related contents and textual commenting. Even then, the sometimes dubious quality of recommendations and the shallow comments leave most of the contents unaltered in terms of how much a user can learn or benefit from other people impressions or from the activity that the surrounds the original creator's readymade idea. In this paper we are interested in exploring the question of reusing these readymade contents by turning them into a focal element of a larger production such as placing them in an interactive system that combines fixed contents with user interaction, creating a framework that actually encourages more active and sustainable user engagement. Using the term of "canned contents" to denote these readymade media, we consider ways of creating open systems that provide some combination of interoperability, portability and open

standards for applying these artifacts to creation of new cultural experiences. This also bears resemblance to open system of learning, where information is sourced from multiple sources to promote active learning that focuses the responsibility of learning on learners. There are several models of instruction that promoting discovery learning by encouraging learners to be engaged during learning and cognitively active.

In applying these ideas to entertainment in general and the arts and music in particular we face with several unique problems. On practical level there is a certain amount of professionalism, both in terms of the technical skills and artistic expression that is required for producing such contents that might be prohibitive to most of the users. One cannot expect large audiences to exhibit the same amount of control and awareness of the impact of media manipulations as someone who had the time and possibility to perfect those details. Moreover, we want to open the readymade contents to include expressions by audiences that might result in shift of attention from the contents themselves to what other audiences think about them. This shift towards social media is current with the burgeoning trend to segment the fans base for artistic creation into smaller groups rather than maintain a central sense of accepted quality and an overall established taste. But rather than considering practicalities, we would like to draw the reader's attention also to the broader questions of the significance and cultural role of public interaction with canned contents. We will show that interaction through commentary and other exegetic practices puts the audience in the role of "significance givers", a practice that had been already a dominant force in shaping cultural and religious thought in early scholastic practices.

2 The Binary Computer and "Blind Thought"

The current success of the computer as a universal information-processing machine essentially lies in the fact that there exists a universal language in which many different kinds of information (text, images, music, videos, etc) could be encoded and that this language could be fully mechanized. Computers are artifacts aimed at storing and manipulating information encoded in various ways, information being basically anything that could be algorithmically generated. In 1697, the German philosopher and mathematician G. W. Leibniz invented binary notation and considered his discovery to be *imago creationis*, that is, at the image of the Creation, see [3] for details. He used the motto *unus ex nihilo omnia*: "from nothing, the One creates everything"... Indeed, with the numbers zero and one only, all others could be constructed and therefore the whole universe of numbers, in the same way as, in Christian theology, God created the world from nothingness. This idea is perfectly at work in current computer entertainment systems and videogames, where complex virtual universes are built from simple (but long) strings of 0's and 1's. Therefore, digital computer technologies have been able to digest the whole world little by little, dimension after dimension, reifying indeed Leibniz' dream. Binary notation thus became the Universal Language to describe non only the linearity of numbers and texts, but also the flatland of images and, moreover, virtual environments in 3D.

The power of the computer lies in its ability to blindly manipulate symbols representing any encoded information; information that will be transmitted,

reinterpreted and decoded later to become humanly intelligible. Indeed, this idea of performing, within a given symbol system, operations on symbolic objects in a purely syntactical manner, without reference to their usual meaning or *modus operandi* (for instance because their meaning is unknown) has also been proposed by Leibniz in the seventeenth century as the basis of combinatorics, and he named it the *cogitatio coeca*, that is, “blind thought”. This reasoning process thus operates on atomic signs (that could not be further interpreted), as in mathematical algebra when the resolution of problem is conducted by a computation involving operations on symbolic variables and arithmetic terms, without any idea of what the variables actually refer to (e.g. in equation solving). Leibniz will use this paradigm in mathematics and combinatorial systems as an *ars inventendi* (art of inventing), and for laying the grounds of modern semiotic systems, but the roots of these concepts has to be dug out further back in the *ars magna* (great art) of the Franciscan monk Ramon Llull (1232 - 1316). Interestingly, some authors relate lullian systems to the Jewish Kabbalah, as the use of combinatorics is at the core of several processes in the Kabbalah, for instance in the ecstatic Kabbalah of Abraham Abulafia (1240-1291) [10]. Anyway, Llull’s ground-breaking work was highly praised by many medieval and renaissance philosophers and there was an abundance of lullian treaties written up to the seventeenth century, in particular by Giordano Bruno (1548 – 1600) [12]. The importance Bruno gave to the syntactic form of signs over the semantics is clearly stated in the *De lampade combinatoria lulliana* (1586): “One should not pay attention to the properties of terms but only to the fact that they define an order, a texture, an architecture” [8]. For further details linking these pioneers to the birth of modern logic at the turn of the 20th century and, see [2].

3 Virtual Worlds, Interaction and Immersion

Computers are massively widespread in the contemporary society and are taking major part in our everyday life. They are certainly, in a subtle but nevertheless very concrete way, developing new cognitive abilities which open the public to a different perspective on the world. Portable wireless handheld devices and GPS-enabled smart objects let us foresee the so-called ubiquitous networking society that will soon be ours, surrounding us with a data-gas of information-filled radio waves, a new ecosystem to be decoded and understood by digital machines only. However let us first look at the last decade of the 20th century and the development of 3D virtual worlds, which popularized computer entertainment and video games and express an important paradigm shift, which is not without resonance with what was happening in contemporary art at the same time.

From the forgotten pioneering work of Morton Heilig (1926-1997) who patented, back in 1960, a device composed of two small TV monitors placed in front of one’s eyes and Head Mounted Displays (HMDs), created at NASA in 1984 following the original ideas of Ivan Sutherland back in 1968 to the CAVE systems [4], an immersive system created by the EVL group at the University of Illinois in 1992, and today’s advanced systems incorporating haptic devices, Virtual Reality technologies are getting closer and closer to the fantasy dream of the “Ultimate Display” [15]. The

scientific experiments or artistic installations developed by using immersive 3D technology compel the spectator to enter within a simulated world instead of just watching it from the outside, as a mere series of photographs, a predefined movie or even a monitor display. On the small screen and in the domain of computer entertainment, the revolution of 3D games with a first-person point of view and subjective camera happened at the same period, with games such as *Wolfenstein 3D* (1992) and the magnificent *Doom* (1993). Despite an incredibly simple scenario (« hunt-and-kill ») and rudimentary graphics, the immersive effect was totally operational, maybe even too much for certain people, as the gamer was totally engaged - mentally if not physically - in the virtual universe. Then some innovative game design tricks appeared which would further enhance immersion, in particular the possibility to not only wander in the virtual space but also to interact and use objects. This was introduced by *Half-Life* (1998) through the use of the « E » key of the keyboard: by using this key and pointing on certain objects, the gamer could trigger some predefined action, such as calling an elevator, opening a valve, or even drive a railway bucket. This “manipulability” of the world, although extremely simple as each object contained its own teleology, induced nevertheless a deep feeling of immersion and of coherence in the virtual world. This was further achieved in games like the superb *Metal Gear Solid* (1999) and *Deus Ex* (2000) where clicking on an object could lead to some complex manipulation : virtual computers could connect to Internet and let you read your email (or hack someone else’s) and virtual cash could be drawn from ATMs ... Thanks to the increasing power of CPUs and GPUs, such interactivity has been vastly improved nowadays, with games incorporating impressive physics-based simulation engines making it possible to manipulate and use objects in a very realistic manner, cf. e.g. *Half-life 2* (2004) or *Metal Gear Solid 4* (2008).

Interestingly, immersive virtual worlds and 3D RFPS games amount to a shift from the cartographic paradigm, found for instance in board games such as *Go*, *Chess* or so-called God-games (*Sim City*, *the Sims*, *Back and White*, etc) to enter the ichnographic paradigm, that is, based on the notion of path. We move away from the metaphor of the map to the metaphor of the path, we move from the third person point of view (God’s eye) to the first-person point of view (our eye). As Morpheus said to Neo in the blockbuster movie *The Matrix* (1999), “There is a difference between knowing the path and walking the path”. We are thus leaving, in the virtual experience and exploration of an unknown artificial world, the Cartesian paradigm of the Euclidean, homogeneous and objective space in which points could be described in an allo-centric manner by coordinates (x, y, z) for a new paradigm of a more constructive, ego-centric, and indeed subjective space. This is somehow backtracking to the ideas of the French mathematician Henri Poincare, founder of modern topology, who considered that a point in space should be described by the transformation that has to be applied to reach it. Hence space is represented as a set of situated actions. No one knows the totality of the map; no one can picture or order the territory in any comprehensive way, even abstractly. The complexity of the structure (“graph” to follow the word of Michel Serres, “rhizome” for Deleuze, or “network” in the contemporary terminology) cannot be conceptually apprehended nor depicted. It is worth noticing that, by moving from a 2D to a 3D model, some information is in some way lost. Lost first is the idea of a comprehensive, full-blown representation –

such as in the map. This is precisely what happens also with the Word Wide Web: no one is able to draw the whole map of the Web, maybe only can we store it on an army of data centers ... We can only explore the Web by following paths, references from one location to another and performing actions, which will open to new paths and possibilities, and so on so forth ...

4 Active Cognition

A key point in immersive environments and computer entertainment is the place of the spectator into the system and not outside, as can be found in normal computer-human interaction. The ability of games to lead to deep integration of the player into the gameplay and to therefore amount to cognitive immersion, by letting him perform intuitive and meaningful *actions*, is certainly the main novelty of this genre. In the paradigm of the interface, interaction has often been considered as a necessarily reduced and incomplete dialogue between the human and the machine, as a means of access that will always be frustrating since it is always limited and imperfect. Immersive virtual environments and now pervasive technologies are doing their utmost to devise interfaces that are more or less natural to allow for an improved interaction with the digital systems. However, one should not forget that immersion is cognitive before being perceptive, the virtual “reality” is clearly being re-invented and re-created by the viewer and not just perceived and undergone. As Marcel Duchamp put it in 1914, referring to his well-known Readymade artworks : “the viewers make the artwork” (in French: « *Ce sont les regardeurs qui font l'oeuvre* ») [7]. By performing actions in a virtual world, or by interpreting (i.e. creating the meaning of) a conceptual artwork, we are engaged in *active cognition*: we learn by doing. In the cognitive philosophy of [11,16], one key idea is the contingency of cognition: “living systems are cognitive systems, and living as a process is a process of cognition.” [10]. Varela proposed the concept of enaction: in the enactive paradigm, the subject is always conceived as situated and interacting with his environment. Cognition is thus inseparable from the experience of the world, cognition is conceived as an embodied action. We should also remember, in the same line of thought, the words of the cybernetics pioneer Heinz von Foerster: “If you desire to see, learn how to act” [17]. For him, “perceiving is making” and all perception is therefore created by the subject’s action upon his environment; perception is active. Experiments have shown that sensory organs (in animals and humans) can be trained to better perceive expected signals before the brain considers them. Therefore, by analogy, it would not be unreasonable to think that a key issue in understanding experiences in virtual worlds would be the ability to perform actions and observe their consequences in order to learn the rules governing the artificial environment – maybe simply by trial and error. This is obviously easier to do in a virtual world than in the real one, and this cognitive process is therefore put to use in many computer games and now intuitively performed by video-game-educated kids. It might be possible that this ability to develop cognition by action is indeed gradually replacing the more classical humanist tradition of learning by books and letters.

5 Exegeses and the Role of a Commentator

In his 1945 article [1], Vannevar Bush described a futuristic device “in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility... When the user is building a trail, he names it, inserts the name in his code book, and taps it out on his keyboard. Before him are the two items to be joined, projected onto adjacent viewing positions. At the bottom of each there are a number of blank code spaces, and a pointer is set to indicate one of these on each item. The user taps a single key, and the items are permanently joined”. The interesting thing about this device is that Bush envisioned a mechanistic method of combining memory trails that only requires a single click. The enactment is limited to some earlier recording of data into the computer that is personally meaningful, while the operation of extracting or creating new meaning is assumed to arise from manipulation of the syntactic form that this data bears, which could be “blindly” handled by a machine. With the advance of pattern matching and machine learning technologies, indeed some progress was made in terms of automatic recombination of cultural artifacts, such as imitation of a personal musical style through Leibnizian-like combinatorics [5]. The advance of artificial intelligence technology actually extended the concept of atomic signs to codebooks of arbitrary size that could be dynamically acquired from observing redundancies and collecting statistics from prerecorded digital artifacts. This mechanical recombination symbolizes the step of moving beyond replication of real life environments through virtual representations of physical objects, to representation of semantics that is encoded through some careful parsing of the same universal binary representation, but now relative to a specific domain, personal memory trail or statistically derived code book.

The act of joining two or more data records into a combined meaningful whole puts an exegetic pressure on the process of combining this data. In early religious traditions we witness the need for reconciliatory techniques that arose from the need to harmonize collections of conflicting canonical texts, generating abstract pantheons of gods, monotheistic deities or cosmological principles. Over many centuries Neo-Platonic, Neo-Confucian, Vedic and Talmudic traditions employed commentarial techniques to show unity in disjoint collections of texts, with increasingly complex and hierarchical visions of reality being constructed through hermeneutical approaches to scripture. Moreover, it is claimed that exegetical techniques are largely culturally invariant, with differences of producing monotheistic deities or linear time versus metaphysical principles or cyclical models of time being dependent on the text they were operating on [9].

6 The Opera of Meaning

The idea of opera of meaning (OoM) is combining the syntactic, “blind” manipulations of the universal binary language of the readymade presentation with enactment of a viewer / user, whose actions serve to validate the cognitive relevance of these digital artifacts. Implemented on the web, OoM employs a variety of

exegetic techniques to direct a group of viewers in synchronized or semi-synchronized way towards more active perception and cognizance of the meaning of the “canned” story, while at the same time using the audience expressions to expand the scope of the presentation contents. Designed originally as a format of presentation of media contents, such as dramatic work, entertainment or lecture, it uses a realization of a script to push to the viewers a collection of electronic objects for interacting with the story meaning. Some of these object are created “on the fly” based on structural analysis of the canned contents and/or analytics of the audience inputs. As all art finally happens in the heads of the audience, the system allow to tap at “the silent conversation” that follows any concert or show, which is an essential element missed by a lot of interactive art. Through processes of linking, association, abstraction, correlation, syncretic fusion and mashup, the semantic scope of the presentation is constantly expanded and contracted, turning a passive viewing into a dynamic scenario of negotiating and contextualizing the meaning of the canned contents according to suggestions of the audience, controlled or orchestrated by a moderator. Exegetic techniques employed in OoM included textual commenting and threaded responses, voting through polls, ranking, image selection, linking and searching, text analytics, sentiment and topic analysis, summarizing, visualization and more. In principle, any web service that provides search, analytics and data mining, can be included in the presentation and triggered from the script at an appropriate moment. The results of these processes are then summarized and shared with the rest of the audience in parallel with the main show.

The following figure shows the elements of OoM, which comprise of several different modalities that keep evolving as we more scenarios to experiment with the system are developed. The main elements in the system are the canned contents, the audience back channel, the moderator function and the shared canvas. In the figure below the presence of the “canned” contents is denoted as “presentation”. The presentation can be a visual, audio or textural rendering of a story that the audiences will perceive. This may be combined with a script that contains structural representation of the story, such as metadata or other descriptors of the canned elements. This information is essential in order to provide more knowledge to the moderator function, as will be explained below, as well as to automate the push of related contents and interactive widgets. In case this information is missing, automatic analysis, such as media semantic processing can be applied to the presentation to try to provide a model according to which the data from the users can be analyzed. For instance, in an ongoing project we use correspondence analysis of the text of a screenplay [12] to create a set of filters or rankings on audience comments “in the context” of the story narrative. In another project [6], association of the user input to the story line is proposed by creation of a metadata space in Mathematical Model of Meaning.

Once the relation between the audience input and the story are established, the communication are bound to shared canvas through a moderator, whose role as an information processing element is to negotiate and prioritize the display of user inputs, possibly by compact visualization or some other summarization process. The additional element of “artificial participant” suggests a possibility for a direct

mechanistic intervention of a computational process in the user back channel without intervention of the moderator, as will be explained later.

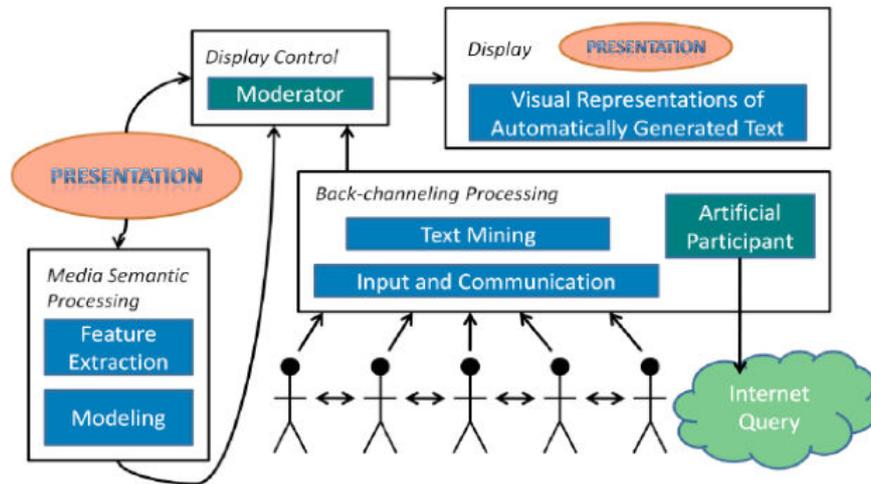


Figure 1. System diagram describing the different OoM components

In terms of functional description, one of the central features of Opera of Meaning is that the public can interact with related contents during the story presentation either privately, or by sharing it with other audiences by posting to designate publicly viewed display areas. The shared canvas provides a way for rendering the user communication in the "back channel" in parallel to watching the main story. Using associative search or semantic database management, additional contents that are relevant to the main story are made available to the users. These contents can be shared at predetermined times on private or shared canvas, depending on their purpose and available screen real estate. It should be also noted that the audience can not directly interact or modify the canned contents of the main presentation, except for cases where a specific role for the shared canvas had been integrated in the design of the main show. Examples of such integration could be use of shared canvas as a back drop in theatrical presentation as part of the set design, or including debate and media sharing episodes intermittently with the canned contents as part of a visual montage, or using audience input to create the remix or collage on the fly, and so on.

Another central role is that of a moderator who "listens" to the input from the presentation and the audience back channel and reports back its results on the shared canvas display, in parallel or in sequence with the canned presentation. To give a concrete example, the moderator can summarize the ongoing chat in the audience and post it to the shared display as a tag cloud. Moderator can also create a tag cloud from the text that accompanies the main presentation. In such a case the tag cloud representation might combine both sources of information – that of the main story and that from the audience. More sophisticated application that we currently explore are automatic creation of web mashups that are derived from topic analysis of the back-

channel, through automatic search and keyword expansion “in context” of the semantics of the canned presentation.

In many cases, asking the viewers to post a comment might be the most difficult part in getting the discussion rolling. The role of the artificial participant is to “break the ice” by providing an initial input regarding the canned contents. An important distinction between the artificial participant and the moderator is that the moderator “summarizes” the overall activity that goes on in the back channel and uses that to alter or augment the presentation, while the artificial participant communicates privately with audience members, either through textual comments or graphical widgets, triggering responses and provoking participation.

An example of the screens and comments generated by OoM are given in Figure 2. This is one of multiple pages that were running during a panel session in the Theater Department in UCSD on activism in the arts. The central image represents a theatrical group called “Guerilla Gilrs”, and it was brought up to demonstrate an example of feminist art activism, embedded from the web after searching for contents related to the topic of the main presentation [art activism, theatre, feminism]. The lower half of the screen shows the audience chat on the right and a dynamic text cloud on the left. Each one of the display elements comprises a separate object in the OoM system.



Figure 2. An example screen of OoM for an event regarding Art Activism

Other functions of the system include administrator functions for manual creation of web pages of the related contents, tools for creating polls and a control panel for arranging the different performance objects both in terms of their layout on the page and timing their appearance in synch with the main show.

As last figure we present a photo of a deployment of such system in a live concert situation. In this piece the deployment of the system was integrated with the design of the projection space in such a way that the displays used for related contents and

shared canvas actually surrounded the audience creating an immersive experience. The main presentation is the performance in the middle, while the audiences with laptops interact with each other and with the system [13].



Figure 3. Use of OoM as a backdrop and immersive environment in a mixed music and media story telling show with audience participation

6 Conclusion

In this paper we presented a system for public interaction during presentations of non-interactive readymade (canned) materials such as films or pre-composed performances. Using concepts of user back-channeling, moderator and shared canvas, the system allows a significant part of an art piece that is actually happening “in the heads” of the audience to be shared in parallel or in sequential manner with the main show. Sharing audience impressions not only contributes to the experience and understanding of the work, but actually creates a new role for the viewer as a second person “significance giver”. We discussed how this type of engagement changes the role of the spectator from being either third or first person, or being either a passive viewer or an direct operator of digital contents, into a second person role whose operative is one of a establishing the meaning of an artwork through a process negotiating with other viewers or addressing other associated contents. The setup allows mixed attendance modes that combine first person interaction among attendees directly or through a moderator, and a third person mode of viewing with regards to the main story, with its associated freedom of imagination and deliberation. This mixed operational mode has both advantages and disadvantages.

We also discussed the broader questions of the cultural role of public interaction with canned contents and showed that interaction through commentary and other exegetic

techniques can become a dominant force in shaping cultural thought, such as the case with religious thought in early scholastic practices. The name “Opera of Meaning” is intended to emphasize the interplay of meanings and exchange of opinions that occur during performance of a story. This approach is related to works that explore social and educational roles of film and theater media, and is inspired by methods of debate and commentary that are common in traditional religious studios situations, as mentioned above.

References

1. V. Bush, "As We May Think", in *Atlantic Monthly*, July 1945.
2. P. Codognet, "Combinatorics, Randomness, and the Art of Invention", In: *New Computational Paradigms for Computer Music*, G. Assayag and A. Gerzso (Eds.), IRCAM / Editions Delatour, 2009.
3. P. Codognet, "Ancient Images and New Technologies: the Semiotics of the Web", *Leonardo*, vol. 35, no. 1, Jan. 2002, MIT Press.
4. C. Cruz-Neira, D.J.Sandin, T.A. DeFanti, "Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE", In: *proceedings of SIGGRAPH '93*, ACM Press 1993, pp. 135-142.
5. S. Dubnov and G. Assayag, "Universal Prediction Applied to Music Generation with Style", In: *Mathematics and Music*, Springer-Verlag, 2002.
6. S. Dubnov, Y. Kiyoki, "Opera of Meaning: film and music performance with semantic associative search", In: *proceedings of Frontiers in Artificial Intelligence and Applications, Information Modelling and Knowledge Bases XX*, Volume 190, pp. 384 – 391, 2009.
7. Marcel Duchamp, *Marchand du sel*, texts collected and edited by M. Sanouillet, Editions Le Terrain Vague, Paris, 1959.
8. U. Eco, *The Search for the Perfect Language*, Wiley-Blackwell 1997.
9. S. Farmer, J.B.Henderson, and P. Robinson, "Commentary Traditions and the Evolution of Premodern Religious and Philosophical Systems: A Cross-Cultural Model", *Kolloquium zu historischen und methodologischen Aspekten der Kommentierung von Text*, University of Heidelberg, July 1997.
10. M. Idel, *The mystical experience in Abraham Abulafia*, State University of New York Press, 1988.
11. H. Maturana and F. Varela, *Autopoiesis and Cognition: The realization of the living*, Reidel, Dordrecht, 1980.
12. F. Murtagh, A. Ganz and S. McKie, "The structure of narrative: the case of film scripts", *Pattern Recognition*, 42, 302--312, 2009.
See discussion: Z. Merali, "Here's looking at you, kid", *Nature*, 453, p. 708, 4 June 2008.
13. D. Ramsey and D. Sutro, "An Opera of Meaning Integrates Live Performance, Internet, Multimedia and Audience Participation at UC San Diego", UCSD News, February 2008.
14. J. Shearman, *Only Connect...: Art and the Spectator in the Italian Renaissance*, Princeton University Press 1994.
15. I. E. Sutherland, "The Ultimate Display", In: *proceedings of IFIP congress*, 1965, pp. 506-508.
16. F. Varela, E. Thompson, and E. Rosch, *The Embodied Mind: Cognitive science and human experience*, The MIT Press, Cambridge, MA, 1991.
17. H. von Foerster, *Observing Systems*, Intersystems Publications, Salinas, Calif., 1984.
18. F. Yates, "Lull and Bruno", in *collected essays I*, Routledge and Kegan Paul, London, 1982.