Chapter 55
Evolution of the Semantic Web Towards the Intelligent Web: From Conceptualization to Personalization of Contents

Blanca Piñeiro Torres and Aurora García González

Abstract While Web 2.0 meant the raising of social media, the semantic dimension of Web 3.0 targets artificial intelligence, operating a method based on ontology to classify webpages, which allows users to find and understand information. Search engines and computers control the process, collecting knowledge from profiles and interactivity of users.

Keywords Artificial intelligence · Ontology · Conceptualization · Personalization · Semantic web · Syntactic web

55.1 Theoretical Framework

55.1.1 Static Web 1.0

The initial version of the website was the static and unidirectional Web 1.0, which was created to provide information about corporate organizations, news or specific content of diverse subjects. Due to the high costs of web publishing, updating was performed infrequently, so the aim of Web 1.0 consisted of a presential strategy on the Internet.

In the new concept of web 2.0 as an interactive online community (Cebrián Herreros 2008), tools like content syndication, messaging services, forums, weblogs, wikis and social networks permit open communication to users.

B. Piñeiro Torres (✉) · A. García González
University of Vigo (UVigo), Vigo, Spain
e-mail: blancapineiro@hotmail.com

A. García González
e-mail: auroragg@uvigo.es

F.C. Freire et al. (eds.), *Media and Metamedia Management*,
Advances in Intelligent Systems and Computing 503,
DOI 10.1007/978-3-319-46068-0_55

55.1.2 Syntactic Web 2.0

The origin of the social web 2.0 is a collaborative work on Internet written by Levine et al. (1999) to examine the impact of the Internet, in both organizations and consumers. Their communication achieved the support of more than a thousand cosigners, including businessmen and Internet users. To define how companies speak with customers through the Internet, these authors published the Cluetrain Manifesto in 2000, consisting of 95 theses.[1]

In 2001, the crisis of digital companies also contributes to the emergence of Web 2.0. However, it was not until 2005 when Tim O'Reilly, founder and CEO of O'Reilly Media, provides theoretical foundation to define the concept in his article "What is Web 2.0" (O'Reilly 2005), which explains the evolution happening in internet: web pages become bidirectional platforms, where the receiver mediates in the network of communication, with the ability to share content, comment and collaborate (Fig. 55.1).

O'Reilly and Battle reviewed the basics of web 2.0 and after five years they redefined collective intelligence, reinforcing the value of communities in the construction of the online media. The Internet collective intelligence depends on the management, understanding and responding to a massive amount of user-generated data in real time.

From a conceptual revision, O'Reilly and Battle announce a transition to the semantic, social, mobile web and virtual reality:

> Is it the semantic web? The sentient web? Is it the social web? The mobile web? Is it some form of virtual reality?
>
> It is all of those, and more (O'Reilly and Battle 2010).

55.1.3 Web 3.0

From a basis on syntaxis and connection between users, Web 2.0 evolves into the semantic web or web 3.0, where a deep understanding is allowed and the user experience is developed by obtaining more accurate results in content search, introducing new applications, such as geolocation or biometrics, and developing artificial intelligence technologies. Berners-Lee (1997) predicted that machines

[1]These conclusions were organized in 8 sections: markets as conversations (thesis 1–6), the hyperlink as subversion of the hierarchy (thesis 7), the connection between new markets and companies (thesis 8–13), the entry in the market organizations (thesis 14–25), the marketing and response organizations (thesis 26–40), the impact of the intranet in the control and structure of organizations (thesis 41–52), the connection between the market for Internet and corporate intranets (thesis 53–71) and new market expectations (thesis 72–95).

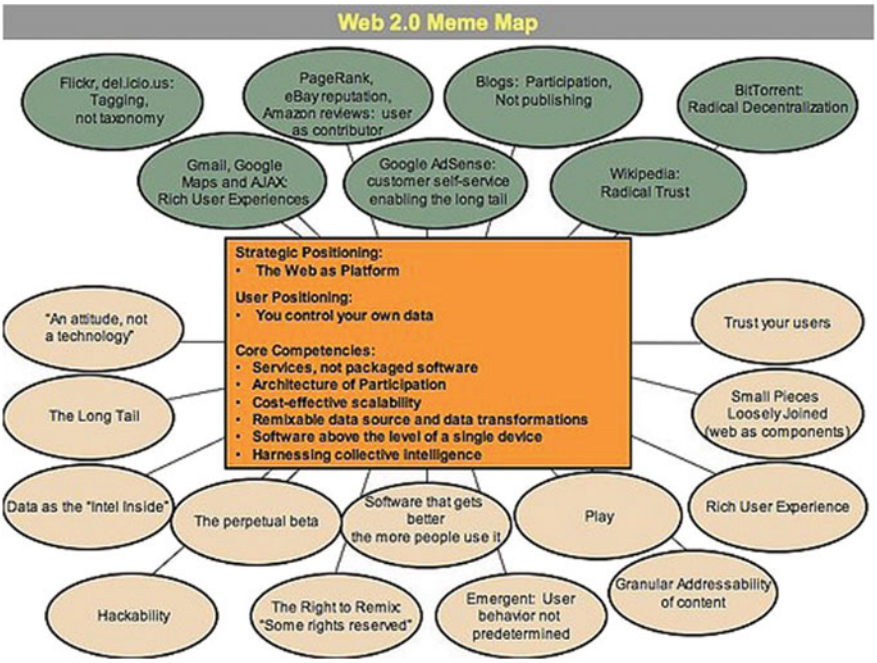

Fig. 55.1 Original map of the web 2.0. *Source* http://www.oreilly.com/

would be able to read web pages as easily as humans in his book 'Weaving the Web'.

Web 3.0 is based on artificial intelligence techniques using natural language in their searches, based on data mining, machine learning and attendance software agents to convert information into collective knowledge. O'Reilly and Battle (2010) explain that the inferential web 2.0 learning method is based on the explicit meaning of the processed data.

Thus, the semantic web uses language to find, share and integrate information. Among the advances in Web 3.0 are cross data or big data analysis, the use of 3D technology, based on the three-dimensional design, augmented reality, facial recognition in images, telematic applications for health monitoring or geolocation systems and triangulation used in mobile telephony.

Semantic web site keeps syntactic principles: decentralization, sharing, compatibility, accessibility and content contribution. But in addition, the semantic web ontology releases the notion of artificial intelligence.

According to Gruber (1993), the definition of ontology used in artificial intelligence application is "a formal, explicit specification of a shared conceptualization" (Gruber 1993: 199). In Gruber (1995), the author specifies the principles of ontology oriented towards knowledge sharing. In the conceptualization of ontology, the following aspects can be identified to understand it:

- Specification: description of characteristics and performance of concepts.
- Explicit: it provides a detailed explanation of these concepts.
- Formal: organization of terms and connections according to normalization.
- Concept: a set of concepts, described by their entities, attributes and relations.
- Sharing: common knowledge accepted by a group of users.

As Urrego-Giraldo and Giraldo Gómez (2005) explained, the ontology used to classify and index information filters out queries from users and facilitates their interaction with processors that can infer relevant information about their needs.

Metadata are required to organize and classify the information contained in the Internet. These texts are structured alphanumeric data representing bibliographic description of electronic resources and provide structured descriptions to locate objects.

55.1.4 Personalization of Contents for Users

Web 2.0 inspires the use of concepts as *emerecs* (emitters—receivers) or *prosumers* (producers—consumers). Both ideas refer to the active involvement of the receiver in the communication process, while that of *emerec* (*emetteur—être—recepteur*) was originally proposed by Jean Cloutier[2] in his '*Petit Traité de communication*' in 1973 and revised to analyze the effects of technology on the processes of human communication (Cloutier 2001).

Fernández Castrillo (2014) recognizes the increasing involvement of users, who called prosumers in content creation UGC or user generated content and the combined use of internet media, which contextualized storytelling in the transmedia, reasoning on the research conducted by Jenkins (2009) at the Massachusetts Institute of Technology, as a process in which the elements of fiction are systematically disseminated through multiple distribution channels in order to create a unique entertainment experience.

55.1.5 The Concept of Cibermedia

The proposal of *Cibermedia* is based on the research group Novos Medios (López García et al. 2003: 40) and refers to the online media using the journalistic

[2]Cloutier's own work has a hypertext structure, composed of six modules (Guylaine Martel 2014), which deal with: (1) the universal communication, (2) communication functions—to inform, educate, encourage and entertain-, (3) *emerec* description (its characteristics, influences, media, resources and technical applications used), (4) the media, languages and messages in the new technologies of information and communication (5) the size of the cyberworld (state, virtuality and interactivity) and (6) the law of the three thirds, favoring the balance of communication, due to a tandem between the *emerec*, the other participants and their status.

techniques, the multimedia language, interactivity and hypertext to update information on the Internet.

From this concept, several perspectives arise:

- The first point of view focusing on the media, with their own editorial, narrative and discursive structures.
- The perspective of the subject who broadcasts informative, commercial and entertainment communication or infomediation.
- The public, which provides a personalized and individualized communication in which "the user leaves the passive role that often plays in the mass media, becoming active (and interactive) subject" (Díaz Noci and Salaverría 2003, p. 41).
- The professional criteria, structure, writing and ethics relating to journalism.
- The hypertext, multimedia technology and interactivity.
- The update.

This definition of online media has recently been reviewed and concretized by Díaz Noci and others, in order to provide a methodology for studying digital media or online media: "A medium that uses a digital interactive online platform in the form of web site or as a mobile web app" (Codina et al. 2014).

Díaz Arias (2009) distinguishes between digital media (which refers to newspapers and magazines) and online media, since the source of the content is directly published on the Internet, the diffusion and the interaction occur in the context of Cyberspace (Díaz Noci et al. 2014).

While characterizing the online media, López García et al. (2003) detected the degree of dynamism of cibermedia, which determines four dimesions: hypertextuality, multimedia, interactivity and refresh rate. With the new version of semantics, this academic proposal adds the personalization of content, which considerates the user profile and preferences.

In content writing for online media, Pavlick (2005) explains that adaptation to the Internet has led to the emergence of contextualized journalism, which meets five conditions: communication modes; hipermediality; increasing audience participation; dynamic content; and customization.

According to Mayoral Sánchez and Edo Bolós (2014), terms such as hypertext, interactive or multimedia have been employed often to characterize cyberjournalism.

55.2 Methodology

The method applied for this research is qualitative content analysis, using as sources academic publications in electronic scientific journals, academic literature and Internet fonts.

55.3 Results

The purpose of the present research is to update the evolution of online media during the period between 2010 and 2015, which corresponds with the arrival and evolution of the semantic web.

Díaz Arias (2009) finds the distinction in the historical development of the online media, for which the author classifies three stages: during the first five years newspapers were the first generation of network information, with minimal input, a second phase in which newspapers, magazines, radio and television create their own Internet publication for professionals who use technical language and multimedia resources, and a third phase that gives prominence to the audiovisual and interactive content, media convergence and online media. The collaborative, dialogic and interactive nature of Web 2.0 opens the way to a new vision of journalism, as Díaz Arias recognizes: the right to freedom of expression and information, namely the right to seek, receive and disseminate ideas, opinions and information. But in the media society, the individual is a consumer of the information. Therefore, access to relevant public information allows him to become a citizen (Díaz Arias 2009: 6).

The web 2.0 users can be editors, customizing their access to information, and community broadcasters, assessing, discussing and sharing content published by consulting sources. In addition, they can create their own content on social networks and blogs.

Many theorists see citizen journalism as an opportunity on global integration, rather than a threat. According to Rheingold (2002), this collective intelligence finds resources in social networks and aims to produce user-generated content (UGC).

The growing influence of social networks has been undertaken by public and private agencies. The user-generated content have recently been integrated by institutions and considered by the political parties in their campaigns, since they consider social participation through networks such as Facebook, Twitter or Youtube to prepare their agendas and electoral programs.[3]

At this level, it is required to introduce the category of digital native cibermedia classified by Salaverría and Negredo (2013).

If the advent of computers and internet represented a paradigm shift with the transition from the mainstream media to online media, the expansion of mobile terminals and tablets also involves changes in the design, editing and digital content consumption, which adapt to portability and multiple screens, according to Aguado and Castellet (2013).

In the same way that the online media had evolved from adaptation of content to the creation of multimedia, hypertext and interactive online creation, the evolution

[3]The main political parties that run for the general elections on 20 December 2015 enhanced the participation of their followers in social networks for making electoral programs and electoral monitoring their activity.

of mobile devices and tablets has experienced two stages: at first, they only offered text messaging services (SMS) or multimedia (MMS), email and web browsing, but similar content to those offered in the web; the launch of the iPhone in 2007 and the advent of digital tablets led to the development of specific content more dynamic and adaptive design for ubiquitous touch devices (responsive design).

Meso Ayerdi et al. (2014) speak of mobile as the fourth screen and conclude that this new channel of communication poses new challenges for journalism to increase competitiveness: the development of narrative forms and adapted genres and effective formulas for participation of active and prosumer audiences. However, mobile applications favor the consolidation of brands, a new form of advertising revenue and user loyalty. Thus, they affirm that the future of online media goes through a strategic approach to multiplatform content.

55.4 Discussion and Conclusions

55.4.1 From Semantic to Intelligent Web

Technological advances in infrastructure and applications allow us to mention a new version of the web, known as intelligent web. The potential uses and applications of this new website are in an experimental phase.

55.4.2 Customization in Cibermedia

The development of applications of artificial intelligence and advances in mobile media and tablets promote the diversity of formats and customization of content in online media or cibermedia, which can improve usability, considering the user-generated content.

55.4.3 Coexistence of Diverse Web Generations

In the current Internet ecosystem, there are four coexistent generations of web (static, syntactic, semantic and intelligent), whose phases of development, characteristics and applications are specified in the Table 55.1.

Table 55.1 Evolution of the web to the artificial intelligence (1990–2015)

Version	Web type	Characteristics	Applications
Web 1.0 1990– 2000	Static	Static: The websites are unidirectional, with static elements and little update	Search engines Websites Databases Email Cibermedia: Digital edition
Web 2.0 2000– 2010	Syntactic	Dynamic: web pages create dynamic content extracting information from one or more databases Social: users contribute to content development. Changes in editing and publishing information procedures Example: Wikipedia, Youtube or Facebook	RSS Weblog Wikis Newsletters Newsgroups Forum SMS/MMS Social networks Cibermedia: Digital publications
Web 3.0 2005– 2020	Semantic	Based on the representation of meaning and connection of knowledge Add to semantic web metadata information by ontologies Collaborative creation Extends the interoperability of systems	Onthology Thesaur and taxonomy Semantic Search Semantic Web Semantic Weblog
Web 4.0 2015– 2030	Intelligent	Development of artificial intelligence Algorithms for natural language processing Semantic development communities Web applications can identify resources and manage them in the context of temporary or permanent sessions	Decentralized communities Artificial intelligence Intellectual property

Source Prepared by the authors

References

Aguado, J. M., & Castellet, A. (2013). Periodismo móvil e información ubicua. In J. M. Aguado, C. Feijóo & I. J. Martínez (Coords.), *La comunicación móvil. Hacia un nuevo ecosistema* digital (pp. 187–218). Barcelona: Gedisa.

Berners-Lee, T. (1997). *Realising the full potential of the web*. Retrieved October 28, 2015 from http://www.w3.org/1998/02/Potential.html

Cebrián Herreros, M. (2008). La Web 2.0 como red social de comunicación e información. *Estudios sobre el mensaje periodístico, 14*, 345–361.

Cloutier, J. (2001). *Petit traité de communication. EMEREC à l'heure des technologies numériques*», *Communication* (online), Vol. 22/1|2003, published 8 of January of 2014. Retrieved October 28, 2015 from http://communication.revues.org/4816

Codina, L., Pedraza, R., Díaz Noci, J., Rodríguez-Martínez, R., Pérez-Montoro, M., & Cavaller-Reyes, V. (2014). Sistema Articulado de Análisis de Cibermedios (SAAC): Una

propuesta sobre el qué y el cómo para estudiar medios de comunicación digitales. *Hipertext.net* (online), 2014, no 12. Retrieved October 28, 2015 from http://raco.cat/index.php/Hipertext/article/view/275560/364530.

Díaz Arias, R. (2009). *Avances y desafíos en los cibermedios.* In I Congreso Modelos Emergentes de Comunicación: de lo análogo a lo digital. Universidad Autónoma del Caribe. Barranquilla (Colombia) (11/5/2009).

Díaz Noci, J., & Salaverría, R. (Coord.) (2003). *Manual de redacción ciberperiodística*. Barcelona: Ariel.

Fernández Castrillo, C. (2014). Prácticas transmedia en la era del prosumidor: Hacia una definición del Contenido Generado por el Usuario (CGU). *Cuadernos de Información y Comunicación, 19*, 53–67.

Gruber, T. R. (1993). A translation approach to portable Ontologies. *Knowledge Acquisition, 5*(2), 199–220.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies, 43*(4–5, November 1995), 907–928.

Jenkins (2009). *Confronting the Challenges of Participatory Culture Media Education for the 21st Century*.The MIT Press Cambridge, Massachusetts.

Levine, R., Locke, C., Searls, D., & Weinberger, D. (1999). *Manifiesto cluetrain*. Retrieved September 13, 2015 from http://personal.us.es/mbmarquez/textos/cluetrain.pdf

López García, X., Limia Fernández, M., Isasi Varela, A., Pereira Fariña, X., Gago Mariño, M. & Calvo Diéguez, R. (2003). Tipología de los cibermedios. In R. Salaverría (Coord.), *Cibermedios. El impacto de internet en los medios de comunicación en España* (pp. 39–81). Sevilla: Comunicación Social Ediciones y Publicaciones.

Meso Ayerdi, K., Larrondo Ureta, A., Peña Fernández, S., & Rivero Santamarina, D. (2014). Audiencias activas en el ecosistema móvil. Análisis de las opciones de interacción de los usuarios en los cibermedios españoles a través de la web, los teléfonos móviles y las tabletas. *Hipertext.net [online], 2014. Núm. 12*. http://raco.cat/index.php/Hipertext/article/view/274309/364489

Mayoral Sánchez, J., & Edo Bolós, C. (2014). Evolución de la producción audiovisual en cinco cibermedios españoles. Fonseca, *Journal of Communication*, *9*(July-December of 2014), 233–262. ISSN: 2172-9077.

O'Reilly, T. (2005). *What is web 2.0. Design patterns and business models for the next generation of software*. Retrieved October 13, 2015 from http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html

O'Reilly, T., & Battle, J. (2010). *Web squared: Web 2.0 five years on*. Retrieved October 13, 2015 from http://www.web2summit.com/web2009/public/schedule/detail/10194

Pavlick, J. V. (2005). *El periodismo y los nuevos medios de comunicación*. Barcelona: Paidós Comunicación.

Rheingold, H. (2002). *Multitudes inteligentes: la próxima revolución social*. Barcelona: Gedisa.

Urrego-Giraldo, G., & Giraldo Gómez, G. (2005). La estructura de servicios y de objetos del dominio: Una aproximación al concepto de ontología. *Revista Tecno Lógicas*, 46–67. December of 2005.

Salaverría, R., & Negredo, S. (2013) Caracterización de los cibermedios nativos digitales. In M. A. Cabrera (Coord.), *Evolución de los cibermedios. De la convergencia digital a la distribución multiplataforma* (pp. 175–180). Madrid: Fragua.