Could semantic web and accessibility be BFF (best friends for ever) in image annotation?

A position paper submitted by

Mireia Ribera (ribera@ub.edu), Professor at the Library and Information Science Department, University of Barcelona, Spain

Bruno Splendiani (splendiani@ub.edu), PhD Researcher the at Library and Information Science Department, University of Barcelona, Spain

Our Perspective

The current web is full of images, and very often they constitute "extremely rich information sources" (Dasiopoulou et al. 2011) providing basic, essential information. Search systems that focus on (or at least include) images are gaining importance in the general web, as well as in the scientific communication field. Some publishers offer tools for searching figures inside academic papers, such as Springerimages (Springer 2012), which search figure captions and retrieve images from journals published by them. Databases like Sciencedirect allow searches for images by legend text and whole article text.

This valuable information poses limitations to its use by all kinds of users and by machines as well. Accessibility and Semantic web have fought this barrier following different approaches, and we think that both disciplines can benefit each other by joining efforts.

Accessibility tackles the "hidden" information of the image by the inclusion of an alternative textual description that must be associated to it. Common solutions such alt text, longdesc, caption or figure-caption techniques offer a mean to describe the semantic content of the image as a whole. Standard formats like Scalable Vector Graphics (SVG), additionally, can describe authorship details and associate text to a specific part of the image. In all cases, the textual alternatives expose the images to search engines, their implementation improves web pages' ranking (Thatcher, 2006) and Google is often referred as the most relevant blind user by webmasters.

On the other hand, the semantic web tackles images "hidden" information with semantic descriptions based on the Resource Description Framework (RDF), or with annotations with microdata, RDFa or Facebook OpenGraph, which could be, in the last two cases, region-specific. With these techniques images have associated metadata to provide information about authorship, technical details or semantic meaning. The objective in this case is to link the image to related content in the web.

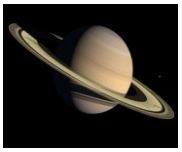
Our Viewpoint

In this position paper we will focus on annotations describing image semantic meaning, being the one which offers most similitude between accessibility and semantic web.

Some use cases that justify the use of annotations in images were described as soon as in 2007 in the report of W3C Incubator Group (W3C Incubator Group Report 2007), specifically the management of personal digital photo collections, cultural heritage, television archive, large-scale image collections at NASA and biomedical images. DIAGRAM recently added the education use-case where a teacher makes comments on an image (DIAGRAM Center 2014). ImageMagick (ImageMagick 2012), a well-known tool for image edition, illustrates another use case for annotation: to point out or highlight some aspect of the image. In all cases we could need whole-part annotations and the power to describe the image at different levels of detail or even give different descriptions to different audiences.

As examples of these use cases and possible descriptions, we managed to create brief annotations -- tags, in fact -- to a selected set of images in the NUSEF dataset (Ramanathan et al. 2010) (table I).

Table I Examples of image annotations according to different use cases

N.	Figure	Use case	Tag
1		Management of Personal Digital Photo Collections	Holidays 2012 (whole image) Vero - Jane (regions)
2		Cultural heritage	Buda stone sculpture (whole image)
3		Television Archive	Ronaldo, alone minute 30’ Best grass (2 tags, whole image)
4		Large-scale Image Collections at NASA	Detail of irregularity (region) Saturn rings (region) Saturn (region) (3 levels)
5		Biomedical Images	Born child with 2.50Kg (whole image)

6		Point out or highlight some aspect of the image	Lungless animal (whole image)
7		Educational Ebooks	Asian elephant (whole image)

As we can see, the applied tags could be relevant for semantic web and for accessibility as well.

Our proposal

In order to find the best solution for image annotation / description, existing techniques could be evaluated under some of the requirements proposed by Watson (Watson 2014) and Dasiopoulou (Dasiopoulou et al. 2011):

- Discoverability: the user can discover and access the description either by a call to action visible for sighted people or by programmatic access for non-sighted people (Watson).
- Structure: the description can have semantic structure within it (Watson).
- Granularity: the description can target the whole image or a part of it (Dasiopoulou).
- Use of ontologies: the linking will be richer if description uses an ontology (Dasiopoulou).

Furthermore, we suggest taking into account the simplicity and the current adoption of the techniques as requirements for the evaluation (table II).

Table II Techniques provided by the accessibility (A11y) and the Semantic Web (SW) disciplines and their conformance to the discoverability, structures, granularity (levels), use of vocabulary, simplicity and adoption requirements.

Technique	Discoverability (call to action)	Structure	Levels	Vocabulary	Simple	Adoption	Discipline
Alt	Programmatic (and visible[1])	No	One, whole image	No	High	Medium	A11y
Actual text[2]	Programmatic	No	One, whole image	No	High	Low	A11y
Caption	Programmatic and visible	Yes	One, whole image	No	Medium	High	A11y
Longdesc	Programmatic (and visible[3])	Yes	One, whole image	No	Medium	Low	A11y

[1] Some user agents provide ways to make alt attribute discoverable to sighted users https://chrome.google.com/webstore/detail/popup-image-alt-attribute/kkbofklgmmbcapfendjbjkjklajldnjj.
[2] For images included in PDF format.

RDFa	Programmatic and visible[4]	Yes	One or more, whole image or region-specific	Yes	Low	Low	SW
Microdata	Programmatic	Yes	One, whole image	Yes	Low	Low	SW
Facebook OpenGraph	Programmatic and visible	Yes	One or more, whole image or region-specific	Yes	High	High	SW

From our point of view, the accessibility community and the semantic web community can joint efforts to:

- Increase the number of described images. Semantic web could benefit from the compulsory use of alt text: the legislation of several countries demands to provide text equivalent for every non-text element in accessible products and services. And reversely, semantic web image descriptions, such as the annotations used in pictures within Facebook, could be translated into alt text.

- Increase discoverability of alt and actual text adopting web semantics mechanisms such as transcoding. This will also avoid bad practices with the title attribute.

- Enrich alternative descriptions within accessibility field to cover the more complex use-cases described in semantic web research.

Currently there is no perfect solution for image annotation. Our proposal aims to suggest developers in Accessibility and Semantic Web communities new ways of progress by sharing different visions and joining common criteria for improving a "universal access" to images by humans and machines.

Bibliography

Dasiopoulou, Stamatia, Eirini Giannakidou, Georgios Litos, Polyxeni Malasioti, and Yiannis Kompatsiaris. 2011. «A Survey of Semantic Image and Video Annotation Tools». In *Knowledge-Driven Multimedia Information Extraction and Ontology Evolution*, Georgios Paliouras, Constantine D. Spyropoulos, and George Tsatsaronis, 196-239. Lecture Notes in Computer Science 6050. Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-642-20795-2_8.

DIAGRAM Center. 2014. «Content Model». http://diagramcenter.org/standards-and-practices/content-model.html (Accessed March 4, 2014)

Imagemagick 2012 «ImageMagick v6 Examples - Annotating Images» http://www.imagemagick.org/Usage/annotating/ (Accessed March 4, 2014)

[3] Some user agents provide ways to make longdesc attribute discoverable to sighted users: http://www.d.umn.edu/~lcarlson/research/ld-ua.html (compatible with Opera 11 and 12)

[4] By transcoding (Mirri, Salomoni, y Prandi 2011)

Mirri, Silvia, Paola Salomoni, and Catia Prandi. 2011. «Augment Browsing and Standard Profiling for Enhancing Web Accessibility». In *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*, 5:1-5:10. W4A '11. New York, NY, USA: ACM. doi:10.1145/1969289.1969297. http://doi.acm.org/10.1145/1969289.1969297.

Ramanathan, Subramanian, Harish Katti, Nicu Sebe, Mohan Kankanhalli, and Tat-Seng Chua. 2010. «An Eye Fixation Database for Saliency Detection in Images». In *Computer Vision – ECCV 2010*, Kostas Daniilidis, Petros Maragos, and Nikos Paragios, 30-43. Lecture Notes in Computer Science 6314. Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-642-15561-1_3.

Springer. 2012. «SpringerImages». http://www.springerimages.com/. (Accessed March 4, 2014).

Thatcher, Jim, 2006. «Web accessibility: web standards and regulatory compliance» Berkeley, Calif.?] : FriendsofED ; New York : Distributed by Springer-Verlag.

W3C Incubator Group Report. 2007. «Image Annotation on the Semantic Web». http://www.w3.org/2005/Incubator/mmsem/XGR-image-annotation-20070814 (Accessed March 4, 2014).

Watson, Léonie. 2014. «The Tink Tank » Solving the longdesc problem». http://tink.co.uk/2013/03/solving-the-longdesc-problem/. (Accessed March 4, 2014)