

GLOCAL: Pro-am collaboration in the news production

Denis Teyssou
Medialab R&D manager

Agence France-Presse (AFP)
{[denis.teyssou](mailto:denis.teyssou@afp.com)}@afp.com
<http://www.afp.com>

Abstract. This paper presents the approach of the Glocal European funded project towards the co-production of news on World events, on an information marketplace involving both amateurs and professionals. It discusses the rise of *user-generated content* amongst worldwide media and how event modelling and technology usage may help to foster this pro-am collaboration. Glocal: Event-based Retrieval of Networked Media [1] is a large scale integrating and collaborative project which started in December 2009 for three years and involves ten partners.

Keywords: event, user-generated content, participatory journalism, citizen journalism, information marketplace, pro-am collaboration

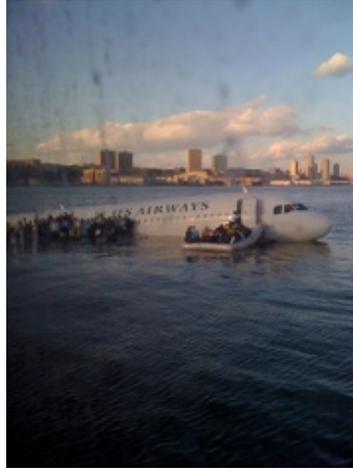
1 Introduction

"There's a plane in the Hudson. I'm on the ferry going to pick up the people. Crazy". This sentence posted with an iPhone on Twitpic [2] is by no means a breaking news according to journalistic standards but the picture which goes with it was definitively a stunning news alert.

It has even become an iconic photograph of user-generated content (UGC). It is now a symbol of how eyewitnesses, amateurs with a camera or a smartphone in their pocket can play a major role in the co-production of news, providing, like professional journalists, that they are in the right place at the right moment, to witness and report on a breaking news event.

This picture of the Hudson miracle sent on Twitpic on January 15th 2009 when the US Airways flight 1549 plane "landed" into the Hudson River in New York City (reproduced here below by courtesy of its author Janis Krums) is a kind of paradigm of this UGC uprising, an event with an happy end in a world of earthquakes, tsunamis, wars, crisis and diseases: a kind of exception to the assertion of Marshall McLuhan who once wrote that "news are bad news" [3].

From the London bombings in July 2005 to the Iranian post-election protests in June 2009, through the Sichuan earthquake in August 2008 in China and the Mumbai attacks in India in November of the same year, participatory journalism has provided raw material, valuable firsthand information and eyewitness reports to mainstream medias professional journalists around the World.



©Janis Krums

Fig. 1. Rescue of passengers of a US Airways flight 1549 in the Hudson River on the 15th of January 2009

Months after months, events after events, news content produced by amateurs continue to flow on Twitter, YouTube, Flickr, and other websites, including media. The public is not anymore a passive reader or listener and participates in the news process, most of the time by commenting, but also as producers.

And this trend is likely to last. According to a recent study from the American Pew Research Center, "Understanding the participatory news consumer" [4], up to 9% of American "internet users have contributed their own article, opinion piece, picture, or video to an online news site".

A figure which seems to be in line with the empirical 90-9-1 principle stating, according to Jakob Nielsen, that "in most online communities, 90% of users are lurkers who never contribute, 9% of users contribute a little, and 1% of users account for almost all the action" [5]. This user participation is also rapidly spreading with the use of smartphones: according to the Pew Research Center study, 40% of on-the-go news consumers on mobile device are willing to participate in online news creation.

"What is important here: Citizen comment and writing have upended the relationship between the professionals, journalists being paid by the Old Media, and the amateurs, the rest of us (...) It's a Pro-Am World (...) Amateur content is expanding rapidly", writes Ken Doctor, a renown media analyst from Outsell research in his recent book *Newsonomics* [6].

In the next section, we will review the rise of *participatory journalism* on news events reporting and in section 3 we will detail the Glocal approach.

2 Participatory journalism

While citizen journalism websites are often struggling to challenge professional media, as the State of the Media special report of 2010 on community journalism [7] shows, the pro-am collaboration in participatory journalism is changing the way professionals are reporting about the events shaping our World.

We understand here the notion of "participatory journalism" not only as "the initiatives undertaken by mainstream media to enhance the integration of all kinds of user contributions in the making of news" [8] but also as a way to build a community around news production, bridging the gap between amateurs and professionals in the field of journalism.

Many mainstream media, from CNN to Fox, through MSNBC, ABC, NPR and Broadcast Interactive Media, have established channels to collect UGC. In Europe, TVs like the BBC UGC hub, France 24 Observers, free dailies such as 20 minutes and Metro are brands using UGC as well as Bild in Germany, Le Parisien in France, The Guardian in England, just to name a few of them.

CNN's iReport, launched in 2006, have more than 100,000 citizen journalists registered and has received more than 300,000 reports, with more than 10,000 per month, 1000 of them being displayed on TV [6].

Some startups have also emerged around UGC as information marketplace to collect that type of content and to offer it to mainstream media. American Demotix and French company Citizenside.com [9] are such companies which have created a two sided-market platform gathering the content and managing communities of several thousands of independant and amateur photojournalists and offering brokered transactions to a global market of media, at a price shared with the copyright owner. Citizenside.com is involved through AFP in the Glocal project.

If dealing with UGC and interacting with the audience are increasingly part of the professional work, sociologists of media have noticed the concerns of the journalists about newsworthiness, trust, legal rights, personal tone, subjective bias, independence and accountability of UGC.

High workload, lack of time and ressources, lack of procedures and tools to verify the reliability and accuracy of user-generated content make professional journalists more cautious, if not reluctant, to give credit to UGC, especially when some hoaxers manage to slip past the gatekeepers by producing fake photos or videos, not to say an entire false event like the Bluewater fake suicide bombing in September 2009 [10].

But these hoaxes do not prevent that amateur pictures like Mr Krums' Hudson River plane are stunning breaking news, while the *networked public sphere* of activists, amateurs, students and security experts, who did the news gathering, analysis and distribution on public concerns about the Diebold electronic voting machines in the US, outplayed with their investigation the mainstream media [11].

In newsrooms, "there is indeed a strong belief that the primary role of journalism lies in the selection stage of the news making process. Their gatekeeping

skills are among the major traits through which professionals distinguish themselves from amateur journalists”, point out Social scientists Steve Paulussen and Pieter Ugille.

How to assess the quality of the content and increase confidence in the users news production to avoid what researchers Alfred Hermida and Neil Thurman called ”a clash of cultures”? [12] How to engage a conversation with the amateurs producers to get more information on the context of their work and take into account the different views of an event? How to avoid hoaxes and fakes which hurt mainstream media reputation and discredit their newsroom as well as UGC?

So far, the answers to these questions rely only on the newsrooms facts checking policies and on the investigation skills, especially online, of their staff. Nevertheless, this is where both social media and innovative computer sciences techniques may help to assess the reliability of UGC and to increase confidence and collaboration between amateurs and professionals on events coverage.

The notion of event is still rather limited in the news industry where it has been mostly used to prepare agendas of planned events, using a common interchange format, instead of indexing unplanned breaking news and to propose a user experience entirely based on events retrieval.

Furthermore, prevention of hoaxes often relies on merely checking the coherence of the Exif/IPTC/XMP metadata with what the picture is about and to eventually detect visible photo retouching.

3 The Glocal approach

We precisely intend to go beyond those limitations in Glocal, a European Commission funded project, standing as a boundary object in the pro-am co-production of news. A boundary object is a concept borrowed from the sociology of Science describing the intersection of different social worlds. It was introduced in 1989 by Social scientists Susan Leigh Star and James R. Griesemer [13] in an ethnographical study of the coordination mechanisms of scientific work between amateurs and professionals, revisiting the Callon-Latour Actor-Network-Theory (ANT) [14].

Science and Technology are propitious to mass collaboration between amateurs and professionals. Astronomy with discoveries of exoplanets, asteroids or supernovas done by amateurs and open source software development, like Linux distributions relying on thousands of voluntary programmers, are well known examples of this so-called Pro-Am revolution [15].

With an innovative user interface fostering the pro-am collaboration and the sharing of different focus on the news, with event modelling based on the International Press Telecommunications Council (IPTC) standard EventsML-G2 [16] to improve the content metadata in a more professional way and to enhance the user search experience, with visual search by similarity, facts checking tools, linked data to other related events on news websites, Glocal aims at building a new kind of news information marketplace.

A study on the top queries performed on AFP's ImageForum photographic database (10 million pictures) [17] over 18 months showed that professional users are primarily querying for locations where a news event takes place (54%), then for persons names (22%), sports (12,5%) and other type of events such as disasters, entertainment (the Oscars Ceremony for instance) or political events (G8, elections...).[18]

This keyword-based queries gives a glimpse of the path to improvement that we can foresee by modeling events and by displaying in a user interface facets clustering the events related named entities (locations, persons, organisations, works) and the relationships between them, as well as relating those events to different thematics such as the IPTC newscodes taxonomy.

We consider Events as a kind of metaclass of an ontology describing *what happens*. If we take the analogy from the five W's of journalism (What, Where, When, Who, Why), we can consider that each newsworthy event describes *what happens* (the action) and its classes: *where it happens* (location), *when it happens* (time range), and *who is involved* (persons, companies, organisations, ...).

This factualization of events is very much at the heart of the recently released EventsML-G2 framework whose syntax is based on XML and also complies with RDF for exchanging event information in the news industry environment.

Why and How the event happens are beyond the scope but we plan to pay special attention to the focus, the point of view of who (actor, eyewitness, victim, direct or undirect source, media) is reporting about the event, a metadata which is the infosource element in EventsML-G2.

Coming back to our example of the Hudson miracle, the event behind the sentence "there's a plane in the Hudson. I'm on the ferry going to pick up the people. Crazy." was actually reported by AFP as the following news alert:

US-air-accident

A US Airways plane carrying commuters crashed on Thursday in New York City's Hudson River and was floating in the water, according to US media.

With an event modelling based on the main 4 W's of journalism (What, Who, When, Where) and proper sourcing, Glocal will provide to the amateurs an interactive framework to describe and categorize their event experience bridging the gap between their local representation and the global one, in an attempt to construct common definitions and meanings and define representativities like in the translation process of ANT.

Furthermore, Mr Krums' photograph as posted on Twitpic does not show any useful Exif metadata such as the camera used to take the picture or the GPS coordinates of the place where the picture was taken. If present, both elements would have increased the confidence in the reliability of the content.

In Glocal, special attention will be given to preserve as much metadata as possible and to check the coherence of this metadata with the photo or video object structure, taking also into account the user profile, uploads history or IP address, as well as the user description of the context of their contribution.

Location-awareness, with the use of geocodes eventually stored in the Exif metadata and/or geotags annotations added by the users or declared in their

registered profile, will help to display events on maps and whenever necessary, Glocal will help the user to disambiguate location tagging with the use of an external knowledge base such as GeoNames [19], minimising interactions between users and the system, by showing a map view of the event at an optimal zoom level.[20]

Although fact checking is the key to ensure the reliability of the content and deliver a guarantee to media willing to use this material, attention will also be paid to protect the user's privacy, especially when the content comes from countries where the freedom of speech is not respected.

Techniques like visual signatures, similarity search allowing to check if a picture or a video presented as a news event has already been published somewhere else on the web will help to increase confidence in the reliability of the content while deeper analysis of the image structure may help to detect if the digital images have been tampered with.

Social media tools will be used to manage, engage and reward the community of users while similarity and event search will also allow for large scale event matching on the web, linking Glocal news to other linked data sources of information.

Glocal will mainly focus on image material, photo and video. "While much of the public feels deficient in writing skills, anyone can take a picture or use a camcorder", remarks Ken Doctor, while in a comparative study on UGC use in The Sun (UK) and Aftonbladet (Sweden), Dr. Henrik Ornebring noted that "the only reader material that is given similar status to material produced by the news organisation is reader photos of breaking news events" [21] like the plane in the Hudson River.

4 Conclusion

In this paper, we presented the goals of the Glocal project, aiming at bridging the gap between the global (media) representation of events and the local (user) focus on those events, and at fostering collaboration between professionals and amateurs in the news production.

By analogy with the pro-am collaboration in Science and Technology, we showed that the rise of user-generated content is currently setting up a new deal in the media industry. Despite concerns about reliability and trust, media are dealing with UGC and encouraging participation and dialogue with the amateurs.

By using events as the primary means for organizing and indexing media, even in the case of unplanned breaking news, Glocal aims to go beyond the state of the art of existing systems in the news industry.

As a pioneer of the concept of the *information marketplace*, Michael Der-touzos, the late Director of the MIT Computer Science Laboratory, once wrote: "the Information Marketplace will inevitably cause us to bring together our humanistic and technical sides, which have been artificially split for centuries" [22].

In Glocal, we will try to bring together social media, news Pro-am co-production and computer science techniques with the goal to participate in the future of the news and of the networked public sphere.

References

1. <http://www.glocal-project.eu/>
2. Twitpic is a web platform allowing to share photos on Twitter <http://twitpic.com/135xa>
3. M. McLuhan. *Understanding Media*. Routledge (1964)
4. <http://www.pewinternet.org/Reports/2010/Online-News.aspx>
5. http://www.useit.com/alertbox/participation_inequality.html
6. K. Doctor. *Newsonomics*. St. Martins Press, New York (2010)
7. http://www.stateofthedia.org/2010/specialreports_community_journalism.php
8. S. Paulussen, P. Ugille. User-Generated Content in the Newsroom: Professional and Organisational Constraints on Participatory Journalism in Westminster Papers in Communication and Culture (2008)
9. Agence France-Presse (AFP) holds a stake of 34% in Citizenside.com
10. <http://www.wired.com/threatlevel/2009/09/bluewater/>
11. Y. Benkler. *Wealth of Networks*. Yale University Press (2006)
12. A. Hermida, N. Thurman. A clash of cultures: The integration of user-generated content within professional journalistic frameworks at British newspaper websites. In *Journalism Practice*, Vol. 2, No. 3, pp. 343-356, 2008
13. S. Leigh Star, J. R. Griesemer. Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39; *Social Studies of Science*, Vol. 19, No. 3 (Aug., 1989)
14. http://carbon.ucdenver.edu/~mryder/itc_data/ant_dff.html
15. C. Leadbeater, P. Miller. *The Pro-Am revolution*. Demos (2004)
16. International Press Telecommunications Council: <http://iptc.org/cms/site/index.html?channel=CH0112>
17. ImageForum: <http://www.imageforum2.afp.com/> and ImageForum diffusion: <http://www.imageforum-diffusion.afp.com/>
18. AFP internal study performed by the author in 2008 and later revised in 2009.
19. <http://www.geonames.org/>
20. P. Serdyukov, V. Murdock, and R. van Zwol. Placing flickr photos on a map. In *SIGIR '09: Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 484-491, New York, NY, USA, 2009. ACM.
21. H. Ornebring. The consumer as producer - of what? User-generated tabloid content in *The Sun* (UK) and *Aftonbladet* (Sweden). In *Journalism Studies* 9 (5) 771-785 (2008).
22. M. Dertouzos. *What will be: How the new World of Information will change our lives*. HarperEdge (1997)