

# CITIZENS AS SENSORS FOR NATURAL HAZARDS: A VGI INTEGRATION WORKFLOW

Bertrand De Longueville, Gianluca Luraschi, Paul Smits, Stephen Peedell and Tom De Groeve  
European Commission, Joint Research Centre, Ispra, Italy

*Recent advances of the Internet, often subsumed under the term Web 2.0, have led to an unprecedented increase of contents created by non-specialist users. The term Volunteered Geographic Information (VGI) is now commonly used when such user-generated contents have an explicit reference to a geographic location. VGI has significant potential as a timely and cost-effective means to enhance our overall understanding of the Earth. However, integrating VGI into specialist-driven Information Systems is a major challenge, as it is often regarded as insufficiently structured, documented, or validated.*

*In this paper, we address the research question of how the credibility of VGI can be increased, so that it becomes a valuable resource, within domains such as natural hazards. This is done by proposing a generic workflow that uses prior information about the phenomenon of interest and reasoning techniques to improve the reliability of the VGI, thus creating a useful source for scientific and technical investigation. This workflow has been developed for a particular case study: the use of pictures from the photo-sharing portal Flickr as a 'signal' that allows the pinpointing of recent flood events in the United Kingdom. A comparison of the workflow's output with independent information provided by scientists and journalists about these floods is also presented in order to uncover some of the strengths and weaknesses of VGI in our case study. We conclude that the proposed workflow allows VGI to be turned into a reliable data source.*

*Les avancées récentes de l'Internet, souvent subsumées dans l'expression Web 2.0 ont mené à une augmentation sans précédent de contenus créés par des utilisateurs qui ne sont pas spécialistes. Le terme information géographique volontaire (IGV) est maintenant utilisé couramment lorsque ce genre de contenu généré par les utilisateurs fait une référence explicite à une localisation géographique. L'IGV a un potentiel considérable comme moyen rapide et peu coûteux d'améliorer notre connaissance globale de la Terre. Toutefois, l'intégration de l'IGV à des systèmes d'information créés par des spécialistes est un défi considérable car cette information est souvent jugée insuffisamment structurée, documentée ou validée.*

*La principale question de recherche de cet article porte sur la façon dont il est possible d'améliorer la crédibilité de l'IGV afin qu'elle devienne une ressource fiable dans des domaines comme les risques naturels. Nous proposons un flux de travail générique qui fait appel à l'information antérieure sur le phénomène d'intérêt et aux techniques de raisonnement pour améliorer la fiabilité de l'IGV, créant ainsi une source utile pour les études scientifiques et techniques. Nous avons établi ce flux de travail pour une étude de cas précise : l'utilisation des photos du portail de partage Flickr comme « signal » qui permet de repérer précisément les récentes inondations au Royaume-Uni. Une comparaison des résultats obtenus avec de l'information indépendante fournie par des scientifiques et des journalistes au sujet de ces inondations est aussi présentée afin de mettre au jour certaines forces et faiblesses de l'IGV dans notre étude de cas. Nous concluons que le flux de travail permet de transformer l'IGV en source de données fiable.*

## 1. Introduction

Initially designed as a one-directional text broadcasting media, the Internet has rapidly become an information sharing platform, commonly referred to as Web 2.0. One of the key consequences of this evolution has been to increase participation, with lowered technical barriers permitting an unprecedented increase of content created by non-specialist users [Rinner et al. 2008]. When such user-generated content has a geographic dimension, it is now commonly referred to as Volunteered Geographic Information (VGI), which has a huge potential to engage citizens and to be a significant, timely, and

cost-effective source for geographers' understanding of the earth [Goodchild 2007].

To emphasize its potential as a novel source of information, O'Reilly [2005] described user-generated content created through blogs, wikis, and media-sharing platform as the *wisdom of the crowds*. However, such information is often regarded as insufficiently structured, documented, and validated to be a reliable source of information to perform scientific and technical analysis [Flanagin and Metzger 2008]. This is why Mummidi and Krumm



**Bertrand  
De Longueville**



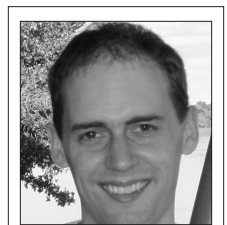
**Gianluca Luraschi**



**Paul Smits**



**Stephen Peedell**



**Tom De Groeve**