

# From Participative Science to Participative Action: Motivation-driven Crowdsourcing Workflows

Constance Thierry, David Gross-Amblard, Simon Chollet

September 19, 2024

## 1 Context

Crowdsourcing involves outsourcing tasks to contributors on dedicated on-line platforms. Tasks can be very varied, ranging from image labelling to the search for new molecules. But existing platforms generally focus on the digital world without asking the contributor for feedback on the physical world. Yet we could gain a great deal of information, particularly for environmental studies, by inviting the contributor to take part in action crowdsourcing tasks. For example, every year over a weekend, the ‘Bretagne vivante’<sup>1</sup> association invites individuals to observe, identify and count birds in their garden<sup>2</sup>. This type of action is essential for a better understanding of the local environment. Unfortunately, there are a number of problems with crowdsourcing actions, starting with the low number of contributors, which is mainly due to the fact that these actions are little-known and voluntary. Then there is the lack of expertise of the participants, who may censor themselves and not contribute information rather than make mistakes. In the case of garden birds, for example, participants are explicitly asked not to record birds that they are unable to identify.

The city of Rennes is already concerned about environmental issues and allows ecologists to use data from the city’s urban area for their studies [4]. Thanks to this data, ecologists can draw conclusions about the impact of decisions taken by the city on the environment. This is a valuable source of information for scientists, but it is limited because their studies involve public spaces, and they have little knowledge of the environmental state of the private gardens of the people of Rennes. Unfortunately, action crowdsourcing is currently not sufficiently exploited. By involving its citizens in participating in crowdsourcing actions of this type, for example, the metropolis of Rennes could have a more detailed knowledge of its environment and act accordingly. It is possible to imagine that following action crowdsourcing campaigns, information campaigns will be carried out among citizens to make them aware of their environment and the actions to be taken to improve it.

## 2 Theoretical approach

Collecting data during crowdsourcing campaigns is a complex task because the quality of the data collected depends as much on the definition of the campaign as on the quality of the data itself [1, 3, 7]. Three approaches can be considered to improve a campaign success. First, various tasks should be proposed to gradually increase one participant’s motivation. These tasks have to be properly orchestrated altogether, using proper workflow management systems. Second, people skills have to be recognized and nurtured by the platform, using a proper skill model. And third, users should be able to participate even if their opinion on a given fact is imprecise or uncertain. These ingredients are available in the HEADWORK platform that we develop at Univ Rennes [6]. Task workflows are realized by finite state relational automata, participant skills are monitored using skills taxonomies, and imprecise and uncertain answers are modeled using so called belief functions [5].

**Example of a complex action crowdsourcing workflow.** The study of vegetation is at the heart of the interests of ecologists. These studies, carried out in an urban environment, are all the more crucial because greenery has an impact on people’s well-being [2]. Although they have access to data from public urban environments, information on private gardens remains inaccessible to them. So although studies have shown, for example, that late mowing in public areas promotes diversity, the experiment has not been carried out

---

<sup>1</sup><https://www.bretagne-vivante.org/bretagne-vivante/>

<sup>2</sup><https://bretagne-vivante-dev.org/coj/>

in private gardens. We could imagine a crowdsourcing campaign in which private individuals are asked to mow several plots of their garden at different frequencies. Individuals would then be invited to report their observations. Another task in this crowdsourcing campaign could be to plant flower strips. Collecting and aggregating this information would then enable ecologists to determine whether, on an individual scale, a space that is not mown is more relevant than a flower strip for biodiversity.

### 3 The internship

Our objective is to define an action crowdsourcing workflow that will motivate people to take part in campaigns to study the biodiversity of their garden. Crowdsourcing campaign flows will be managed by a state machine. The challenge is to allow imperfect responses so as not to limit the responses of non-expert contributors. We will use the theory of belief functions to model imperfect responses in a crowdsourcing context [8]. To do this, the trainee will begin by carrying out a state of the art in the field and familiarising himself/herself with the theory of belief functions.

In order to gain a better understanding of the needs and expectations of ecologists in terms of the data to be collected, the student will spend part of the placement working with them. In particular, they will carry out field campaigns with them.

#### Outline of the internship.

- Bibliographical study on action facilitation in participative science
- Simple action scenario in a local setting (Beaulieu campus for example)
- Proposition of a cost/incentive model for crowdsourcing workflows AND/OR generalization of imprecise/uncertain answers
- Design (if relevant, implementation) within the HEADWORK platform
- Evaluation of the campaign using volunteer students
- Report writing, towards publication

The internship is open to computer science master 2 level students. The student must have mathematical and development skills; knowledge of automata is an asset.

### References

- [1] M. Allahbakhsh, B. Benatallah, A. Ignjatovic, H. R. Motahari-Nezhad, E. Bertino, and S. Dustdar. Quality control in crowdsourcing systems: Issues and directions. *IEEE Internet Computing*, 17(2):76–81, 2013.
- [2] S. Bonthoux and S. Chollet. Wilding cities for biodiversity and people: a transdisciplinary framework. *Biological Reviews*, 2024.
- [3] A. I. Chittilappilly, L. Chen, and S. Amer-Yahia. A survey of general-purpose crowdsourcing techniques. *IEEE Transactions on Knowledge and Data Engineering*, 28(9):2246–2266, 2016.
- [4] S. Chollet, C. Brabant, S. Tessier, and V. Jung. From urban lawns to urban meadows: Reduction of mowing frequency increases plant taxonomic, functional and phylogenetic diversity. *Landscape and Urban Planning*, 180:121–124, 2018.
- [5] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 38:325–339, 1967.
- [6] D. Gross-Amblard, M. Tommasi, I. Rakotoniaina, C. Thierry, R. Singh, and L. Jacoboni. Headwork: a data-centric crowdsourcing platform for complex tasks and participants. In *EDBT*, pages 778–781, 2024.
- [7] M. Lease. On quality control and machine learning in crowdsourcing. *Human Computation*, 11(11), 2011.
- [8] C. Thierry, J.-C. Dubois, Y. Le Gall, and A. Martin. Modeling uncertainty and inaccuracy on data from crowdsourcing platforms: Monitor. In *Proceedings of the 31st International Conference on Tools with Artificial Intelligence*, 2019.