

Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination

By Aniket Kittur and Robert Kraut

This paper tries to understand what makes wikipedia so good and how other peer production systems can learn from it. Wikipedia is both a living proof and a model for a successful peer production system. But what makes it as good as a traditional expert curated encyclopaedia ? Is it just the large number of contributors ? It is definitely a possibility - many contributors result in higher quality and less biased articles. But, this doesn't explain the phenomenon convincingly.

Most models of collective intelligence are based on aggregating contributions from a large number of people and Internet makes this kind of aggregation easy. However, unlike tasks like estimating the weight of an oxen or predicting the number of gems in a jar, writing in an article is a cognitively taxing task and the quality doesn't automatically improve with increasing the number of contributors. In fact , the reverse holds true - as the popular software engineering adage goes "Adding manpower to a late software makes it later", just increasing the number of contributors will have a negative result. A sense of collaboration is needed to make it work.

The authors point out that two types of collaboration takes place - explicit coordination (such as discussions in a forum) and implicit coordination (such as workgroup structure and assignment of roles). Both kinds of collaboration improve the article quality in its formative stage but however interact differently as the age of the article increases.

Explicit coordination through discussions on article forums are needed in the formative stages of the article to set the structure of the article. As the article's age increases , more implicit coordination takes place - in the sense that the core of the article is fixed and most of the editing tasks are grammar checking and checking for vandalism. The paper's results suggests that even explicit coordination involves an implicit component - planning the discussion is done by a small subset of contributors.

This kind of concentration of contributors is a common feature in all peer to peer systems - a very small set of people do the planning and initial work and the rest add on to it incrementally. This can be observed in Reddit forums, in the Linux project etc. The conclusion is that having more editors work on an article increases the article quality if and only if an effective combination of explicit and implicit coordination is used.

Going beyond the paper, why does implicit coordination work in the later stages of article ? One possibility is that most of the contributors are not motivated enough to do the work of deciding the structure of an article, the topics it should include etc. However, they are comfortable with fact

checking and basic editing tasks. This is because the implicit coordination workflow gives them a framework and a rigid set of ways to cooperate. The next question is to ask what other mechanisms apart from implicit coordination can we use to harness contributions from a larger set of people?. Would it be possible to learn what kind of tasks and topics a particular author is interested in and can we use that to suggest articles to work on?

I did not like the part of the paper where the methodology used is outlined. I found it to be slightly poorly written - I had to struggle to get what the numbers meant in real life.

Even a free and open environment like Wikipedia needs coordination and peer to peer communication is ineffective if too many contributors are involved. This shatters some of the beliefs Bitcoin enthusiasts believe in.