Wikipedian: a social identity between work and contribution.

Léo Joubert
Aix-Marseille Univ., LEST, CNRS
Aix-en-Provence, France
joubert.leo@gmail.com

ABSTRACT

Contributors to the Wikipedia "free encyclopedia" identify themselves and are identified as "Wikipedians". A Wikipedian does not leave his job when he becomes a Wikipedian. Nor does he become a Wikipedian in his workplace. The worker's identity and the Wikipedian identity coexist in the social identity of an individual. On which patterns does this coexistence between worker's identity and Wikipedian identity operate? Beyond the differences specific to the social identity of each contributor, we will try to show that singulars transactions all take place according to a finite number of patterns that it is possible to count. At this stage of our analysis, we are able to distinguish five identity patterns: employment, learning center, alternative development, continuity in upset, parallel arena. Our model aims to better understanding of why a contributor stay in Wikipedia and identifies himself as a contributor.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative and social computing**; • **Social and professional topics** → **User characteristics**;

KEYWORDS

Wikipedia, work, identity, social, role, contribution

1 QUESTIONS

Contributors to the Wikipedia "free encyclopedia" identify themselves and are identified as "Wikipedians". In this exploratory work, we will look at Wikipedian as a kind of social identity that is hold differently by different individuals.

[3] encourages us to think of social identity as a transaction between personal categories and institutions. To explain this theoretical guideline in our particular context, we can take the example of an academic who decides to contribute to Wikipedia. The academic institution does not promote the same definition of knowledge production as Wikipedia. Wikipedia and the university are two institutions that will meet personal categories of the individual and

enter into transactions with them. Then categories will either be adapted or stay out of institution rules. In the latter case, contribution practice will probably stop. For instance, our very estimated professor will sooner or later come up against the "no original knowledge" rule that everything published on French Wikipedia must first have been written and demonstrated elsewhere. Whatever his employment, it is likely that the purpose of his work is to produce knowledge that has *never* been written or demonstrated elsewhere. In cases like this, Wikipedian identity radically subverts worker's identity.

More generally, [5] has established the fundamental role of work in the building of social identity. From the surgeon to the politician, when we read these words we not only have in mind activities but also characters, roles, social functions, ways of doing but also ways of being. [7] described the motivations for contributors to contribute. There is a high prevalence of "internal self-concept" motivations: one most efficient motivation is to be able to see himself as a contributor.

These results are consistent with those of [1] which show two-stage model of Wikipedian career. First, the individual makes minor editions - such as correcting spelling mistakes - on pages where he feels he knows the subject very well. In this first stage of his career, the actor see himself as quite isolated from others. The more involved the individual become, a subjective transformation occurs as a result of which he ends up perceiving himself as a member of a community.

So if either Wikipedia and work are strong place for social identity building, on which patterns does the arrangement between worker's identity and Wikipedian identity operate? Of course, our academic does not leave his job when he becomes a Wikipedian. Nor does he become a Wikipedian in his workplace. The worker's identity and the Wikipedian identity coexist in his social identity. Beyond the differences specific to the social identity of each contributor, we will try to show that singulars transactions all take place according to a finite number of patterns that it is possible to describe.

2 METHOD

Our study focuses on the French version of Wikipedia. We interviewed 30 Wikipedians during an event: the WikiConvention which took place in Strasbourg, France from 19 to 21 October 2017. Of these contributors, 15 were interviewed on site and 15 were interviewed by phone call. Interviews last an average of 45 minutes. The minimum duration was 31 minutes, the maximum duration was 2 hours. The interviews are transcribed in full.

Each interview began with the same instruction: "explain to me how you came to contribute to Wikipedia". When necessary, we insist and go deeper around 3 particular themes: the work of the individual, his first contact with Wikipedia and his opinion

of the rules of the wiki. This method produced what [2] called biographical interviews. These interviews provoke an attempt by the subject to narrate his own world.

This paper is a first contact with the collected data, in which we have analyzed the relative place of Wikipedia and work in the narrative of each of the actors. As is customary in inductive work, we have constructed our categories over the course of the analysis. In the state in which we are currently presenting them, we are beginning to reach a saturation phenomenon: the new cases can be described with the categories built on the cases already analysed.

3 RESULTS

At this stage of our analysis, we are able to distinguish 5 identity patterns: employment, learning center, alternative development, continuity in upset, parallel arena.

These categories must not be seen as iron cage in which all contributors are trapped for all their Wikipedian life. A Wikipedian can witch between identity patterns as he or she progresses as a contributor. Each identity pattern will be described in three stages:

- **How it articulates** worker's identity and Wikipedian identity;
- **This articulation will be documented** by examples from the interviews;
- **Each pattern is based on a tension between work and contribution.** As interview are read as narratives[2], then this tension reveals the plot of the identity pattern.

Using contributions narratives show us quite clearly that the question "why a contributor *come* in ?" is not the only relevant one when asking ourselves about incentives of contribution. We should also answer the question "why a contributor *stay* in ?". At this stage of our investigation, our response is the following : **because he can solve the plot that contributing create in his social identity**.

3.1 Employment

This first pattern is that of an embedding between the Wikipedian identity and the worker's identity. This embedding prevents the distinction between worker's identity and Wikipedian identity other than theoretically.

In our corpus, these individuals were all employees in the public relation sector. Often inserted in an agency, the contributions of these individuals respond to the order of a client seeking to upgrade his page. We do not have precise information on the remuneration of their activity, therefore Jean-Luc mentions an annual figure of 60,000 euros for a specialized agency he founded with a partner. However, the paid contribution is not his only activity and it is therefore impossible to calculate his income exactly using Wikipedia.

The particularity of this pattern is that it must prove its legitimacy within the community. Many important contributors to the wiki have created a strong monitoring climate for paid contributions:

> "It's very stressful to do paid editing because you yourself as a Wikipedian are in control and others as you declare your conflict of interest, the bar is much higher. I think there are things not allowed for paid publishers that are totally tolerated and accepted for the average publisher." **Jean-Luc**

Two strategies are implemented by contributors to continue to act according to this identity pattern:

- **Emphasize their seniority**. A paid contributor is more likely to produce sustainable contributions if he has been contributing for a long time;
- **Promote respect for community values**. The main challenge facing paid contributors is to show that their contributions are neutral and sourced.

Now it is possible to understand a paradox: while the paid editing is poorly viewed on Wikipedia, many central contributors engage in it with others knowing that they do. In fact, it must be understood that seniority give them a capital of recognition. In particular, being able to show that one respects the rule on Neutral Point of View (NPOV) seems very important in all interviews. For example, the page on the theory of evolution must mention *both* Darwinian theory and those of Christian creationist groups, and it must also mention the criticisms of one *and* of the other. To be an experienced Wikipedian is to benefit from the presumption that you are able to be neutral. One can imagine that paid editing have been completely forbidden by monitoring contributors. But it seems that powerful contributors gave enough proofs that they were neutral even being paid. This is why we observe a rule that deals with conflicts of interest instead of a forbidding one:

> "If you are paid - including benefits in kind - to contribute to Wikipedia in French, you must mention the identity of your client, your employer and your relationship or affiliation with that client or employer, either on your user page, on the discussion page of the articles in which you are involved, or in summary of the changes concerned." **Aide:Contributions rémunérées page ; 5 of April 2018**

Contributors who build their identity on this pattern must be able to match the identity of a public relations officer's work with a Wikipedian identity as a neutral contributor. The plot of this pattern is thus as follows: will a contributor succeed in facing the suspicion of conflict of interest ? And possibly the stigma of non-neutrality if other Wikipedians judge his contributions biased ?

3.2 The parallel arena

In the parallel arena identity pattern, Wikipedian identity is an extension of worker's identity.

This is the example of Marie, a librarian who makes visible on Wikipedia the objects of the museum that employs her. In his institution, the use of Wikipedia opened a debate with his hierarchy and the formal scope of his job evolved as she used Wikipedia:

> "when I arrived at the Musée a year ago, I adapted it. Each time I propose and I adapt a little my professional practices to the employee, to the institution, and I see what we can do on Wikimedia projects." **Marie**

As in the employment pattern, Wikipedian identity and worker's identity are closely linked. The parallel arena is the opposite : contributions are made within the goal of findings tools allowing to carry out tasks defined in a job perimeter. But, the case of Marie show us an frequent example of how a Wikipedian career can evolve in this pattern :

> "For example, when I'm on vacation, it's a bit like deformation... When I am on vacation I will take pictures, create articles

on Commons that will not necessarily have a link with my professional frame. Here, I'm going to go correct some pages... Did I already write an article that is really related to my professional framework no I don't think. But it is true that it is not in the service of my institution either, here is... I don't see it that way... After that, it's the time it takes, but... Yes, it's on personal time, that's it... Even for professional contributions, it happens to me, so less now because now it's a little more official too that I contribute to the Wikipedia project in the professional framework too, but here it is... I did it on the weekends, at night, because it's fun, it's nice..." **Marie**

This identity pattern blurs the distinction between work and other dimensions of social identity. By strengthening the commitment in the work, it allows at the same time to build a more adequate identity in the work with the self that it represents. But it is also through this commitment that a risk arises: social identity would end up becoming *only* a worker's identity. Being the only way of self-expression, work then becomes *at the same time* what defines us and what takes us out of ourselves. The plot of this identity pattern is thus the following: will a contributor be able to avoid this risk while preserving the greatest expressiveness in the work that the use of Wikipedia allows him?

3.3 The learning center

With the help of the image of the learning center we wish to insist on the fact that the Wikipedian identity makes it possible to build a worker's identity. This is the example of Angélique who, excluded from university, was able to develop on Wikipedia skills related to documentation work.

"And so the first article I submitted to the label, I think I stayed contributing because at the time the community was much smaller and much more friendly in general than now. My article which was thus sourced among other things with blogs, I proposed it to the label "good article" naively saying: "look I put everything that was written on this topic!!!!!". And there were contributors who were very kind and very helpful who said to me, "Uh... no, that's not how we do it. Sources are not of equal value or quality." And that's how I acquired in the end, I almost acquired a "researcher" method. It may not be the right term, but thanks to the other contributors who had these methods. And I owe them a lot for that..." **Angélique**

After long periods of uncertainty, Angélique found a job as a documentation assistant in a research team. It seems important to us to stress that Wikipedia has played the role of a school in the building of her social identity.

"after the event I mentioned, so the library exclusion. It was hell, the accumulation of situations of exclusion that became unbearable. I think Wikipedia helped me a lot because I found that I could talk almost on an equal footing with people who had degrees, PhD, I had nothing." **Angélique**

In the learning center pattern, Wikipedia is an arena to show the virtuosity of his work: it is not enough to be a good contributor, you must be seen as a good contributor. So the plot of this transaction is this: will a contributor manage to show that he is good, and will this demonstration be worth beyond Wikipedia to translate it into an opportunity on the job market?

3.4 Alternative valuation

The alternative valuation pattern is the opposite of the employment pattern. In employment, work identities and Wikipedia identities are indistinguishable. In alternative valuation, they are strictly parallel.

This is the example of Marc, risk manager in the human resources division but also passionate about history and geography since school.

"I think I have a frustration, I wanted to study history and geography. And in my environment, the parents, the others, my other friends said to me "ah but no! not that, it leads to nothing, tralala...". And finally I studied computer science, nothing to see... And I think I always kept that frustration of not having done those studies. And I think it was something I was really interested in, and I started contributing to it. And as the contribution gets better and better. If I look at my old contributions, I'm not ashamed of them but they're not encyclopedic, so now I think my contributions are encyclopedic." **Marc**

This passion for history is well known among his family. For example, he tells us this anecdote about his family:

"I was known a bit in my family as "ah, yes Marc, that's the guy you ask him a question about a capital city, a forest, a mountain, a political party he's capable of answering you..." but that's all it was. It was limited to conversations, discussions, but there was nothing more... Wikipedia... I didn't realize it right away, but I jumped into the rift." **Marc**

The contribution to Wikipedia allows Marc to open a space to create a historian's public work. In an informal discussion, Marc will confide to us that Wikipedia according to him is "a landmark for frustrated people". It seems clear to us that he wanted to emphasize that Wikipedia is for him and for others the means to make exist in the eyes of others elements of his social identity that he could not enhance through work.

Identity transactions based on this pattern are caught in a permanent tension between work and Wikipedian tasks. Of all the contributors to the corpus, these are the ones who invoke the most often the lack of time as a brake to contribute more or as justification for their exit from Wikipedia. The plot is thus the following: will these contributors manage to resolve the tension, even the contradiction between worker tasks ans Wikipedia's ones?

3.5 Continuity in the upset

The last pattern is that of a failure of worker's identity that Wikipedian identity can support.

This is Sylvie's example. After studying engineering, she was unable to find work in this field. It will finally be used by a bank for the maintenance of its computer system. In her first job she will experience a burn out that she attributes to her "stressful responsibilities". She is now assigned to a "less stressful" position.

Wikipedia will act as a guiding thread throughout this ordeal. During the interview, Sylvie adopts a contradictory posture: on the one hand she expresses her attraction for the values of the Wikimedia movement, on the other she repeats several times that the community is "unhealthy". She makes the decision to become an arbitrator because she sees neutrality as a rule of behavior and as an intelligent way of resolving conflicts. But this commitment is sometimes painful:

"It wore me out, spending 25 hours on an arbitration reading logs, things... To see the people who send you e-mails saying you must defend it, thing... Not to attack me personally but to try to influence me in one way or another on the arbitration. Then saying it's actually useless... We'll tell that person you stop, but in three months I'm sure their contributions slightly biased but still at the border of what it takes will start again."
Sylvie

If Wikipedia wore her out, why didn't Sylvie make the decision to stop contributing during all these years? Thinking the contribution in terms of identity transactions offers an explanation for this paradox: the contribution allows Sylvie to build a more satisfactory social identity. It doesn't matter that she doesn't like some Wikipedians and that some Wikipedians don't like her. What matters is that she is able to ensure the consistency of her social being. While her work no longer allows her to be defined, Wikipedia offers her elements of definition of what she is in the eyes of others. The plot of this pattern is thus the following: is Wikipedian identity enough alone to compensate for the failing elements of social identity?

3.6 Some remarks on the pertinence of patterns.

Thinking the contributor trajectory in terms of identity transaction opens a research program not only on the contributors themselves, but also on the evolution of Wikipedia itself. To conclude this exploratory work, we will put forward the following hypothesis: the categories in force on the wiki (administrator, vandal, wikilove...) only last if they are congruent with the identity transactions made by contributors. For the institutionalization of Wikipedia to continue, it must allow the various contributors to advance in their Wikipedia narrative. In other words, it must make it possible to advance in the resolution of the plot that we revealed in the analysis of each of the patterns.

For example, becoming a Wikipedia administrator can help contributors in the learning center pattern to value their practice of contributing to the labour market. In other identity pattern, becoming administrator has other meaning. It is because it restores the plurality of identities present that the administrator category is sufficiently stable and that it has regulated the wiki for more than 15 years. On the other hand, on the French version of Wikipedia, the decline of arbitration committees between 2006 and 2012 shows that the arbitrator was a figure that no longer allowed contributors to solve plot in their narratives.

It also shows us the peril hanging over Wikipedia. Anyone with minimal computer skills can put a wiki online; however, for it to persist over time it must be able to attract diverse identities into the contribution and allow its identities to unfold.

4 FURTHER EXPLORATIONS

At the end of this first data exploration, we will come to formulate two questions in addition to those that might come to the reader's mind and that are welcome.

4.1 What pattern for what contribution to page quality?

While everyone can contribute to Wikipedia, not everyone can produce contributions that remain permanently present on the pages. Priedhorsky *et al.* come to the following conclusion: 40% of sustainable contributions on English Wikipedia are produced by 0.4% of contributors[4]. **One may therefore wonder whether there is a strong difference in the sustainability of contributions depending on the identity pattern that the contributor adopts**.

Methodologically, answering such a question requires checking the lifetime of the words forming a contributor's contributions on the pages to which he contributed. We have already built such tools gathered in the R package called WikiSocio. We still have to answer a question that will determine the data extraction: which page to take for which contributor? This question is all the more important as a contributor who publishes 10 000 pages in his trajectory contributes seriously to 10 or 15 of them.

4.2 Is it possible to quantitatively measure qualitatively established identity patterns?

Welser *et al.* show that a role system has developed over time on Wikipedia. They discern four roles: the vandal fighters, the technical editor, the substantial expert and the net-worker[6]. For example, the expert contributes more than the vandal fighters on content pages while the latter is mainly active on discussion pages and monitoring tools. These roles correspond to quantitative patterns that can be measured on numerous contributors after isolating them on a small number by qualitative techniques. This quite convincing example allow us to search for **a quantitative signature of identity patterns**.

REFERENCES

[1] Susan L. Bryant, Andrea Forte, and Amy Bruckman. 2005. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. 1–10.
[2] Didier Demazière and Claude Dubar. 2004. *Analyser les entretiens biographiques*. Presses Université Laval, Québec.
[3] Claude Dubar. 2015. *La socialisation* (5e édition ed.). Armand Colin, Paris.
[4] Reid Priedhorsky, Jilin Chen, Shyong Tony K. Lam, Katherine Panciera, Loren Terveen, and John Riedl. 2007. Creating, destroying, and restoring value in Wikipedia. In *Proceedings of the 2007 international ACM conference on Supporting group work*. ACM, 259–268.
[5] Renaud Sainsaulieu and Norbert Alter. 2014. *L'identité au travail* (4e édition ed.). Les Presses de Sciences Po, Paris.
[6] Howard T. Welser, Dan Cosley, Gueorgi Kossinets, Austin Lin, Fedor Dokshin, Geri Gay, and Marc Smith. 2011. Finding social roles in Wikipedia. In *Proceedings of the 2011 iConference*. ACM, 122–129.
[7] Heng-Li Yang and Cheng-Yu Lai. 2010. Motivations of Wikipedia content contributors. *Computers in Human Behavior* 26, 6 (Nov. 2010).