

# governance

In order to meet our movement’s goal of making all the world’s information available to everyone, we have more work to do than human editors can do alone. We need help in the form of augmentation, which is when humans and algorithms work together. Though augmentation in the wikis is not new, it will be a growing part of the future of the wikis. To ensure that the contributions made by algorithms are productive, unbiased, and fair, we will need to stick to our movement’s principles of openness, transparency, and the ability for anyone to contribute. We should build closed-loop infrastructure and interfaces that allow anyone to contribute new algorithms, and for even non-technical editors to participate in training and tuning those algorithms. These principles would apply to all types of augmentation, whether it is in the aspect of content generation, content curation, or governing interactions between people.

Governance is a word meant to capture a broader scope than “Code of Conduct”. It refers to all the ways that people interact with each other on wiki projects, in both constructive and unconstructive situations. Current newcomers rarely contribute past their initial edits because of bad reactions to quality control mechanisms, algorithmic tools (bots) or policy. We see augmented governance practices as the vehicle that will safeguard, and simultaneously empower the Wikimedia community to become that desired safe haven for knowledge discourse through a set of human-centered principles.

## **Sections**

[Definition of governance](#)

[Aspiration](#)

[Augmented governance](#)

[Governance Strategy](#)

[Guiding Questions](#)

[Notes](#)

[Sources](#)

In order to meet our movement's goal of making all the world's information available to everyone, we have more work to do than human editors can do alone. We need help in the form of augmentation, which is when humans and algorithms work together. Though augmentation in the wikis is not new, it will be a growing part of the future of the wikis. To ensure that the contributions made by algorithms are productive, unbiased, and fair, we will need to stick to our movement's principles of openness, transparency, and the ability for anyone to contribute. We should build closed-loop infrastructure and interfaces that allow anyone to contribute new algorithms, and for even non-technical editors to participate in training and tuning those algorithms. These principles would apply to all types of augmentation, whether it is in the aspect of content generation, content curation, or governing interactions between people.

## Definition of governance

Governance is a word meant to capture a broader scope than "Code of Conduct". It refers to all the ways that people interact and communicate with each other on wiki projects, in both constructive and unconstructive situations. This is inclusive of mentorship, harassment, and straightforward conversations about content generation and content curation.

## Aspiration

The Wikimedia movement wants the sum of all knowledge to be available to everyone in the world. We also want the process to assemble that knowledge to be inclusive,

balanced, and safe for all participants. Current newcomers rarely contribute past their initial edits because of bad reactions to quality control mechanisms, algorithmic tools (bots) or policy. We see augmented governance practices as the vehicle that will safeguard, and simultaneously empower the Wikimedia community to become that desired safe haven for knowledge discourse through a set of human - centered principles .

## Augmented governance

Wikipedia wants to attract and retain a person who edits in good faith and has a relatively high quality of edits (a desirable newcomer). Therefore, our goal for an augmented governance is to leverage artificial intelligence to create a safe space for humans to collaborate with each other to craft and compile the worlds' knowledge.

Currently the interplay between augmentation and governance has been explored through bots such as HostBot, which welcome new contributors to Wikipedia, and through processes such as the Bot Approval Process. The reception of these machine-generated greetings and process guidance has been initially cold due to impersonal and sometimes ineffective guidance. However, the way forward is leveraging the technology to create a partnership spectrum in which bots and algorithmic tools support and enhance the role of humans within the Wikimedia communities.

Below are some specific examples of augmented governance activities that can exist in the future, drawing from all along

the spectrum of the human-machine partnership.

<b>Activity</b>	<b>Algorithm role</b>	<b>Human role</b>
Welcoming	Identify newcomers and send welcome message	Follow up machine-generated messaging with a real offer for mentorship
Flag content for quality review	Identify and list poor quality articles with associated reason	Correct machine-generated reviews
Suggest content revisions for inappropriate behavior	List common ways to rectify the poor quality condition (ie: vandalism, harassment, sockpuppetry)	State if recommendation was useful and rectify the content dispute within the article
Provide wayfinding	Identify in-context opportunities and provide recommendations for users to consult resources (ie: norm guidance or policy disputes)	Follow wayfinding recommendations and consult resources.
Connect like-minded contributors	Identify when contributors are inhabiting similar spaces on the wikis and suggest contact	Accept suggestion and make human contact
Expertise and good behavior recognition	Identify when contributors have routinely provided their expertise within a wiki and provide a micro-credential such as a badge assertion.	Endorse micro-credentials. Once issued micro-credential, choose how to display it.

## Governance Strategy

In the 1987 film RoboCop[1], a picture is painted of a dystopian city on the decline that is dealt with by employing an army of robots to brutally police the city. This is not the future of governance that we want at Wikimedia. The majority of AI systems and related tools that are being created in the world today are being put in place with minimal oversight, few accountability mechanisms and little research into their broader implications. Although artificial intelligence is a powerful editing aide, it

also has the potential to powerfully magnify the problems of bias and unfairness [2] that already exist in the wikis, and has the potential to discourage new editors. [3] Currently, there are no internally agreed-upon methods to measure and assess the social implications of governance augmentation. Therefore, to ensure that the systems that we use and create are responsive to the complex social domain of Wikipedia communities, we will need to develop ways to measure, audit, analyze and improve them.

There are two ways that we can concretely do this:

1. The Wikimedia community needs to generate a set of guiding principles to ensure that the wikis are authentic and sincere representations of the worlds' knowledge. The following are topics that will be addressed:

- > **Human-centered AI:** AI should be "human centered", meaning that it nudges and opens the door to human connection instead of closing it. Wikimedia augmentation champions you – your security, privacy, and the quality of your online activity within our tools. Wikis got where they are today because of people working together, so there is something good about human connection that we should be mindful of preserving.
- > **Learnability as a core metric:** Wikimedia is optimizing for learning, which is a fundamentally different use of algorithms than recommendation systems built for revenue. Therefore, the core metric for success for all of the augmentation tools within Wikimedia's tools and projects is learnability, how well did the tool or service assist the learning experience?
- > **Transparency:** While augmentation tools (hopefully) make sensible decisions for you that respect your time, data, and attention, you should always be made aware when augmentation tools are being utilized.

2. Developers and administrators will create and utilize open governance algorithms. If anyone can contribute to these tools, and if Wikimedia makes it possible for non-technical editors to watchguard them so that they aren't allocation or representation biased, we have the opportunity to define the role of human computer interaction within the context of governance within the communities that we create and run.

## Guiding Questions

- > What do WE mean by fairness?
- > How might augmentation help us detect 'fake news' and misinformation campaigns from powerful malicious actors, like nation states?
- > How can we ensure that individual editors (new, experienced) feel that their contributions are still VALUED in a wiki where so much is created and curated by machines?
- > How can machines enhance the governance work done by editors?

## Notes

- [1] <https://en.wikipedia.org/wiki/RoboCop>
- [2] [Basic description of JADE by Aaron Halfaker](#)
- [3] Halfaker, A., Gieger, R. S., Morgan, J., & Riedl, J. (2013). **The Rise and Decline of an Open Collaboration System: How Wikipedia's reaction to sudden popularity is causing its decline.**
- [4] American Behavioral Scientist 57(5) 664-688.

## Sources

M. Marshall and J. Klein [Research and Insights](#), Other contributors: A. Halfaker, D. Garry, D. Horn, J. Katz, J. Minor, J. Morgan, T. Negrin, M. Novotny, N. Pangarkar

Aaron Halfaker. 2017. **Interpolating Quality Dynamics in Wikipedia and Demonstrating the Keilana Effect.** In Proceedings of the 13th International Symposium on Open Collaboration (OpenSym '17). ACM, New York, NY, USA, Article 19, 9 pages. DOI: <https://doi.org/10.1145/3125433.3125475>

Halfaker, A., Gieger, R. S., Morgan, J., & Riedl, J. (2013). **The Rise and Decline of an Open Collaboration System: How Wikipedia's reaction to sudden popularity is causing its decline.** American Behavioral Scientist 57(5) 664-688.

Buowamlini, Joy (2018). **The Dangers of Supremely White Data and The Coded Gaze** [Video from Wikimania 2018]. Retrieved from <https://youtu.be/ZSJXKoD6mA8>.  
[https://www.mediawiki.org/wiki/File:Conversations\\_Gone\\_Awry\\_\(slides\).pdf](https://www.mediawiki.org/wiki/File:Conversations_Gone_Awry_(slides).pdf)