OBJECTIONS TO FACEBOOK'S AUTOMATIC ALTERNATIVE TEXT

PATRICK STEPPAN

"Is AAT better than nothing? Perhaps, but the alternative is not nothing. The alternative is an opportunity to increase awareness of alt-text, the role it serves, and how to effectively write it."

In 1999 the Web Content Accessibility Guide was first published, and its first of 14 guidelines is to 'provide equivalent alternatives to auditory and visual content.'[1] Images on the internet have since been associated with an alternative text, or alt-text. Alt-texts are meant to serve the same purpose as the image they're associated with.

For people who have impaired vision, these texts are vital to reading things on the web. They are often communicated by screen-readers, braille displays, or other means, such that the reader will understand the image's purpose with or without the ability to see the image itself.

Social media platforms handle alt-text differently. Twitter has an opt-in alt-text composition setting, that allows the user to provide an alt-text for the photos they post. Instagram provides no official alt-text support, however, the caption area can easily be used for the same purpose as an alt-text.[2]

Since 2016, Facebook defaults to generating alt-text for images uploaded by it's users. They call it automatic alt-text, or AAT.[3]

Facebook's approach suffers from a number of problems:

1. Object detection is posited as a valid method of generating automatic alt-text.
2. Their object detection is too vague to be useful.
3. 'Objectivity' is framed as the default.
4. Users who need alt-text are considered a problem to be cured.
5. It undermines care and the process of normalizing accessibility.

Images and texts are modes of communicating information. When describing an image as text to another person, there is no one-to-one translation. What an image is, and what a text is, depends largely on its context.

Context is what frames an image; meaning arises from their interaction. What is gray against a black background, what is gray against a white background? This is a reductive approach, but it provides a helpful framework for understanding what Facebook's AAT algorithm is capable of. The algorithm, simply put, is based on object detection; meaning the alt-text is based solely on the the image, and not its context. What is gray now?

Facebook's Accessibility team based the AAT on a modified version of a fast region-based convolutional network (Fast R-CNN)[4, 5] which processes an image in two main steps: image segmentation, and image classification. The algorithms for these steps were once named DeepMask and SharpMask, but have since evolved into MultiPath, itself part of a broader framework at Facebook known as Lumos[6].

Those details of the algorithm are not relevant, because the approach is inherently flawed. Facebook could have used any neural net and the problem would be the same: an image is not its objects. Object detection is not a valid algorithm for generating alt-text. Regardless of that, the product of the algorithm is often vague and of little substance. How is 'Image may contain: One or more people, smiling' or 'Image may contain: text' helpful?

Is AAT better than nothing? Perhaps, but the alternative is not nothing. The alternative is an opportunity to increase awareness of alt-text, the role it serves, and how to effectively write it.

Rather than fostering a user-experience of care, Facebook offloads the labor to server farms. Facebook engineers justify their approach to AAT, because it "[provides] unmatched coverage and convenience for photos on large-scale services such as Facebook."[7] At its best, AAT is another example of technologists' misguided vision to solve or *cure* accessibility; and at its worst it subverts the social responsibility of accessibility by framing it as an unnecessary concern. It skirts social-responsibility in favor of driving higher engagement, and building a 'seamless experience'. A seamless experience for who?

AAT directly benefits Facebook in other ways that have nothing to do with accessibility. Search engine optimization, spam detection, flagging objectionable content, and streamlining user experience. Driving user engagement and therefore ad revenue are clear motivations.

Alternative meanings for AAT may contain:

- Artificially Altruist Technologists
- Awful Ableist Thinking
- Anti Access, Truly.

Facebook's process for adding human-authored alt-text starts with adding a photo, selecting **edit photo** and then **Alt Text,** and finally **Override generated alt text**. Upon clicking **Clear** the AAT is restored. One reading of this UI is how Facebook posits AAT as neutral; or that the objects detected are the inherent description of the image.

AAT ought to be secondary to human-authored alt-text. For the third most accessed platform in the world, Facebook has incredible potential for bringing awareness to alt-text, and surfacing the option for human-authored alt-text as the primary model. Contextual information and nuanced reading of images are important for writing effective alt-text.

Algorithms aren't a substitute for social responsibility. As an alternative to AAT, I argue Facebook should prioritize human-authored alt-text; working towards normalizing accessibility and placing care over cure.

WORKS CITED:

1. Techniques for Web Content Accessibility Guidelines 1.0. Accessed November 04, 2018. https://www.w3.org/TR/WAI-WEBCONTENT-TECHS/#Techniques.

2. "Accessible Social Media." Accessible Social Media | Accessible U. Accessed November 04, 2018. https://accessibility.umn.edu/tutorials/accessible-social-media.

3. "Under the Hood: Building Accessibility Tools for the Visually Impaired on Facebook." Facebook Code. June 26, 2018. Accessed November 04, 2018. https://code.fb.com/ios/under-the-hood-building-accessibility-tools-for-the-visually-impaired-on-facebook/.

4. Zagoruyko, Sergey, Adam Lerer, Tsung-Yi Lin, Pedroo. Pinheiro, Sam Gross, Soumith Chintala, and Piotr Dollar. "A MultiPath Network for Object Detection." *Procedings of the British Machine Vision Conference 2016*, 2016. doi:10.5244/c.30.15.

5. Lerer, Adam. "A Torch Implementation of the Object Detection Network from "A MultiPath Network for Object Detection" (https://arxiv.org/abs/1604.02135)." GitHub. November 28, 2016. Accessed November 04, 2018. https://github.com/facebookresearch/multipathnet.

6. Candela, Joaquin Quiñonero. "Building Scalable Systems to Understand Content." Facebook Code. June 26, 2018. Accessed November 04, 2018. https://code.fb.com/ml-applications/building-scalable-systems-to-understand-content/.

7. "Under the Hood".