Landscape Analysis of Commercial Visual Assistance Technologies

Emma Sadjo
University of Washington
Seattle, Washington, USA
esadjo@uw.edu

Leah Findlater
University of Washington
Seattle, Washington, USA
leahkf@uw.edu

Abigale Stangl
University of Washington
Seattle, Washington, USA
astangl@uw.edu

ABSTRACT

We present a landscape analysis of commercially available visual assistance technologies (VATs) that provide auditory descriptions of image and video content found online, as well as those taken by people who are blind and have visual questions. Through structured web-based searches, we identified 20 VATs released by 17 companies, and analyzed how these companies communicate to users about their technical innovation and service offerings. Our results can orient new researchers, UX professionals, and developers to trends within commercial VAT development.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in accessibility**.

KEYWORDS

visual assistance technology, image description, image caption, blind, landscape analysis

ACM Reference Format:
Emma Sadjo, Leah Findlater, and Abigale Stangl. 2021. Landscape Analysis of Commercial Visual Assistance Technologies. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21), October 18–22, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3441852.3476521

1 INTRODUCTION AND BACKGROUND

Visual assistance technologies (VATs) provide auditory descriptions of images and videos to answer blind people's visual questions about their physical surroundings and visual media found online and in documents. Research on VATs has focused on technical development (e.g., [19, 50, 51]), blind people's access requirements (e.g., [31, 42, 48]), privacy concerns (e.g., [9, 10, 14, 42]), and ethical questions representation (e.g., [15, 25, 41]).

In parallel, companies that develop commercial VATs, such as Seeing AI [33], Be My Eyes [20], Aira [8] are driving innovation. These commercial VATs can have far reach, as exemplified by Be My Eyes, who describe their product as being used by "295,170 blind & low-vision people" in "150+ countries" speaking "180+ languages" [20]. Yet, for researchers, UX professionals, developers, and users who are interested in advancing the design and development of commercially available VATs, it can be challenging to identify and assess the range of VATs and the trends guiding their development. In turn, we conducted a landscape analysis of 20 VATs from 17 different companies using materials found on VAT company websites, industry blogs, social media, and digital application stores.

We discovered that VAT companies use five primary concepts to describe their service offerings, are highly oriented to leveraging AI to provide visual assistance on iOS-based hardware, and are more likely to collect and store users' visual data on "the cloud" than on users' devices—though the trend to process visual data on device is growing. Our work may be used to spur further investigation and critical reflection about the influence of AI on the way visual assistance is provided and how VAT companies' messaging is aligned with users' values, goals, and decision making.

2 METHODS

Landscape analyses examine key aspects of products and services that can be used to observe technology sector evolution and opportunities for new innovations [5]. Two sighted US-based academic researchers spent approximately 60 hours from January to April 2021 collecting data based on web search results. We identified 20 commercially available VATs by querying both Google and Bing using the phrase *visual assistance technology for blind*, which denotes the technical domain of our investigation and the primary user group of the technology. After reviewing each platform's first 30 search results, we extended the search by using other keyword phrases we found on product websites and blogs[1]. For each new search term, we reviewed the first 10 search results.

Next, we created a codebook with nine parent codes to guide our data collection. The parent codes emerged as we observed trends in the information companies readily provide to potential users on the web. Five of the codes—*Form Factor*, *Visual Media Type*, *Operating System*, *Intelligence Type*, and *Visual Data Storage*—were informed by previous work on VATs [6, 10, 28, 40, 42, 50]. We included four additional codes—*Mission Statement/Slogans*, *Keywords*, *Release Dates*, and *Use Metrics*[2]—based on our observations that each VAT company uses varying language to describe its products, and with the intent to provide more context to product positioning within the spectrum of commercially available VATs. Together the parent categories help frame how individual VATs are positioned in relation to the others, and enable us to observe how their messaging

ASSETS '21, October 18–22, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8306-6/21/10.
https://doi.org/10.1145/3441852.3476521

[1] Other keywords include: *image description technologies, technologies for people who are blind [blind people], assistive technologies and [low vision, blind], top apps for blind or visually impaired people, assistive apps for blind.*

[2] During data analysis, we found that only eight VATs posted use metrics, and therefore, we cannot report which VATs are the most popular. Metrics reported included: (A) number users, (B) number of images processed or tasks completed, (C) global reach of their technology by country and/or language, and/or (D) download counts via app stores.

relates to best practices as described in access technology research, e.g., [13].

We visited each VAT company's home page to collect data, and copied relevant text into a data sheet for each of the nine parent codes. When data was not available, we searched Twitter, Facebook, Google Play, and Apple App Store for VAT product and company profiles. We also used external blogs and the Wayback Machine [30], an internet archive tool[3]. During data collection, trends began to emerge, such as VATs take *image files, video files*, and *live feed/images* as input, which became the sub-codes for *Visual Media Type*. In our final code book[4], each parent code included two to four sub-codes. To observe notable trends for each code, we calculated the number of instances for each sub-code. In some cases, we observed trends across multiple codes. Notably, the data under *Mission Statement/Slogan* and *Keyword* was rich, and required creating affinity diagrams [18] to identify the range full range of sub-codes.

3 FINDINGS

We identified 20 commercial VATs, developed by 17 companies, including: *Adobe Accessibility* by Adobe [1], *Aira* by Aira Tech Corp [8], *Amazon Rekognition* by Amazon [11], *Apple VoiceOver Recognition* by Apple [12], *Be My Eyes* by Be My Eyes [20], *BeSpecular* by BeSpecular [17], *TapTapSee* by CloudSight Inc. [46], *Envision AI* by Envision Technologies B.V. [4], *Envision Glasses* by Envision Technologies B.V. [23], *Automatic Alt Text* by Facebook [21], *Google Lookout* by Google [3], *LookTel Money Reader* by LookTel [38], *LookTel Money Recognizer* by LookTel [29], *Supersense* by Mediate [44], *Seeing AI* by Microsoft [33], *OrCam MyEye* by OrCam [36], *OrCam Read* by OrCam [35], *KNFB Reader* by SensoTec [2], *Sullivan+* by TUAT Inc. [37], and *AI Poly* by V7 Ltd [47]. We have provided an overview of VATs according to their *Form Factor*, with details related to the *Operating Systems* they run on, their *Release Dates, Intelligence Types*, and *Visual Data Storage*. We then present findings from the *Mission/Slogan* and *Keywords* parent codes. Please see the **Supplementary Materials** for a tabulated version of these findings with the online data sources we used.

Findings by VAT Form Factor. VATs are provided as (A) mobile applications, (B) embedded software (on different types of devices), and (C) on web-based services. Here we present an analysis of our findings according to this categorization. VATs are most commonly offered as *mobile applications* (13 of the 20). Of these mobile applications, nine come in both iOS and Android versions, three offer only an iOS app (*LookTel Money Reader, LookTel Recognizer*, and *Seeing AI*) and one (*Lookout*) offers only an Android app. Of all 20 VATs, the *LookTel Money Reader* was the first to be released, in 2011 on iOS. *TapTapSee* was the first Android app, debuting in 2014. Seven companies released their iOS version first, on average more than six months ahead of the Android version. Of the 13 mobile apps in our dataset, we see that 11[5] use AI to provide visual assistance; *Be My Eyes* and *BeSpecular* employ humans and *Aira* is the only VAT that uses both AI and humans (humans deliver the visual assistance, yet the dashboards they use to gather information for users is AI-powered). Of the ten VATs that use AI, only two take video as input (*TapTapSee* and *Aira*); the other 12 take image files and/or live images as input. Five of the ten utilize users' on device storage, whereas seven process users' data on "the cloud". (We could not find data regarding data storage two of the AI-powered mobile application-bases VATs: *Supersense and AI Poly*, and we learned that *Google Lookout* and *Seeing AI* "do not store users data"). Importantly, we found data indicating that *Envision AI* and *Aira* use both on device and cloud-based storage, but it is unclear when each approach is used. Envision AI processes live images, where as Aira processes videos. Of three VATs that employ human intelligence, two take video files only as input (*Aira* and *Be My Eyes*), and one takes image files and video files as input. All three store data in the cloud (though *Aira* also stores data on device).

Four of the VATs run as *embedded software*, including *Apple VoiceOver Recognition*, which was released in 2017 and uses AI to provide image descriptions. *OrCam MyEye* and *OrCam Read* run on hand-held devices, while *Envision Glasses* runs on wearable technology; These three VATs employ AI exclusively, supporting both images and video for *Envision Glasses*, but only images for the *OrCam* products. All three store data both on device and in the cloud, while we could not find information on data storage for VoiceOver Recognition. Finally, three of the 20 VATs run as *web-based services* (*Adobe Accessibility, Amazon Rekognition, Facebook Automatic Alt Text*). Amazon Reckognition was released in 2015, Facebook Automated Alt Text was released widely in 2016, and we could not find an Adobe Accessibility release date. All employ AI and take image files as input. *Adobe Accessibility* and *Amazon Rekognition* store data in the cloud, and we could not find storage information for *Facebook Automatic Alt Text*.

Company Messaging. From the "Keywords" analysis we learned that VAT companies use 18 different terms to describe their VATs, including "visual assistance"; see footnote for full list[6]. By analyzing the Mission Statement/Slogan data, we were also able to identify five sub-codes for how companies discuss their VATs, as described below.

Computational Innovation. Six companies use language that highlights the computational techniques used to provide visual assistance, such as *"An AI Assistant for the blind"* [7], *"Computer vision for the blind"* [44], and *"Automate your image and video analysis with machine learning"* [11]. Moreover, three companies incorporated AI into their VAT product names: Aira = AI+RA (Remote Assistance), *Seeing AI*, and *Envision AI*.

Focus on the Senses. VAT companies commonly mention vision and hearing in their product descriptions. Seven companies emphasize vision, such as *"...let it become your eyes to seeing the world"* [37]

[3] While VAT companies' data storage practices were of interest, we ultimately deemed examining privacy policies to be out of scope, based on the understanding that people commonly do not read these legal documents as they are difficult to understand [27, 32, 39].

[4] *VATs Form Factor* [Mobile, Glasses, Other], *Operating System* [iOS, Android, Other], *Intelligence Type* [Human, AI], *Data Storage* [On Device, Cloud, No Storage, Unknown], *Use Metrics* [Number of Users, Tasks Completed, Region, Language], and *Release Date* [Year(s)].

[5] *AI Poly, Aira, Envision AI, Google Lookout, KNFB Reader, LookTel Money Reader, LookTel Money Recognizer, Seeing AI, Sullivan+, Supersense, and TapTapSee*

[6] The 17 addition keywords/phrases we observed include: *"AI assistant for the blind", "AI assistive technology", "AI digital reader", "AI-powered assistive glasses", "Assistive reading device", "Automatic alternative text", "Currency identifier", "Intelligent camera app", "Mobile assistant", "Object and color recognizer", "Object recognition technology", "Visual Assistance" "Visual-aid app", "Voice-activated device", "Wearable Assistive Device", "Wearable assistive technology"*, and *"Wearable technology"*.

and orienting vision as a resource shared by sighted people (e.g., *"Let blind people see through your eyes"* [17]). Companies that use the terms "hearing," "speech," or "audio" do so with respect to: (A) transmuting vision into language, (e.g., *"Seeing AI turns the visual world into an audible experience"* [34]), (B) mediating communication between the senses (e.g., *"Hear what you want to see"*), and (C) the transformative ability of the technologies (e.g., *"Hear what's happening on your screen"* [12]).

Enhanced Experience. Companies reference user experience in a variety of ways. Six use the word **"experience"** directly, such as *"Supersense wants to kindle a new generation of experiences and opportunities for every blind individual using technology..."* [44]. Two companies highlight **personalized** experiences, i.e., *"Built-in features that work the way you do. Make them yours, and make something wonderful"* [12]. Seven companies focus on increasing **convenience** for users, such as *"to provide instant access to information to anyone, no matter where they are in life's journey."* [8]. Three companies framed their VATs as creating **independence** for users, such as *"[Supersense...] helps blind and visually impaired users to read, find objects, and explore places independently"* [44]. Only four companies used terms that highlight **connection and community**, such as *"Facebook's mission is to make the world more open and connected, and that means everyone, including the visually impaired community"* [21].

Object of Access. Drawing attention to the object of access, six companies use the word **world** (e.g., *"Let it become your eyes to seeing the world"* [37]), while four companies focus on **everyday life/objects**, (e.g., *"TapTapSee is designed to help the blind and visually impaired identify objects they encounter in their everyday lives"* [45]).

Language of Assistance. Companies also use language to describe how VATs support people who are blind. Four use the term **"help(s)"**, such as *"an object and color recognizer that helps the blind, visually impaired, and color blind understand their surroundings"* [47], and *"Supersense - helps blind and visually impaired"* [44]. Three others use words similar to "help", such as **"allows"**, a word that implies that the products assist blind and visually impaired users, e.g., *"LookTel Recognizer allows users [...] to instantly recognize everyday objects"* [29]. Finally, two other companies use the term **"empower"**, such as *"Orcam. Our Vision. Empowering People"* [36].

4 DISCUSSION

Our landscape analysis demonstrates the range of VATs commercially available and trends related to their technical development and messaging to consumers. Our results show a dominant trend—17 of the 20 VATs use machine learning. While academic researchers are well aware the application of AI to visual assistance (e.g., [24, 26, 43, 50]), to our knowledge we are the first to show the extent to which this is occurring in industry. These findings are important for several reasons. First, people who are blind are concerned about the accuracy of automated visual assistance (e.g., [15, 22, 31, 41]). Future research may compare the algorithms used by the 17 to provide visual assistance and assess how they ensure accuracy in image and video description. Second, our findings indicate that users' personal data is likely collected and stored on the cloud to improve VAT services or otherwise. Researchers, UX professionals, developers, and users alike can use our findings to evaluate whether the companies policies are in alignment with user values regarding data regulation and preferences for the preservation of their visual privacy.

Finally, our analysis of how companies describe their VATs reveals that the language of independence and help is prevalent. In contrast, recent research work has identified that "interdependence" may be a more suitable way to describe the interaction between users and assistive technologies [13] and the care work that occurs when tasks are approached collaboratively [16]. In addition, we observed that some VATs use terms such as "help" or "allow" which conflict with concepts in ability-based design [49] and challenges to normative use of language that marginalizes the abilities of people with disabilities. Future work may be conducted to evaluate whether these terms and concepts align with users' perceptions of VAT service offerings and their values surrounding the provision of assistance.

REFERENCES

[1] [n.d.]. Adobe Accessibility. https://www.adobe.com/accessibility.html
[2] [n.d.]. KNFB Reader App features the best OCR. Turn print into speech or Braille instantly. iOS 3 now available. | KNFB Reader. https://www.knfbreader.com/
[3] [n.d.]. Lookout by Google - Apps on Google Play. https://play.google.com/store/apps/details?id=com.google.android.apps.accessibility.reveal&hl=en_US&gl=US
[4] 2021. Envision App. https://www.letsenvision.com/envision-app
[5] Barak S. Aharonson and Melissa A. Schilling. 2016. Mapping the technological landscape: Measuring technology distance, technological footprints, and technology evolution. *Research Policy* 45, 1 (Feb. 2016), 81–96. https://doi.org/10.1016/j.respol.2015.08.001
[6] Tousif Ahmed, Patrick Shaffer, Kay Connelly, David Crandall, and Apu Kapadia. 2016. Addressing Physical Safety, Security, and Privacy for People with Visual Impairments. 341–354. https://www.usenix.org/conference/soups2016/technical-sessions/presentation/ahmed
[7] AIpoly. 2020. *AIpoly Homepage.* https://www.aipoly.com/
[8] Aira. 2020. Aira. https://aira.io/ Library Catalog: aira.io.
[9] Taslima Akter, Tousif Ahmed, Apu Kapadia, and Swami Manohar Swaminathan. 2020. Privacy Considerations of the Visually Impaired with Camera Based Assistive Technologies: Misrepresentation, Impropriety, and Fairness. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Virtual Event Greece, 1–14. https://doi.org/10.1145/3373625.3417003
[10] Taslima Akter, Bryan Dosono, Tousif Ahmed, Apu Kapadia, and Bryan Semaan. 2020. "I am uncomfortable sharing what I can't see": Privacy Concerns of the Visually Impaired with Camera Based Assistive Applications. https://www.usenix.org/conference/usenixsecurity20/presentation/akter
[11] Amazon. 2021. Amazon Rekognition – Image and Video Analysis – Amazon Web Services (AWS). https://aws.amazon.com/rekognition/?nc=sn&loc=1&blog-cards.sort-by=item.additionalFields.createdDate&blog-cards.sort-order=desc
[12] Apple. 2021. Accessibility - Vision - Apple. https://www.apple.com/accessibility/vision/
[13] Cynthia L. Bennett, Erin Brady, and Stacy M. Branham. 2018. Interdependence as a Frame for Assistive Technology Research and Design. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Galway Ireland, 161–173. https://doi.org/10.1145/3234695.3236348
[14] Cynthia L. Bennett, Jane E, Martez E. Mott, Edward Cutrell, and Meredith Ringel Morris. 2018. How Teens with Visual Impairments Take, Edit, and Share Photos on Social Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–12. https://doi.org/10.1145/3173574.3173650
[15] Cynthia L. Bennett, Cole Gleason, Morgan Klaus Scheuerman, Jeffrey P. Bigham, Anhong Guo, and Alexandra To. 2021. "It's Complicated": Negotiating Accessibility and (Mis)Representation in Image Descriptions of Race, Gender, and Disability. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–19. https://doi.org/10.1145/3411764.3445498
[16] Cynthia L. Bennett, Daniela K. Rosner, and Alex S. Taylor. 2020. The Care Work of Access. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–15. https://doi.org/10.1145/3313831.3376568
[17] BeSpecular. 2020. BeSpecular. https://www.bespecular.com/
[18] Hugh Beyer and Karen Holtzblatt. 1998. Contextual design: defining customer-centered systems. 1998. *San Francisco: Morgan Kauffman* (1998).

[19] Jeffrey P. Bigham, Tom Yeh, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Aubrey Tatarowicz, Brandyn White, and Samuel White. 2010. VizWiz: nearly real-time answers to visual questions. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A) - W4A '10*. ACM Press, Raleigh, North Carolina, 1. https://doi.org/10.1145/1805986.1806020

[20] Be My Eyes. 2020. Be My Eyes - See the world together. https://www.bemyeyes.com/ Library Catalog: www.bemyeyes.com.

[21] Facebook. 2021. Facebook Automatic Alternative Text. https://www.facebook.com/watch/?v=1082033931840331

[22] Cole Gleason, Patrick Carrington, Lydia B. Chilton, Benjamin M. Gorman, Hernisa Kacorri, Andrés Monroy-Hernández, Meredith Ringel Morris, Garreth W. Tigwell, and Shaomei Wu. 2019. Addressing the Accessibility of Social Media. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*. ACM, Austin TX USA, 474–479. https://doi.org/10.1145/3311957.3359439

[23] Google. 2021. Lookout by Google - Apps on Google Play. https://play.google.com/store/apps/details?id=com.google.android.apps.accessibility.reveal&hl=en_US&gl=US

[24] Danna Gurari, Qing Li, Abigale J. Stangl, Anhong Guo, Chi Lin, Kristen Grauman, Jiebo Luo, and Jeffrey P. Bigham. 2018. Vizwiz grand challenge: Answering visual questions from blind people. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3608–3617.

[25] Margot Hanley, Apoorv Khandelwal, Hadar Averbuch-Elor, Noah Snavely, and Helen Nissenbaum. 2020. An Ethical Highlighter for People-Centric Dataset Creation. *arXiv:2011.13583 [cs]* (Nov. 2020). http://arxiv.org/abs/2011.13583 arXiv: 2011.13583.

[26] Rabia Jafri, Syed Abid Ali, and Hamid R Arabnia. 2013. Face recognition for the visually impaired. In *Proceedings of the International Conference on Information and Knowledge Engineering (IKE)*. The Steering Committee of The World Congress in Computer Science, Computer, 1.

[27] Patrick Gage Kelley, Lucian Cesca, Joanna Bresee, and Lorrie Faith Cranor. 2010. Standardizing privacy notices: an online study of the nutrition label approach. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. ACM Press, Atlanta, Georgia, USA, 1573. https://doi.org/10.1145/1753326.1753561

[28] Sooyeon Lee, Madison Reddie, Chun-Hua Tsai, Jordan Beck, Mary Beth Rosson, and John M. Carroll. 2020. The Emerging Professional Practice of Remote Sighted Assistance for People with Visual Impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–12. https://doi.org/10.1145/3313831.3376591

[29] LookTel. 2020. *Instant Recognition Apps for Persons with Low Vision or Blindness*. http://www.looktel.com/

[30] Wayback Machine. 2021. Wayback Machine. https://web.archive.org/

[31] Haley MacLeod, Cynthia L. Bennett, Meredith Ringel Morris, and Edward Cutrell. 2017. Understanding Blind People's Experiences with Computer-Generated Captions of Social Media Images. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 5988–5999. https://doi.org/10.1145/3025453.3025814

[32] Aleecia M. McDonald and Lorrie Faith Cranor. 2008. The Cost of Reading Privacy Policies. (2008). https://kb.osu.edu/handle/1811/72839 Accepted: 2015-09-30T22:41:07Z Publisher: Ohio State University. Moritz College of Law.

[33] Microsoft. 2020. Seeing AI App from Microsoft. https://www.microsoft.com/en-us/ai/seeing-ai Library Catalog: www.microsoft.com.

[34] Microsoft. 2021. Microsoft Azure Cognitive Services. https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/concept-describing-images

[35] OrCam. 2020. *Never Forget a Face*. https://myme.orcam.com/

[36] OrCam. 2021. OrCam MyEye, Now With the Groundbreaking Smart Reading Feature - YouTube. https://www.youtube.com/watch?v=bbEEmc0xtvw

[37] Sullivan Plus. 2021. Sullivan Plus. https://www.mysullivan.org/

[38] Looktel Money Reader. 2020. *LookTel Money Reader for iPhone, iPod Touch and Mac*. http://www.looktel.com/moneyreader

[39] Florian Schaub, Rebecca Balebako, and Lorrie Faith Cranor. 2017. Designing Effective Privacy Notices and Controls. *IEEE Internet Computing* (2017), 1–1. https://doi.org/10.1109/MIC.2017.265102930

[40] Rachel N. Simons, Danna Gurari, and Kenneth R. Fleischmann. 2020. "I Hope This Is Helpful": Understanding Crowdworkers' Challenges and Motivations for an Image Description Task. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (Oct. 2020), 1–26. https://doi.org/10.1145/3415176

[41] Abigale Stangl, Meredith Ringel Morris, and Danna Gurari. 2020. "Person, Shoes, Tree. Is the Person Naked?" What People with Vision Impairments Want in Image Descriptions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. https://doi.org/10.1145/3313831.3376404

[42] Abigale Stangl, Kristina Shiroma, Bo Xie, Kenneth R. Fleischmann, and Danna Gurari. 2020. Visual Content Considered Private by People Who are Blind. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Virtual Event Greece, 1–12. https://doi.org/10.1145/3373625.3417014

[43] Jeremi Sudol, Orang Dialameh, Chuck Blanchard, and Tim Dorcey. 2010. Looktel—A comprehensive platform for computer-aided visual assistance. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*. 73–80. https://doi.org/10.1109/CVPRW.2010.5543725 ISSN: 2160-7516.

[44] Supersense. 2021. Supersense - AI for Blind / Scan text, money and objects. https://www.supersense.app/

[45] TapTapSee. 2020. *Blind and Visually Impaired Assistive Technology - powered by CloudSight.ai Image Recognition API*. https://taptapseeapp.com/

[46] TapTapSee. 2021. TapTapSee - Blind and Visually Impaired Assistive Technology - powered by CloudSight.ai Image Recognition API. https://taptapseeapp.com/

[47] AiPoly Vision. 2021. Aipoly Vision - Apps on Google Play. https://play.google.com/store/apps/details?id=com.aipoly.vision&hl=en&gl=US

[48] Violeta Voykinska, Shiri Azenkot, Shaomei Wu, and Gilly Leshed. 2016. How blind people interact with visual content on social networking services. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 1584–1595.

[49] Jacob O. Wobbrock, Shaun K. Kane, Krzysztof Z. Gajos, Susumu Harada, and Jon Froehlich. 2011. Ability-Based Design: Concept, Principles and Examples. *ACM Transactions on Accessible Computing* 3, 3 (April 2011), 9:1–9:27. https://doi.org/10.1145/1952383.1952384

[50] Shaomei Wu, Jeffrey Wieland, Omid Farivar, and Julie Schiller. 2017. Automatic Alt-text: Computer-generated Image Descriptions for Blind Users on a Social Network Service. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, Portland Oregon USA, 1180–1192. https://doi.org/10.1145/2998181.2998364

[51] Yu Zhong, Walter S. Lasecki, Erin Brady, and Jeffrey P. Bigham. 2015. RegionSpeak: Quick Comprehensive Spatial Descriptions of Complex Images for Blind Users. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, Seoul, Republic of Korea, 2353–2362. https://doi.org/10.1145/2702123.2702437