Speaking Web

Present by Xiang Li, Rui Liu, Zhuo Shan

Autumn 2018 INFO 360

Table of Contents

Problem Statement

The problem

Online websites is one of the major resources for people to access information. However, people who are visually impaired or blind have to put much more effort to use websites and most time end up receiving less information than normal people.

The chances for them to succeed in school or work is therefore less than normal people. Recent research shows that in the US, around 7,297,100 (2.3% of population) people are diagnosed with blindness. Over 70% of them are not employed full-time and one of the major cause to that is their inability to use technical tools including websites and apps as well as us.

Why current solutions failed

Currently, efforts have been put to help blind people to use website. The most common tool now is called screen reader. Screen readers are software applications that attempt to convey what people with normal eyesight are able to see on a display via non-visual means, like text-to-speech or sound icons.

However, it is still hard to eliminate the information gap between normal people and visually impaired people only using a screen reader.

Our literature review shows 5 things that frustrated people from using a screen reader: complicated page layout which leads to confusing screen reader feedback; conflicts between screen reader and application; poorly designed/unlabeled forms; no alt text for pictures and the 3-way tie between misleading links, inaccessible PDF, and a screen reader crash. All in all, although there are some existing solutions, we think these solutions cannot address the problem in a decent manner. (Lazar, Allen, Kleinman, & Malarkey, 2007)

What cause the problem

After having an overall understanding to this problem, we focused our research on two popular screen readers -- Google chrome reader and Voice Over from Apple to try to understand why the productions fail to solve the problem. According to our own user research, there are two main problems for both screen readers.

- The learning curve for using screen reader is too high. There is a 22 pages document for Apple's screen reader instruction that involving memorizing lots of shortcuts and usages. From our own experience, it took 2 hours to just read and memorize the instruction and we still cannot remember how to use it after learning.

- The design is not user-centered. We know that when we looking at websites, our attention is jumping from one place to another to look for the information we need. However, this is not how screen readers do. Most screen readers do is to read everything from top to bottom. In this case, it is hard for users to skip the information they do not need or to quickly skim over the whole website. The way most screen readers do is contrary to the human nature and thus cause them inefficient to use.

According to our research, there are also other minor issues causes the problem.

- The first is that some websites are not optimally design for people who are visually impaired. For instance, some websites use unclear explanation to non-textual content (video, picture), complicated website layout, or confusing tags that make screen reader hard to identify proper content.
- The other reason is the low usability of screen readers. The screen readers we investigate only help people in terms of the functionality, but failed to create a user-friendly experience.

What we will address

In our design, we plan to focus on two causes of the problem according to our research. The first one the the high learning curve, we hope to melt the instruction into the output of screen readers instead of asking our users to remember all hotkeys before being able to use the product.

More importantly, we plan to mimic how people with normal eyesight use a website by changing the way current screen reader present information. We aim to propose a new hierarchical category which split a webpage into sections. By doing so, users will first have an understanding of how this website is organize, then jump directly into the section they want instead of listening to every detail of the given webpage.

All in all, our product hope to help blind, or visually impaired, users to perceive information online in an easier and more efficient way.

Solution

The core idea for our design is to restructure current screen readers to better imitate how normal people perceive information from the websites. Instead of reading all content from top to bottom, our design focuses on a new way for users to interact with screen readers and a new structure to present the information. Our design create a hierarchy to analyze the structure of websites since normal people often obtain an overall structure of the website instead of detailed content. Any webpage will be analyzed and presented by sections, for example, menu section, content section and input section. Users will first hear the description of the highest strategy and choose which section they are going to by pressing numbers correlated to each section. Numbers will be presented in a way that is similar to telephone voice prompts -- “1 menu, 2 content, 3 input.” Users can also choose to go back to upper level or exit the page using hotkeys. The naming of sections in our prototype comes from html file but hopefully it can be more descriptive by using techniques like machine learning. When it comes to the lowest level of the hierarchy, which can be the search result page in this prototype, numbers will not be presented since there is no structural information involved and no need to make selection after hearing all sections. Users can go to the page by pressing “return” when hearing the prompt they are looking for.

Tree Map

Figure 2.1

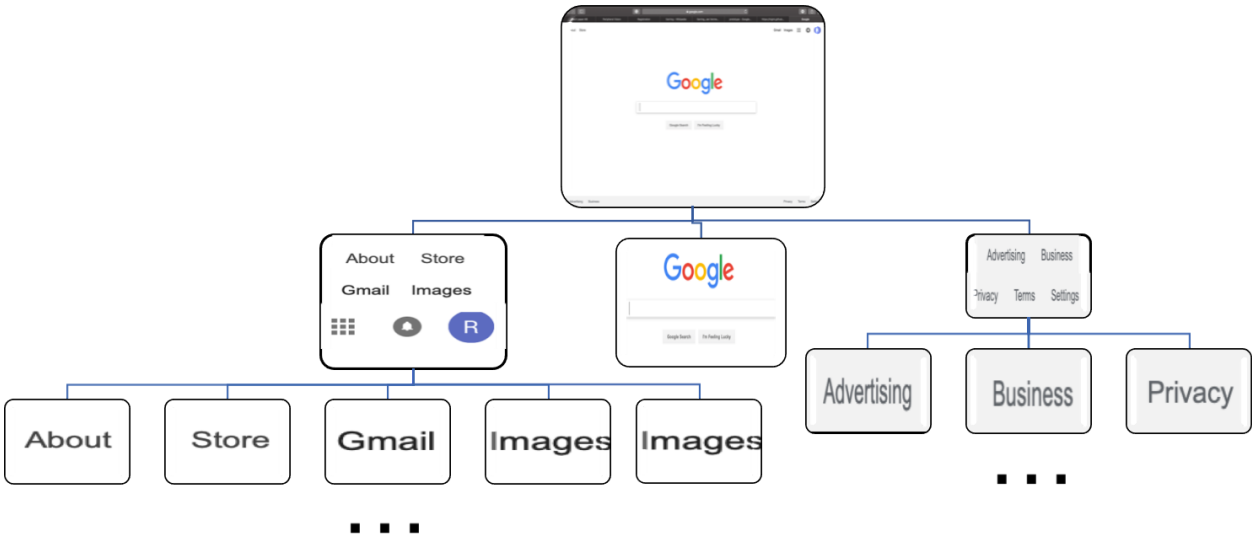

Figure 2.1 is a tree map that conclude the overall strategy we are proposing. It shows how our product will split a google homepage into sections. The first level contains 3 sections -- header, input and footer. The available sections within each level will be labeled by numbers. By pressing the numbers correlating to each section, users can go to the section they are interested in. After selecting a section, the screen reader will abandon the previous level and jump into next level to create new sections within the section selected. Again, these new sections will be labeled by numbers. To go back to upper level to switch to another section, users have to use hotkeys to achieve this action since the numbers for previous level are now used to label sections in current level only.

Prototype

This prototype presents the process of a visually impaired user trying to search for a cupcake recipe using google.

Figure 2.2

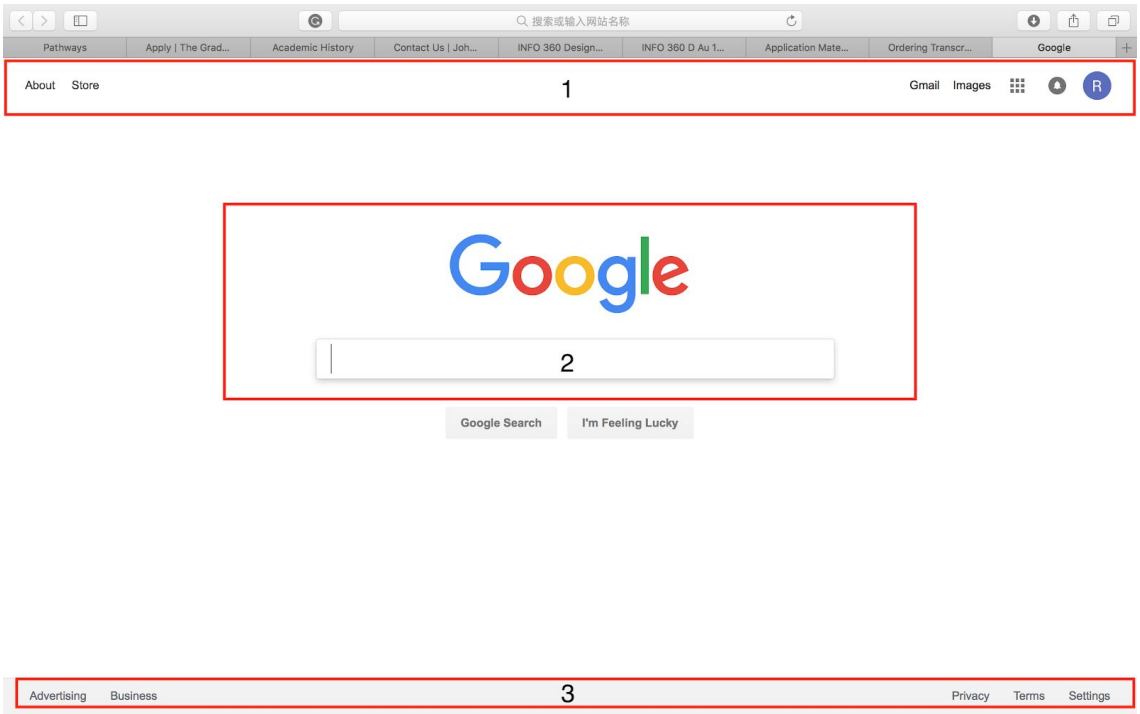

Figure 2.2 is the homepage of google. When users open this page the automatic audio output will be:

"There are 3 sections. 1 header, 2 input, 3 footer."

The prompt will keep repeating until a selection is made.

By pressing 1, users will then first hear a audio feedback of the selection, then section names in this new level.

"1 (feedback) 1 About, 2 Store, 3 Gmail, 4 Images, 5 Options, 6 Account."

By pressing 2, users will be notified the current status since there is no further levels.

"Input (feedback)."

Section 3 is similar to section 1.

"3 (feedback) 1 Advertising, 2 Business, 3 Privacy, 4 Terms, 5 Settings."

In this prototype, users will press 2 to input their research keyword.

Figure 2.3

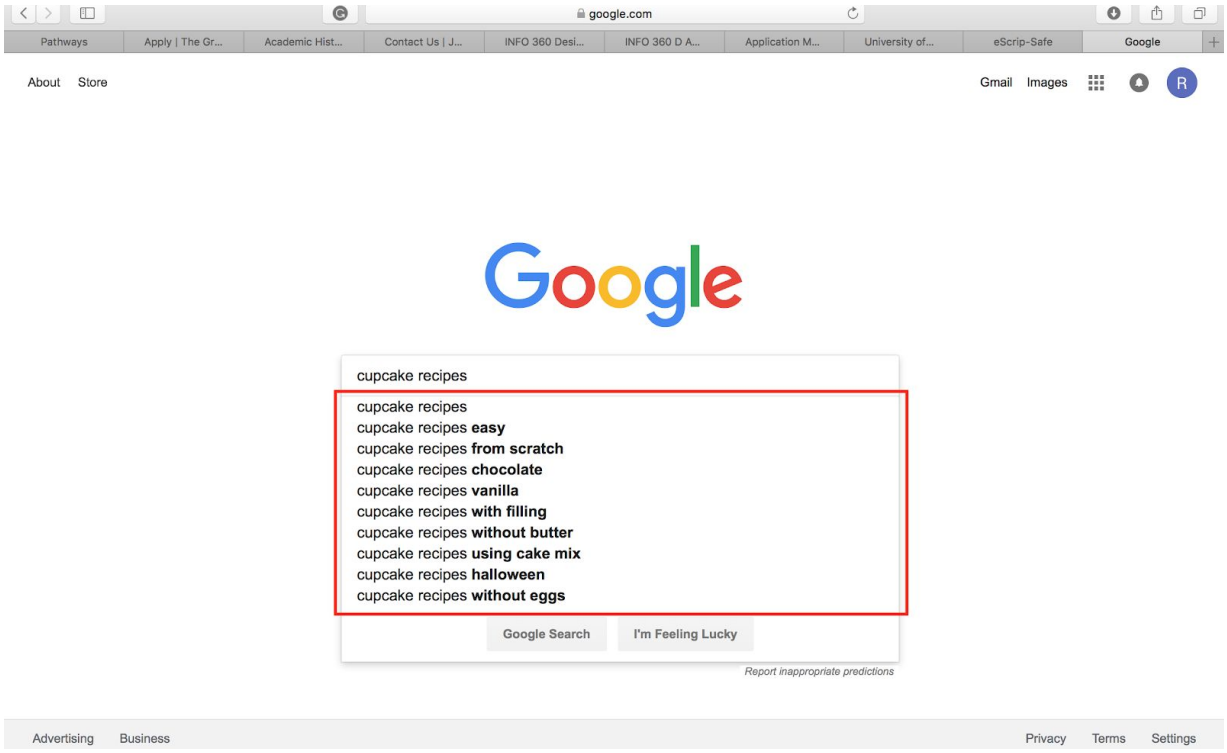

Figure 2.3 is the result of a user pressing 2 in the previous page. User can type in keyword in this status.

For this prototype, an user will type in “cupcake recipes”. There will be an real-time audio feedback of user’s input so that users can make corrections if they make a mistake.

“C-u-p-c-a-k-e space r-e-c-i-p-e-s”

Users can use arrows to listen to search suggestion and hit “enter” to go to result page. This process is identical to all current websites so that users do not have to learn any new hotkey to replace what they already know.

After hitting enter, the website will jump to result page after a audio prompt (“enter”).

Figure 2.4

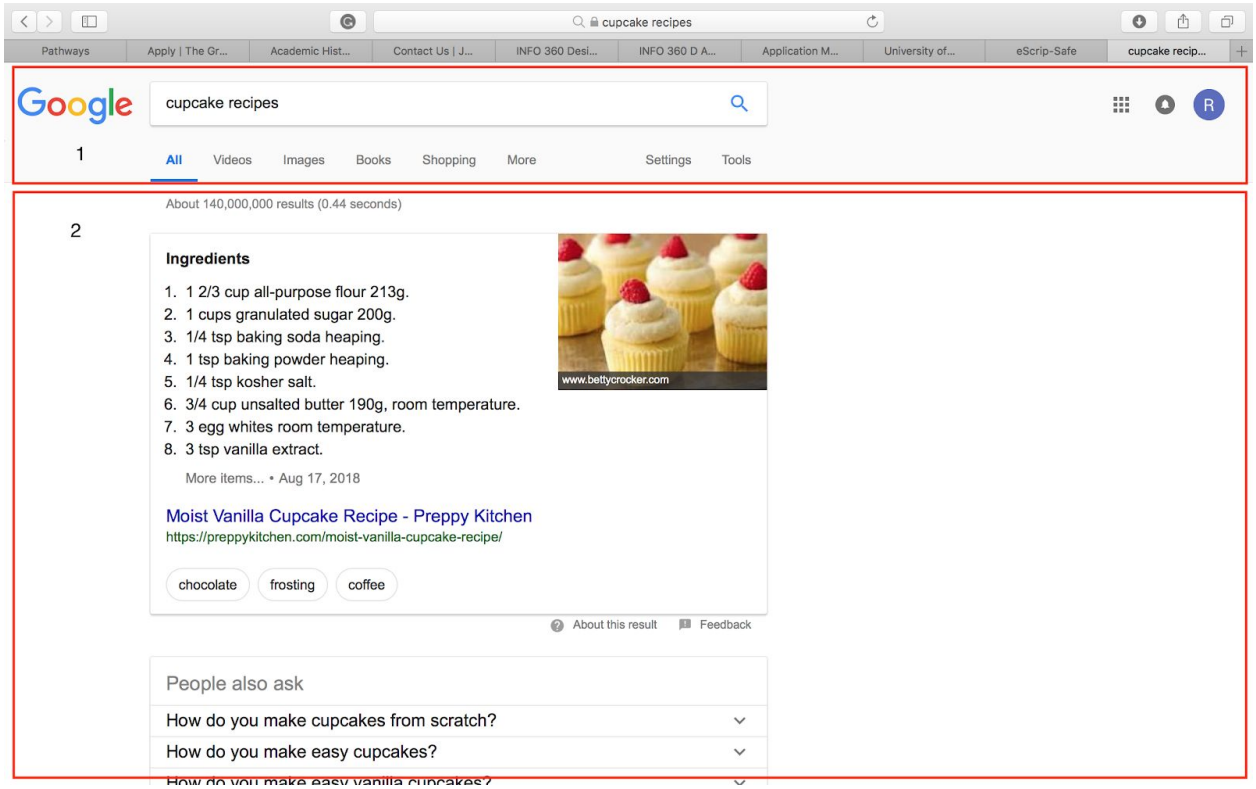

Figure 2.4 is the result page for “cupcake recipes” and is being separated into 2 sections, navigation bar and results. After the page is completely loaded, the automatic audio output will be:

“1 navigation bar, 2 results.”

By pressing 1, users will be able to interact with the web with more function, such as going back to homepage (the google icon), searching for news, images or videos. Since this section is not the targeting section for this prototype, further levels and selections will not be presented here.

By pressing 2, the result part will be further separated into sections according to results. For this specific page, the audio output will be:

“2 (feedback), 1 Moist Vanilla Cupcake Recipe - Preppy Kitchen, 2 People also ask…”

Users can select the result they are interested in and then be transferred to the corresponding website.

By selecting 1, users will enter the page for the first result.

Figure 2.5

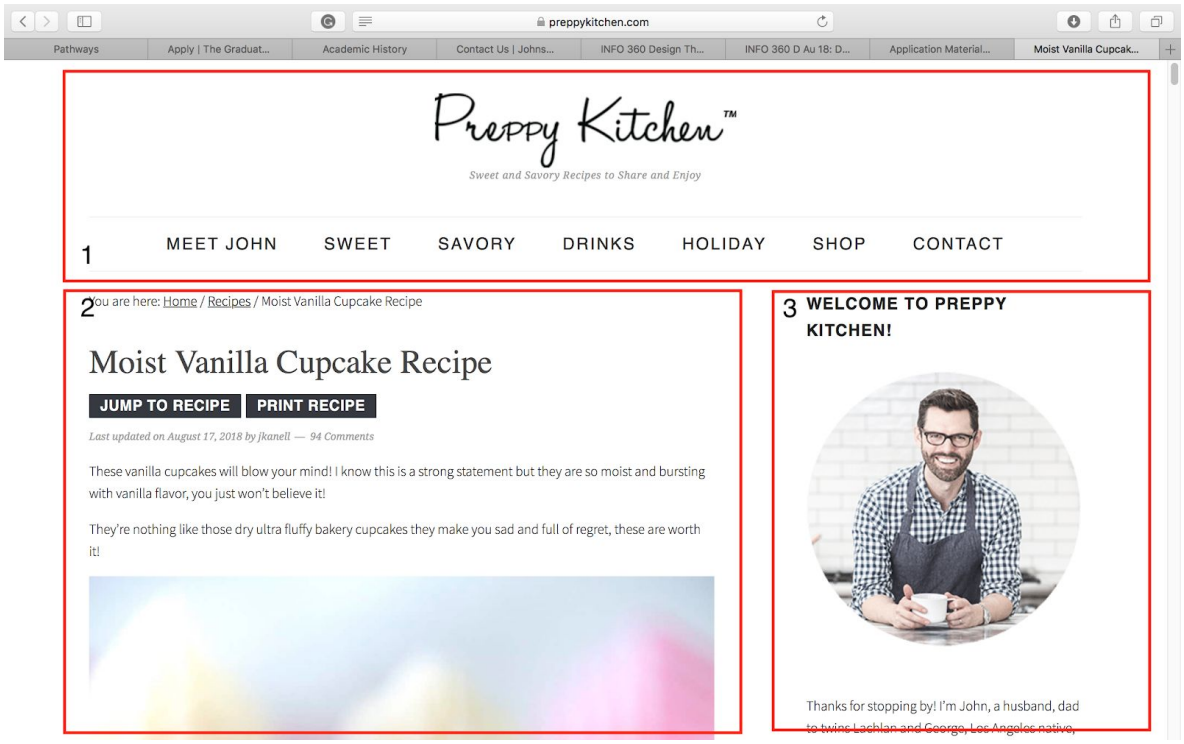

Figure 2.5 is the page for a cupcake recipe and will be separated into 3 sections. The automatic feedback will be:

"1 (feedback), 1 navigation bar, 2 content, 3 introduction"

The targeting section for the recipe should be 2, so explanation for other sections will not be included here.

By pressing 2, users will hear title- button - content. Since this section is mixed will both content and button. When presenting each information, there will be a feedback of what kind of information this is and buttons will be assigned with number to allow for selections.

"Title (information type) Moist Vanilla Cupcake Recipe, Button (information type) 1 (number assigned) JUMP TO RECIPE, Button 2 PRINT RECIPE, Content These vanilla cupcakes will blow your mind! …"

The shortcut to this structure is that users can directly jump to the ingredient part of the article so that they do not have to listen to all the content!

By selecting 1, users will be directed to the recipe section and hear automatic output of ingredients and instructions.

Figure 2.6

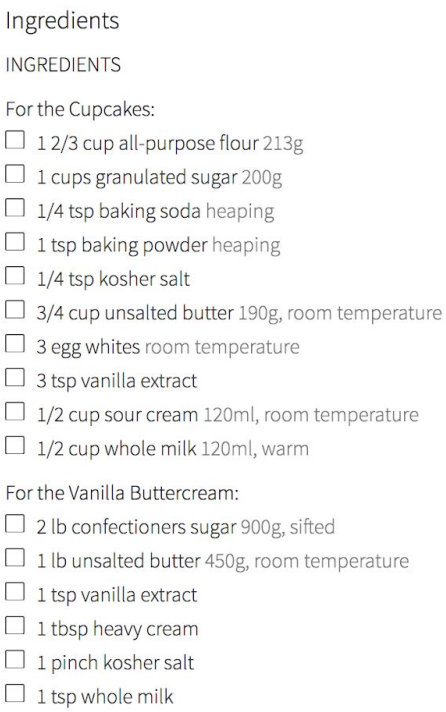

Figure 2.6 is the final goal for this prototype.

Evaluation

Cognitive Walkthrough

Since our product could not be fully built due to the professional skills required, it is hard to conduct evaluations like A/B test. Also, our product is designed for minority population which makes it hard to conduct evaluations on specific target users. Taking these into consideration, we will use cognitive walkthrough.

The focus of the cognitive walkthrough is on understanding the system's learnability for new or infrequent users. It is particularly well suited for evaluating systems that are primarily used for audio or display based such as ATMs.

During the cognitive walkthrough, group members will go through each step of the prototype and ask the following questions:

- Will users want to produce whatever effect the action has
- Will users see the control for the action?
- Once users find the control will they recognize that it will produce the effect that they want
- After the action is taken, will users understand the feedback they get so they can confidently continue on the next action?
- Will the user notice that the correct action is available to them?
- Will the user associate the correct action with the outcome they expect to achieve?

Evaluation Result

We went through the evaluation process among group members and add a few feedback functions to improve the learnability for our product. Firstly, we added audio feedback in multiple conditions to help users see the control for their input. To be more specific, we currently do not have any input feedback which might be confusing and problematic when users type the wrong number when selecting categories. In order to solve this, the screen reader will read users input just like a talking calculator. We also added audio feedback in certain conditions so that users can fully understand the consequences of the actions. For irreversible actions including closing the page and jumping to a new website, there will first be a audio feedback to inform the user what action he or she is attempting and then ask for a double check by pressing “Y” key (Yes).

We did notice that compared with current screen reader, the use of hierarchical output will result in more “switching between levels” action but we failed to mention any information of actions like “back to upper level”. So a instruction section of all related actions will be added when users learn how to use the screen reader for the first time. This will also be available to users in the “Help” section of screen reader itself. The reason we did not provide such information throughout the whole using process is due to the fact that our target users will use

the product very often so there is no need for a frequent hint. Also provide such information so frequently will reduce the efficiency of the screen reader output since out users care the most about the content that they are looking for.

Limitation:

- When the website contains non-contextual information like pictures, the output might not be as descriptive and well-organize as it is to users will normal eyesight. Hopefully is can be solved by the development of machine learning.
- A complicated website with poor html grouping may cause the screen reader produce output that is not descriptive enough.
- Users that are not familiar with keyboard may experience difficulties when looking for the right number key.
- A complicated website flow will result in a messy hierarchy which contains too many levels, thus influence efficiency and usability.

Reference

Lazar, J., Allen, A., Kleinman, J., & Malarkey, C. (2007). What frustrates screen reader users on the web: A study of 100 blind users. International Journal of human-computer interaction, 22(3), 247-269.

AFB Accessibility Resources. (n.d.). Retrieved from http://www.afb.org/info/accessibility/afb-accessibility-resources/23

Screen reader. (2018, November 21). Retrieved from https://en.wikipedia.org/wiki/Screen_reader