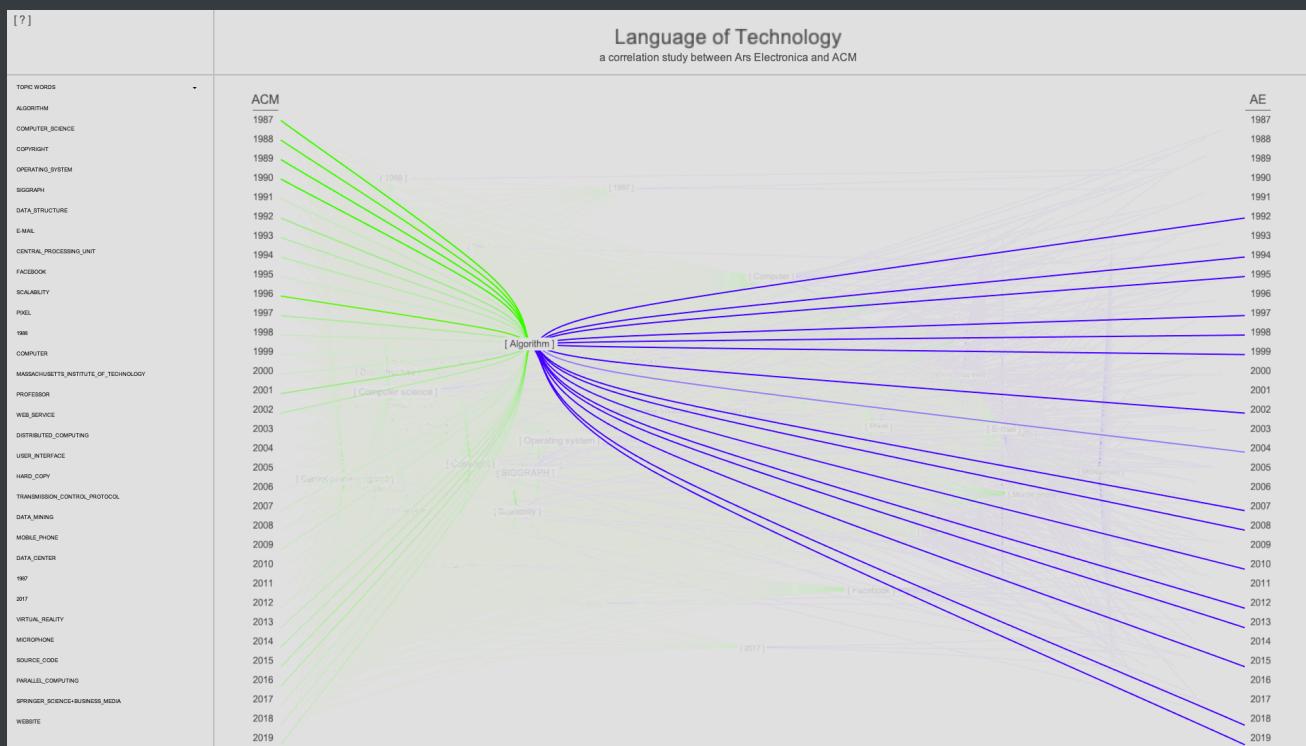


Language of Technology

by Ariella Gilmore



Language of Technology is an interactive data visualization where users can learn and analyze the historical trends that happened between Ars Electronica and the Association of Computing Machinery.

Background

Ars Electronica is an annual exhibition over the last 30 plus years where "artists, scientists, developers, designers, entrepreneurs, and activists" submit work guided by the philosophy of "never asking what technologies can or will be able to do, but always what it should do for us." The Association of Computing Machinery, on the other hand, is the world's largest computing society and strives "to raise awareness of computing's important technical, educational, and social issues around the world." Seeing how these both explore and investigate similar technologies and ideas, this visualization strives to show how media artists and technical researchers have discussed topics over time.

Data

Both these resources, AE and ACM, have digital libraries that contain information of submitted projects or papers over time. The AE archive includes artist descriptions for each awarded submission from 1987 to 2019. ACM's digital archive, which is much larger than AE, contains all the papers within the association. In order to extract this information from both archives, the websites were digitally scraped and converted into two different databases.

Analysis

With two different sets of texts it was crucial to find a way to properly compare these works. UCSB's digital humanities department has been working on a tool called the Wikifier, which made this comparison possible. The Wikifier is a tool which reads in the given text and then is able to select certain keywords or phrases and match them to a corresponding Wikipedia title. The Wikipedia titles act as a "third party system," which is able to connect these two different texts to one source, Wikipedia.

Running the Wikifier analysis led to a list of topic words (Wikipedia titles) for each database. After removing the non-overlapping topic words, it was now possible to dig deeper and learn how each topic word was being discussed from each database and how it changed over time. Using a collocation method, the analysis furthered by looking at the surrounding words when a topic was mentioned in each art description or research paper. By choosing the twenty surrounding words for each topic mention and further filtering by choosing the top twenty surrounding words from each database, the analysis starts to gain a larger perspective of how each topic was being discussed over time and highlights whether or not AE and ACM were discussing the topics similarly or differently.

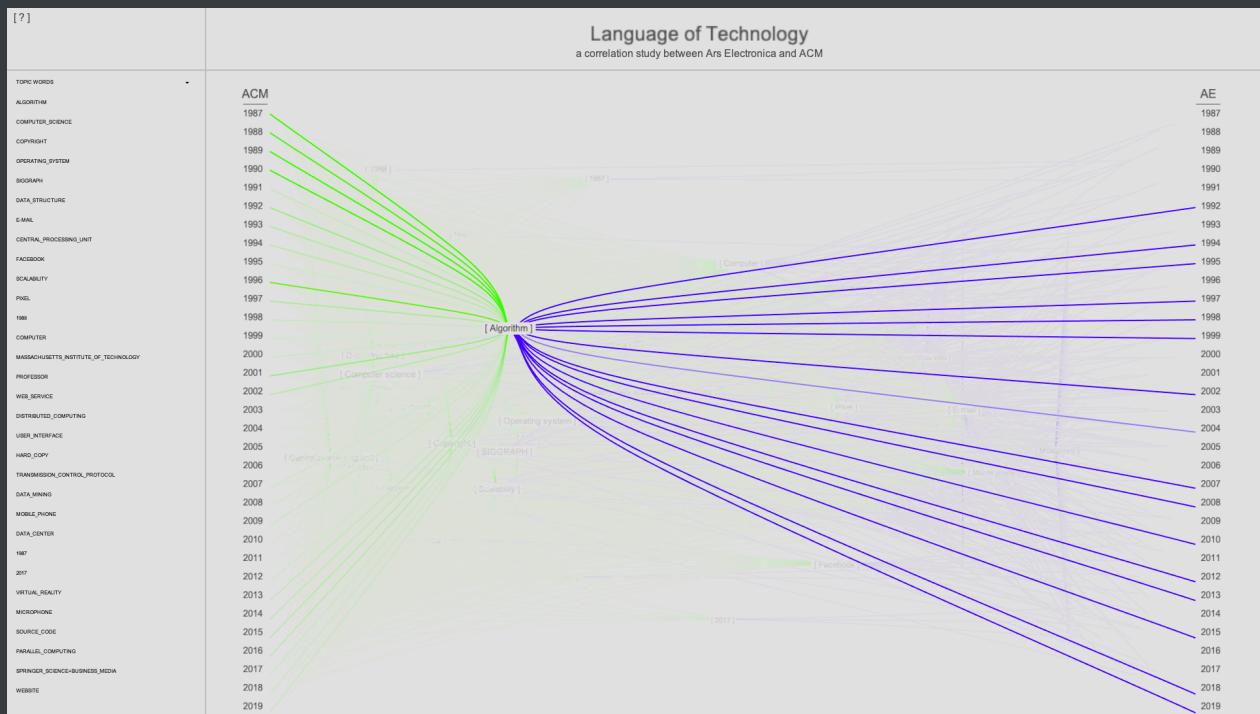
| | A | Y | Z | AA | AB | AC | AD | AE | AF | AG | AH | AI | AI | AK | AL | AM | AN | AO | AP | AQ | AR | |
|--|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----|----|
| Algorithm | 2010 ae | 2011 ae | 2012 ae | 2013 ae | 2014 ae | 2015 ae | 2016 ae | 2017 ae | 2018 ae | 2019 ae | 1987 acm | 1988 acm | 1989 acm | 1990 acm | 1991 acm | 1992 acm | 1993 acm | 1994 acm | 1995 acm | 1996 acm | | |
| "Algorithm": (⊕) | | | | | | 1 | | | | 2 | 10 | 18 | 10 | 23 | | | | 16 | 104 | 9 | | |
| "Computer_science": (⊕) | | | | | | | | | | 7 | 590 | 560 | 507 | 866 | 132 | 250 | 49 | 368 | 104 | 550 | | |
| "Copyright": (⊕) | | | | | | | | | | 1 | 2 | 2 | 1 | 10 | 14 | 29 | 65 | 6 | 41 | 12 | | |
| "Operating_system": (⊕) | | | | | | | | | | 1 | 1 | 1 | 1 | 8 | 8 | 5 | 5 | 6 | 6 | 7 | | |
| "SIGGRAPH": (⊕) | | | | | | | | | | | 2 | 1 | 59 | 52 | 76 | 40 | 11 | 46 | 39 | 2 | 16 | |
| "Data_structure": (⊕) | | | | | | | | | | | | 4 | 245 | 154 | 311 | 142 | 28 | 35 | 17 | 30 | 36 | |
| "E-mail": (⊕) | | | | | | | | | | | | | 245 | 154 | 311 | 142 | 28 | 35 | 17 | 30 | 36 | |
| "Facebook": (⊕) | | | | | | | | | | | | | | 245 | 154 | 311 | 142 | 28 | 35 | 17 | 30 | 36 |
| "Central_processing_unit": (⊕) | | | | | | | | | | | | | | | 2 | 2 | 1 | 2 | | | | |
| "Scalability": (⊕) | | | | | | | | | | | | | | | | | | | | 9 | | |
| "Pixel": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "1988": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Computer": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Massachusetts_Institute_of_Technology": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Professor": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Hard_copy": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Web_service": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Distributed_computing": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "User_interface": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Data_mining": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Transmission_Control_Protocol": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Mobile_phone": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Data_center": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "2017": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "1987": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Virtual_reality": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Microphone": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Source_code": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Parallel_computing": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Springer_Science+Business_Media": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "XML": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Website": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Application_programming_interface": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Database": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Electronic_commerce": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Laptop": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "World_Wide_Web": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Information_systems": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "1990": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Blog": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Human_factors": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Cambridge_Massachusetts": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "File_system": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Dissertation": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Workstation": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Microsoft": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "OpenGL": (⊕) | | | | | | | | | | | | | | | | | | | | | | |
| "Twitter": (⊕) | | | | | | | | | | | | | | | | | | | | | | |

With this, the analysis now includes information of the frequency of overlapping topics as well as including much finer detail of the word frequencies that are discussing these topics. All of this information has exactly which year and which AE art description or ACM paper that it's being pulled from. The data can be structured in a way, so that it can both show how AE and ACM were discussing each topic as well as show the evolution over time.

Visualization

The visualization includes four different modes, which allows the user to properly navigate the space and be able to explore how these topics changed over time.

- First Mode



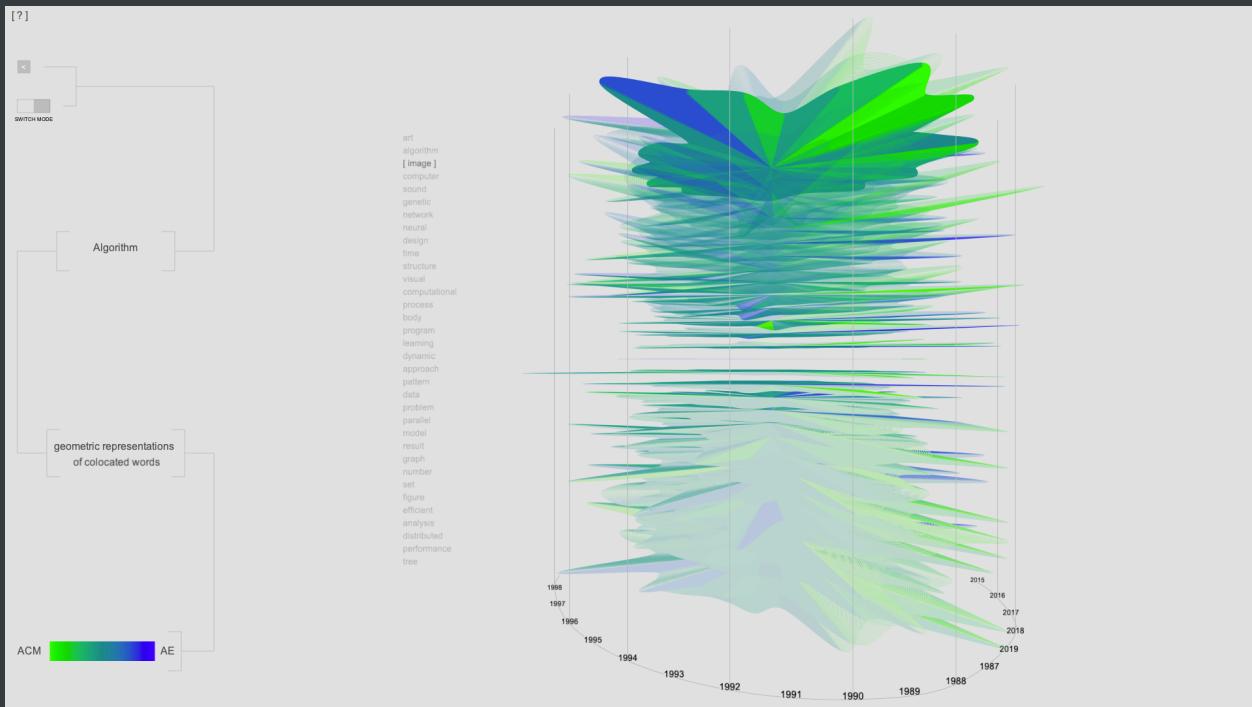
- Display

- The first screen displays the overlapping topics discussed above and positions each topic in an x and y axis where the x axis corresponds to ACM vs AE and y axis corresponds to years.
- Using the spring and pulley method in section 3.9 in [The Nature of Code](#), each topic word has springs attached to each year for both ACM and AE. The spring force is calculated based on the year frequencies from each database of each topic word. This allows the topic word to fall in the average x and y coordinate that it's being pulled in.
- Each topic word is separated by an off set in the z axis and ordered by the most frequently mentioned topic words starting with "Algorithm."

- Interaction

- One can either click on the left scrollbar to view that particular topic word or click within the visual itself.
- In order to move on to the next mode, one must double click on a topic word within the scrollbar.

- Second Mode



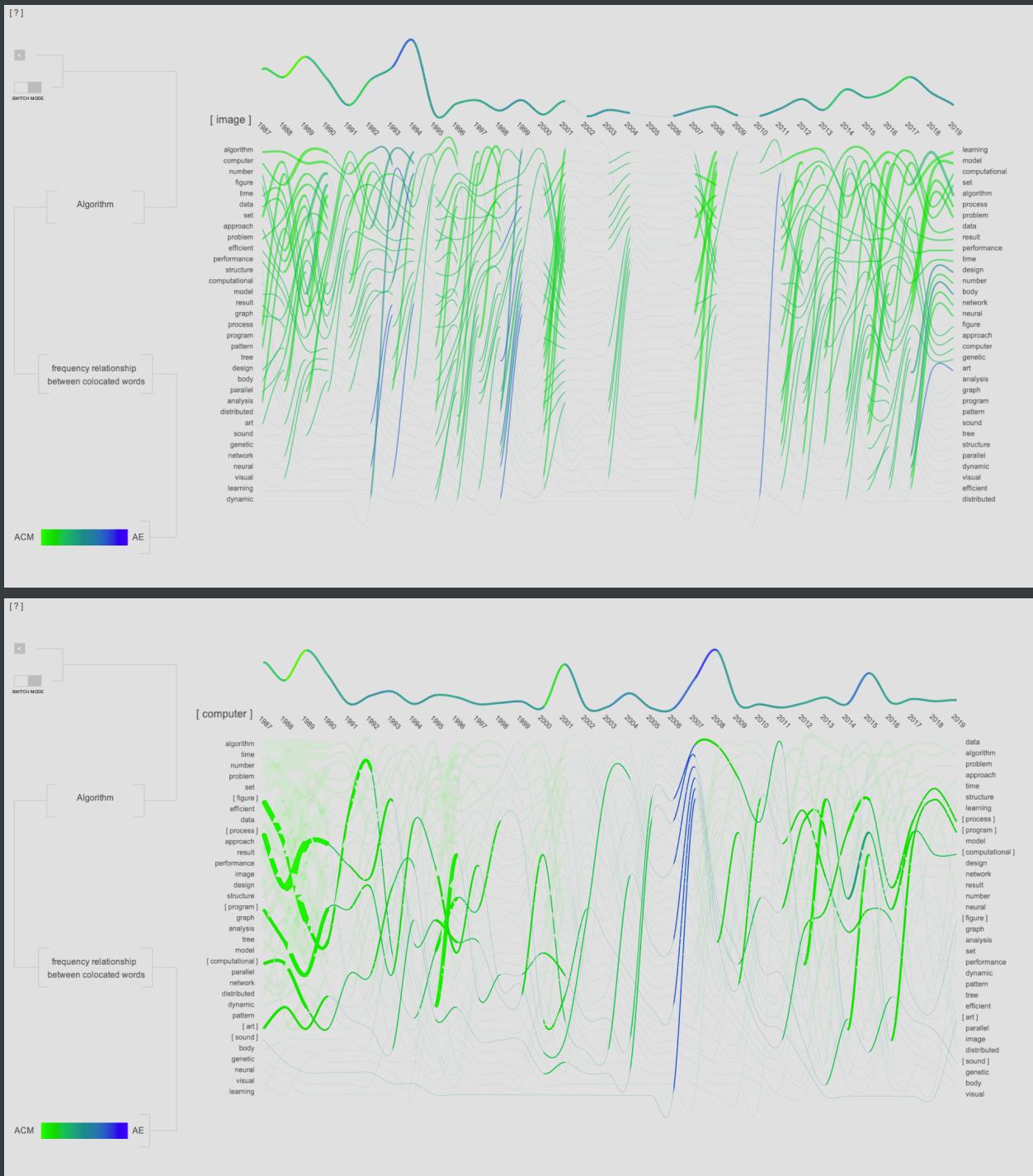
- Display

- This mode now displays the descriptive words of the topic that you previously chose. In an arbitrary order, the words are displayed vertically lining up with their resulting geometric representation.
- The shapes are made up of each words frequency when mentioned with that particular topic and each frequency is arranged in a radial timeline axis. Applying a color to each frequency from each database is represented in the bottom left gradient. This allows users to see spikes in color when one dataset is more frequent and a more neutral color when the two datasets have similar frequency.

- Interaction

- One can hover over each word to better reveal the shape and use left and right arrow keys to rotate the shape and explore the spikes over time.
- In order to learn more about each descriptive word, one can click on the text.

- Third Mode



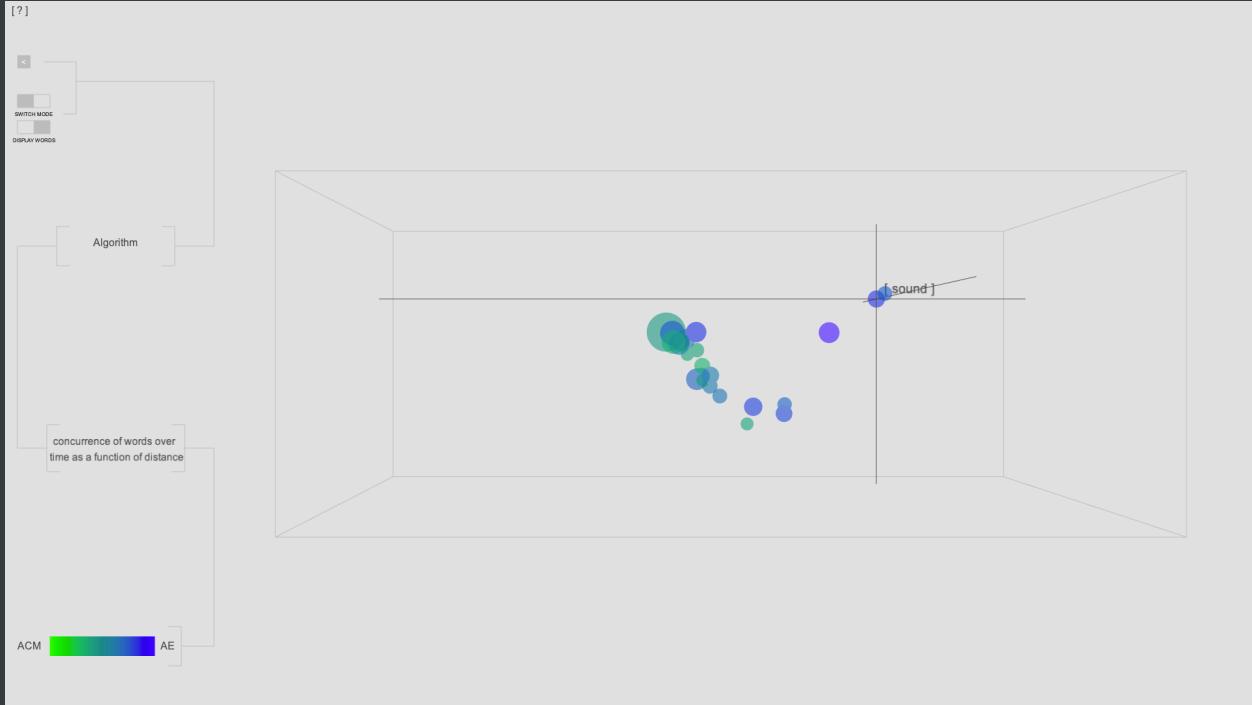
■ Display

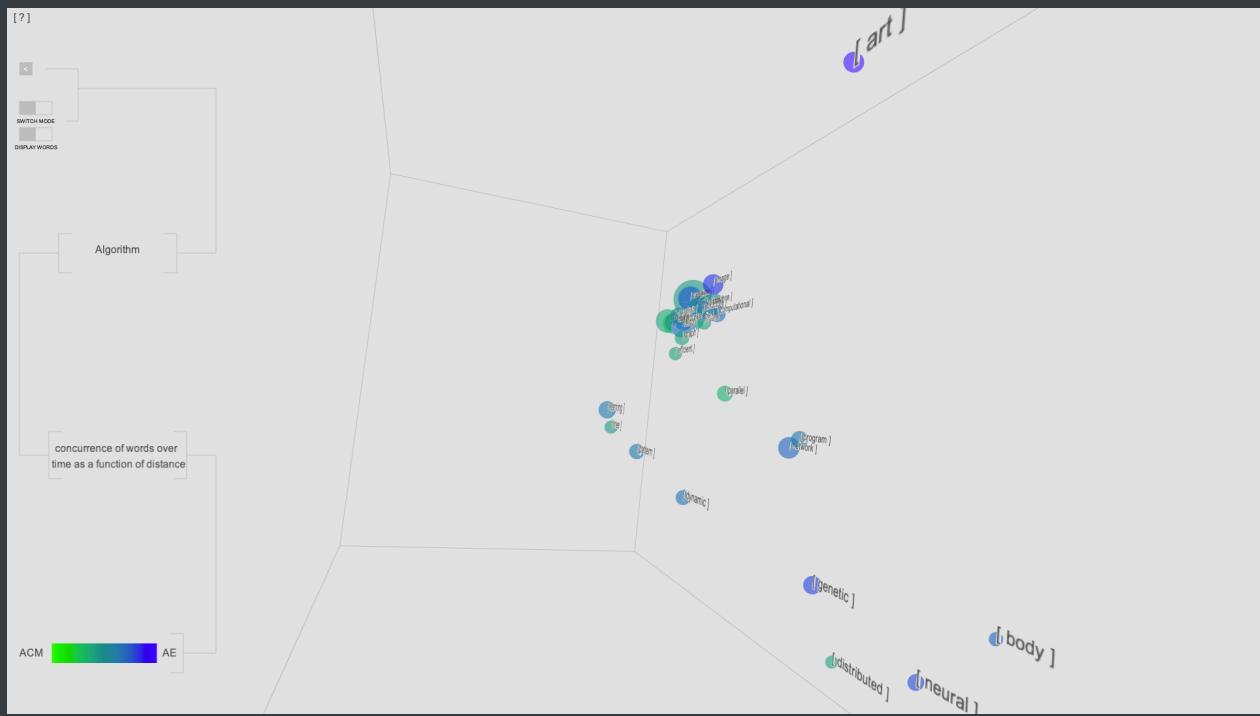
- This visual unravels the previous shape in order for the user to better see how a descriptive word changes over time.
- The color stays the same as the previously mentioned mapping, but now the alpha will decrease to illustrate that neither ACM or AE mentioned the particular word in the context of this topic word in that year.
- Below the unraveled shape, the other descriptive words are shown. The frequencies of these words show in what year the current descriptive word overlapped with the others. The order of the other descriptive words change over time with the top being that they frequently

appeared together. The thickness of each line also represents this by having the frequency these words appeared together adjust the stroke weight.

- Interaction
 - One can click to select or deselect a particular word's line frequency to learn how that frequency lines up with the selected descriptive word. Hovering over the selected line will also clearly reveal which word the user is looking at.
 - Using right click, the user can click on a particular word's line frequency at a specific year and it will open a new tab in their browser with the corresponding ACM or AE articles that contain the topic word, descriptive word, and supplemental descriptive word at the specific year clicked.
 - The topic word on the left is also clickable and returns the user back to the previous geometric shape mode.

- Fourth Mode





- Display

- Places the descriptive words in a 3D space where concurrence of words over time act as a function of distance. This is using a PCA algorithm, which reduces the large dimension of data being shown to just 3D.
- Each dot is the average of a descriptive word over time, so the larger the dot means the more frequent the word was used within the specific topic. The color represents the average of which database it came from with the color scale staying the same as previously mentioned.

- Interaction

- The previous two modes include a toggle which when turned on bring you to this final mode.
- Another toggle is added in this mode, which displays the words on top of each point. One can also hover to view the word and see the x, y, and z axis in which that point is placed.
- Each dot is clickable and will return you back to the previous mode of the unraveling shapes.

- All Modes

- Interaction

- Top-left "?" can be hovered over and explains the interactions within each mode.
- Pressing the "r" key will reset the camera distance intended for that mode.

Conclusion

With this unique data, new visualizations are created where both the data and the visual are of equal importance. This structure allows users to understand the history of topics as well as further learn how each topic changed over time. As the user begins analyzing the visualization and understanding the connections to history it is making, then the user can open up particular ACM papers or AE descriptions

and better understand how these texts arrived at their particular moment in history.

The previous details on the visualization are innovative in the fact, that no other visualizations are able to show these trends in history over time when looking at a text based system comparing two different datasets. This innovation leads the user to understanding topics over time and being able to use a tool where specifically new comers to Ars Electronica will not be able to use the current implemented archive, but can use a tool like this and initially start by looking at things they are interested and then link to the particular AE projects and gain a perspective of how AE has changed over the last 30 years in comparison to ACM.