Milestone 2: Sketchbook

Our project aims to make academic research in Computer Science (abbreviated CS) more visible and accessible by presenting a comprehensive, interactive overview of the field. We want users—students, educators, or curious minds—to quickly grasp key areas, influential papers, notable authors, and leading publication venues.

1 Core Visualization (Minimal Viable Product)

We begin with a timeline visualization (see Figure 1) highlighting major breakthroughs in computer science. Each milestone includes a brief description revealed on hover, offering users a structured, chronological entry point into the field.

Once this historical foundation is established, we shift focus to the contemporary research landscape. The next visualization (see Figure 2) illustrates the rise and fall of computer science subfields over the past two decades (2000–2022, based on our dataset). This view makes academic trends clearly visible—for example, the notable ascent of Machine Learning.

To enable more detailed exploration, a field-subfield selection interface (see Figure 3) allows users to focus on specific domains. Clicking a field button dynamically updates the subsequent visualizations. In particular, some subfields may appear under multiple fields, emphasizing their interdisciplinary nature.

Based on the selected area, users can view the most cited papers and authors within a specified timeframe (see Figure 4). Key publication venues are also highlighted, either via a bar chart (see Figure 5) or a Sankey diagram (see Figure 6), which maps where top-cited authors publish their work.

Several of these visualizations are already live on the website, though some are currently populated with placeholder data for demonstration purposes and may not yet follow the described narrative sequence.

2 Additional Ideas

One potential extension is a connectivity graph of papers, visualizing reference or citation networks. This could serve as a valuable tool for researchers by surfacing related papers and providing a broader context around a given work (see Figure 7). We also see potential in expanding this concept to represent entire subfields rather than focusing on individual papers. But this addition introduces greater complexity—not only visually but also in terms of the data preprocessing required. As such, it will demand significant more development effort and computational resources.

Another idea is an "ancestral gallery" of Turing Award winners (see Figure 8), which would further enrich the historical context. Each laureate would be represented by a portrait, with a brief summary of their contributions displayed on hover. Clicking the image could redirect users to the official ACM Turing Award page for more detailed information.

3 Tools and Technologies Used for Visualizations

For our visualizations, we will use the D3.js library to create dynamic, interactive data graphics. To simplify styling and layout, we will use Tailwind CSS instead of plain CSS.

We have already set up the project repository with Node.js, which gives us a development environment with features such as live reloading.

To maintain consistent code style, we use Prettier for code formatting. Linting rules are enforced via our GitHub CI pipeline, which also handles automatic deployment of commits to the main branch directly to the website hosted on GitHub Pages.
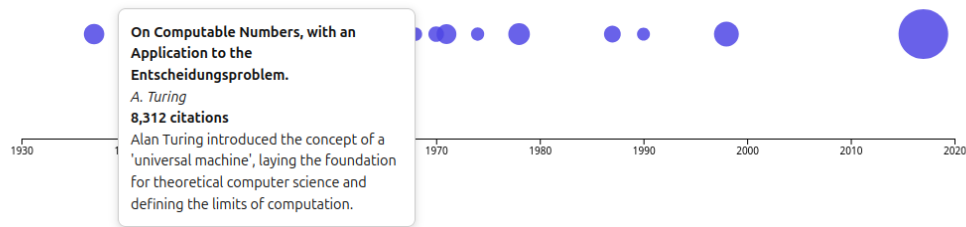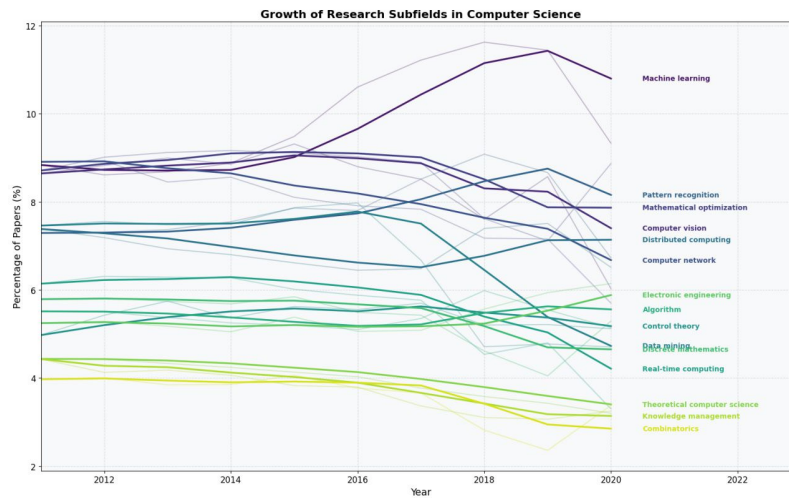
Figure 1: Timeline of major breakthroughs

Figure 2: Growth of research fields

4 Relevant Lectures

Dos and Donts cover important general advice we will try to follow. The lectures about D3.js and interactions with D3.js as well as their exercises are helpful in using the chosen library. To show the relationships between authors or to have reference graphs the future lectures about graphs could be useful.
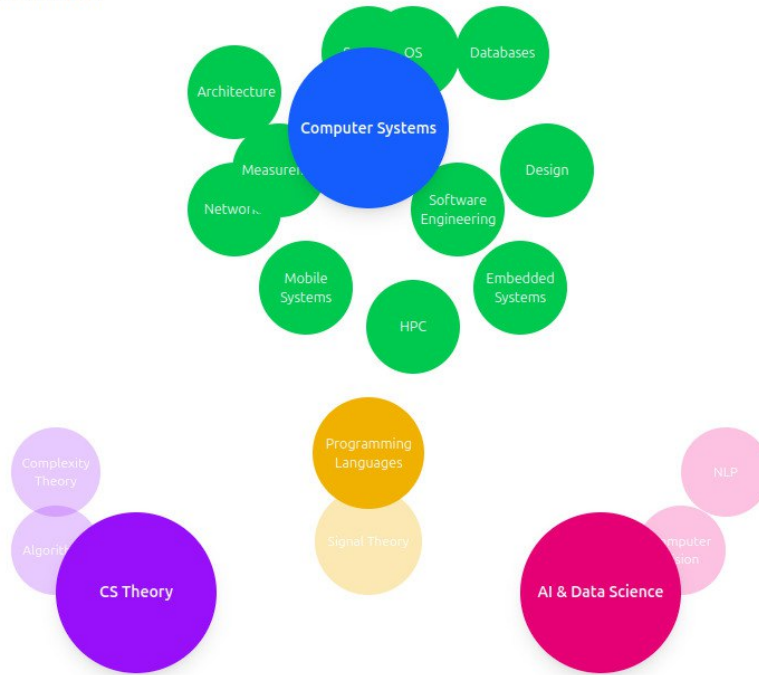
🗺️ **Main Fields**

Figure 3: Fields and subfields of Computer Science

📚 **Top Papers**

1.
Fast-track cardiac care for adult cardiac surgical patients.
Citations: 235

2.
Environmentalism in surgical practice.
Citations: 46

3.
Embedded Computer Systems: Architectures, Modeling, and Simulation
– 24th Int. Conf., SAMOS 2024
Citations: 43

4.
A systematic review of large language models and their implications in
medical education.
Citations: 41

5.
Protocol for national monthly survey of smoking and alcohol use in
England, Scotland & Wales.
Citations: 38

👩‍🔬 **Top Authors**

Dr. Alice Thompson
Total Citations: 3,250

Prof. John Wu
Total Citations: 2,980

Dr. Maria Reyes
Total Citations: 2,540

Dr. Omar Farouk
Total Citations: 2,310

Dr. Helena Fischer
Total Citations: 2,190

Figure 4: top papers and authors
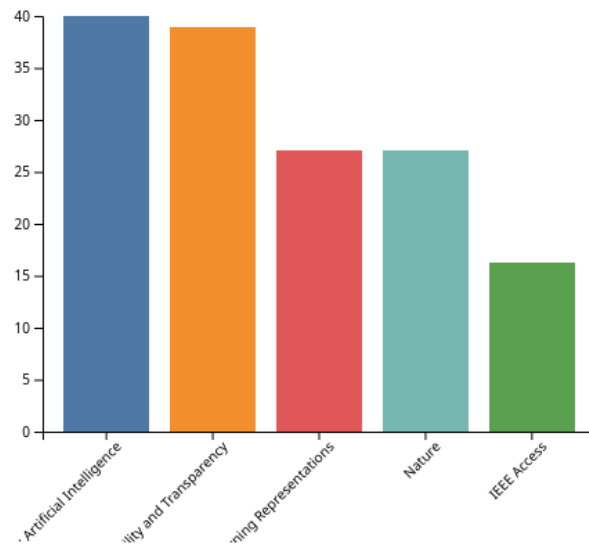
Figure 5: Top publication venues
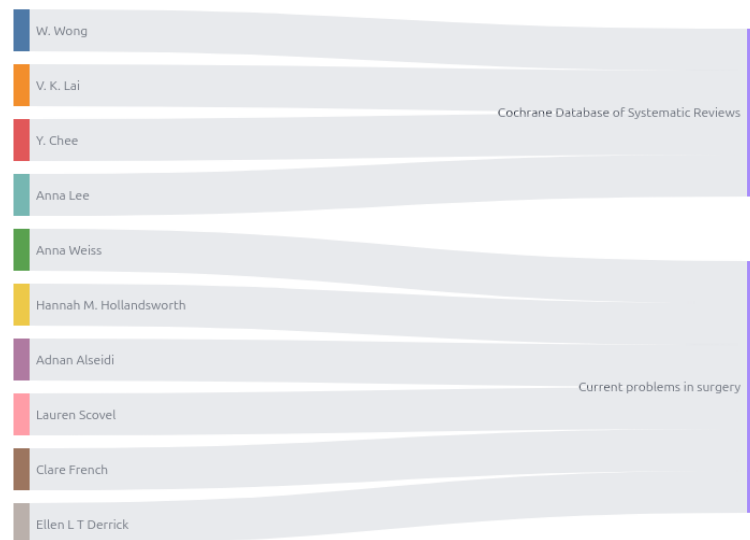
Figure 6: Where most cited authors have published
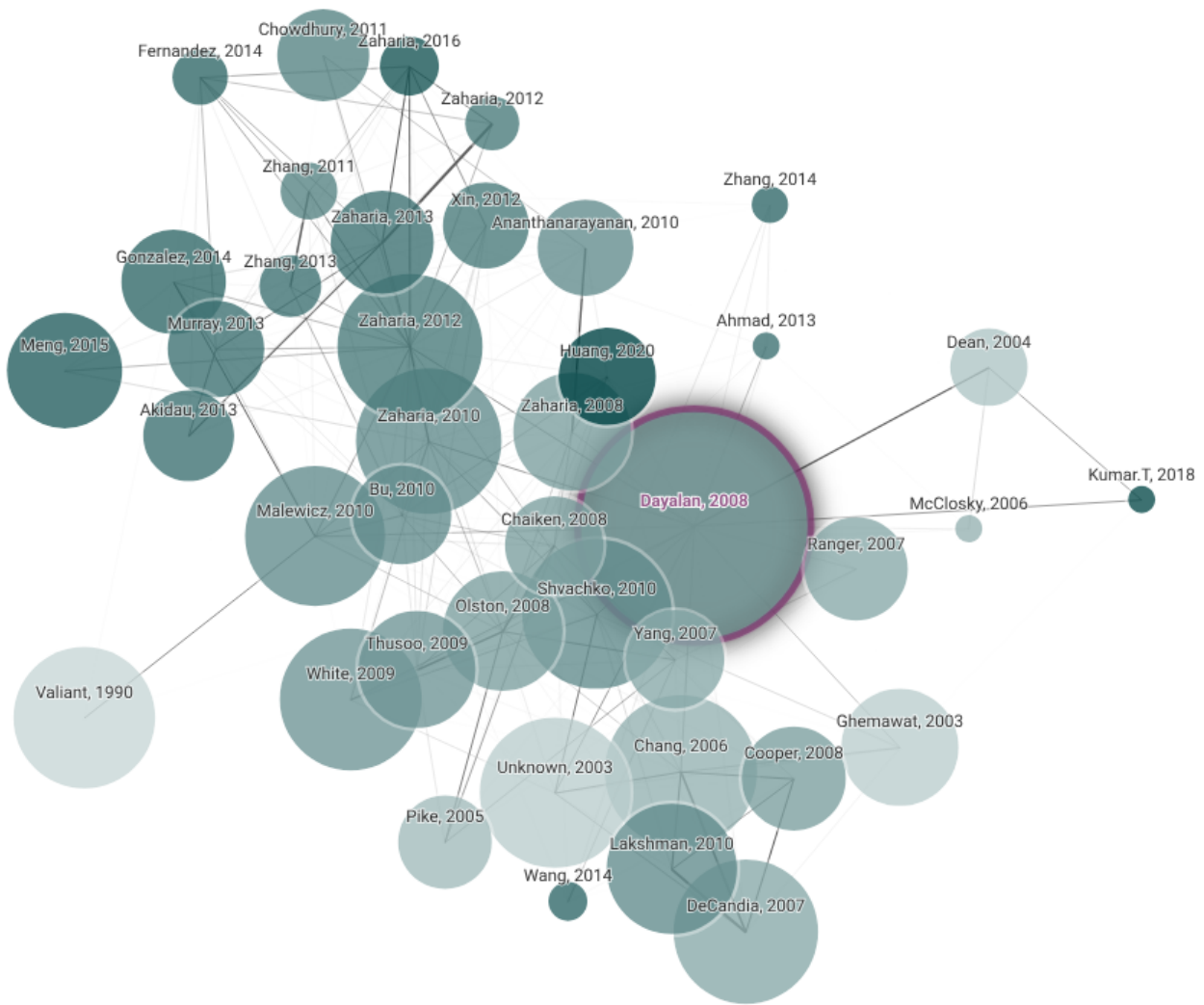
Figure 7: Connectivity graph of papers (screenshot from https://www.connectedpapers.com)

Figure 8: Ancestor gallery (screenshot from https://amturing.acm.org/)