OED Labs: exploring the Oxford English Dictionary's prototype tools

Presented by

James McCracken
Language Engineering Manager

Tania Styles
Editor: Revision and Etymology

Please note that this presentation will be recorded, including the Q&A
The OED Labs initiative will help to shape the future of the Oxford English Dictionary, and we believe the research community is an integral part of this evolution.

The wealth of language data in the OED can currently be searched only through the oed.com website. Our aim is to offer researchers new, more direct and more flexible ways to access this massive curated dataset and to gain richer insights in the English language than ever before.

OED Labs invites scholars to contribute to the creation of new functionalities for the exploration of the OED dataset.

We want to work with those at the cutting edge of academic research across all disciplines to review and refine new ideas for the benefit of this community.

Through OED Labs, researchers will help shape the development of new digital tools that harness the power of the OED and help us to reimagine how this unique resource can be shared and accessed to aid the evolution of academic research.
<table>
<thead>
<tr>
<th>Experiment</th>
<th>Sort by date</th>
<th>Sort by popularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>App Inventor for Android</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Google Books Ingram Viewer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Image Swirl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Google Reader Play</td>
<td></td>
<td></td>
</tr>
<tr>
<td>City Tours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Google Shared Spaces</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast Filp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tachlew</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other experiments at Google

- Calendar Labs
- Labs FAQ
- Labs alumni
- Stay informed
- RSS
- Google gadget

Meet the engineers

- David Petrou: Worked on **Google Maps**, the visual cortex of the internet.
- Sharon Perl: Worked on **Phone Invented for Android**, Non-programmers have great app ideas, and she wants to make them real.
- Charles L. Chen: Worked on **Android Text-To-Speech**, His work speaks for itself.
- Yang Li: Worked on **Gesture Search**, Yang builds interactive systems that make computers and data more accessible to users.
- Trevor Johnstone: Worked on **City Tours**, By day, Trevor keeps Google’s systems running but by night, he’s a tour guide for the world.

Labs favorites

1. Google Body
Labs

Twitter Developer Labs

Test new Twitter API previews, provide feedback, and get the latest news on what we’re building.

Join Twitter Developer Labs

What is Labs?

Twitter Developer Labs is where you’ll have early access to new API endpoints, features and versions. We’ll use Labs to test out new ideas and invite our developer community to share their feedback to help shape our roadmap.

Why sign up?

- Test out the latest API endpoints and developer tools from Twitter.
- Provide feedback to help shape the future of Twitter’s APIs.

You will need a developer account to use Labs. Apply now to get started.

Sign up to learn more.

Be the first to know when we release early versions of new endpoints and features in Twitter Developer Labs. We’ll send occasional promotional emails to keep you informed and ask for product feedback.
OED Labs principles

- experiment and prototype in public
- engage and collaborate with real users
- use feedback to drive development of products like oed.com
- discontinue experiments that don't work or don't engage users
OED Labs projects

- OED API
- Text visualizer
- Etymology visualization
OED Researcher API

We believe that the OED should be able to serve all scholars who want to investigate the sociological, historical, economic, or geographical record of the English language or use this data to fuel research in the wider humanities disciplines.

Our goal is to further develop the OED’s offering to actively support the needs of academic research as they evolve in the coming years.

The OED Researcher API will provide unprecedented access to the OED’s huge and uniquely curated English language dataset.

https://languages.oup.com/research/oed-researcher-api/
https://languages.oup.com/research/oed-researcher-api/
Why visualize etymologies?

- Users value the OED's etymologies highly.
- They would like to do more with the data.
- High demand for tools that create etymology visuals.
- OED is well-placed to improve on existing tools.
1. Clarification of etymology text

"Summarize the key points of this etymology for me."

**Problem:** User finds the **Origin** statement insufficient but can’t see the wood for the trees in the main **Eymology** text.

**Solution:**

- Generate simple map of main transmission.
- Use alongside etymology text to aid accessibility.
- Level of detail intermediate between **Origin** & full **Etymology**.
- Gives the gist of the etymology at a glance.
1. Clarification of etymology text

‘Summarize the key points of this etymology for me.’

**Problem:** User finds the **Origin** statement insufficient but can’t see the wood for the trees in the main **Etymology** text.

**Solution:**
- Generate simple map of main transmission.
- Use alongside etymology text to aid accessibility.
- Level of detail intermediate between **Origin** & full **Etymology**.
- Gives the gist of the etymology at a glance.

**Features:**
- Timeline format, moving forward in time L to R.
- Lexical items colour-coded by language.
- Etymological relationships shown by coloured arrows.
- Specific processes indicated in text below.
- Uncertainty shown by faded arrows & text above.
2. Tracing a word’s etymology across multiple entries

‘I want to trace the etymology of this word right back to its ulterior etymon.’

**Problem:** Transmission pathway of a single word is often split across a number of OED entries. Piecing the information together by clicking between cross-references is onerous.
2. Tracing a word’s etymology across multiple entries
2. Tracing a word's etymology across multiple entries

**Problem:** Transmission pathway of a single word is often split across a number of OED entries. Piecing the information together by clicking between cross-references is onerous.

**Solution:** Automate the process, presenting entire transmission pathway in a single visual.

Information from OED entries:
- melon, n.1
- melopaeon, n.
- melo, n.1
- pepin, n.
- coak, n.1
3. Exploring etymological connections between words

‘I’m interested in this word in an OED etymology, and want to know which English words are derived from it.’

- Follows transmission pathways *forwards* from a node, to identify words that include it in their transmission pathways.
- Enables user to extend multiple pathways forward to create their own map of etymologically-related words.
- They can then use this map to explore the OED, navigating between entries via clickable links.
https://languages.oup.com/research/oed-labs/

Thank You

Questions?