From black box to glass box: Making UMAP interpretable with exact feature contributions

We transform UMAP from a black box into a glass box. By learning the embedding function with a certain type of deep network, we can compute equivalent linear mappings of the input features that exactly reconstruct each embedding, revealing the heretofore hidden logic of UMAP.

The **full pub** is available here.

The **source code** to generate it is available in this GitHub repo (DOI: 10.5281/zenodo.17478720).

In the future, we hope to host notebook pubs directly on our publishing platform. Until that's possible, we'll create stubs like this with key metadata like the DOI, author roles, citation information, and an external link to the pub itself.

Contributors (A–Z)

- **Audrey Bell**: Critical Feedback
- **James R. Golden**: Conceptualization, Formal Analysis, Investigation, Software, Visualization, Writing
- **Evan Kiefl**: Validation
- **George Sandler**: Critical Feedback, Visualization

- **Ryan York**: Supervision