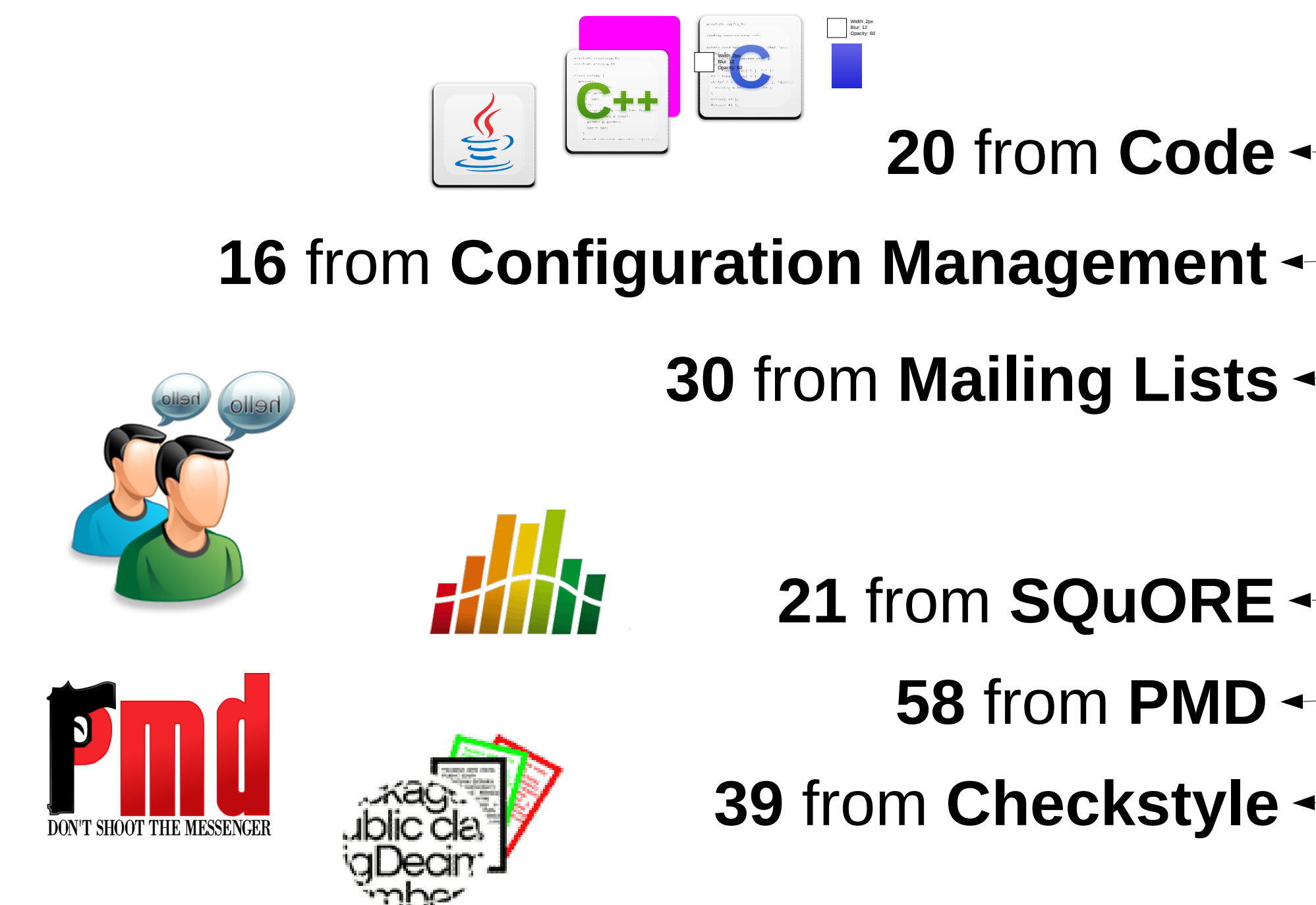


Understanding Software Evolution : The Maisqual Open Data Sets

Boris Baldassari, Philippe Preux, Université de Lille 3, LIFL (UMR CNRS) & INRIA, Villeneuve d'Ascq, France

We propose **data sets** composed of ...



... **66 Metrics** ...

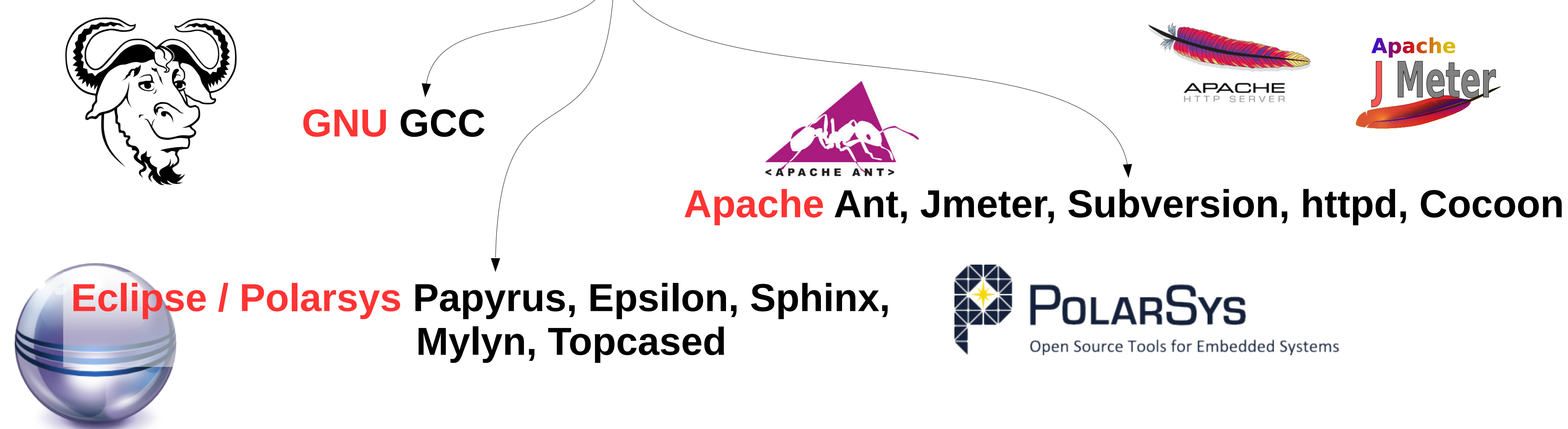
... **118 Rules Checked.**

At the **application, file and function** levels.

With different **time ranges** :

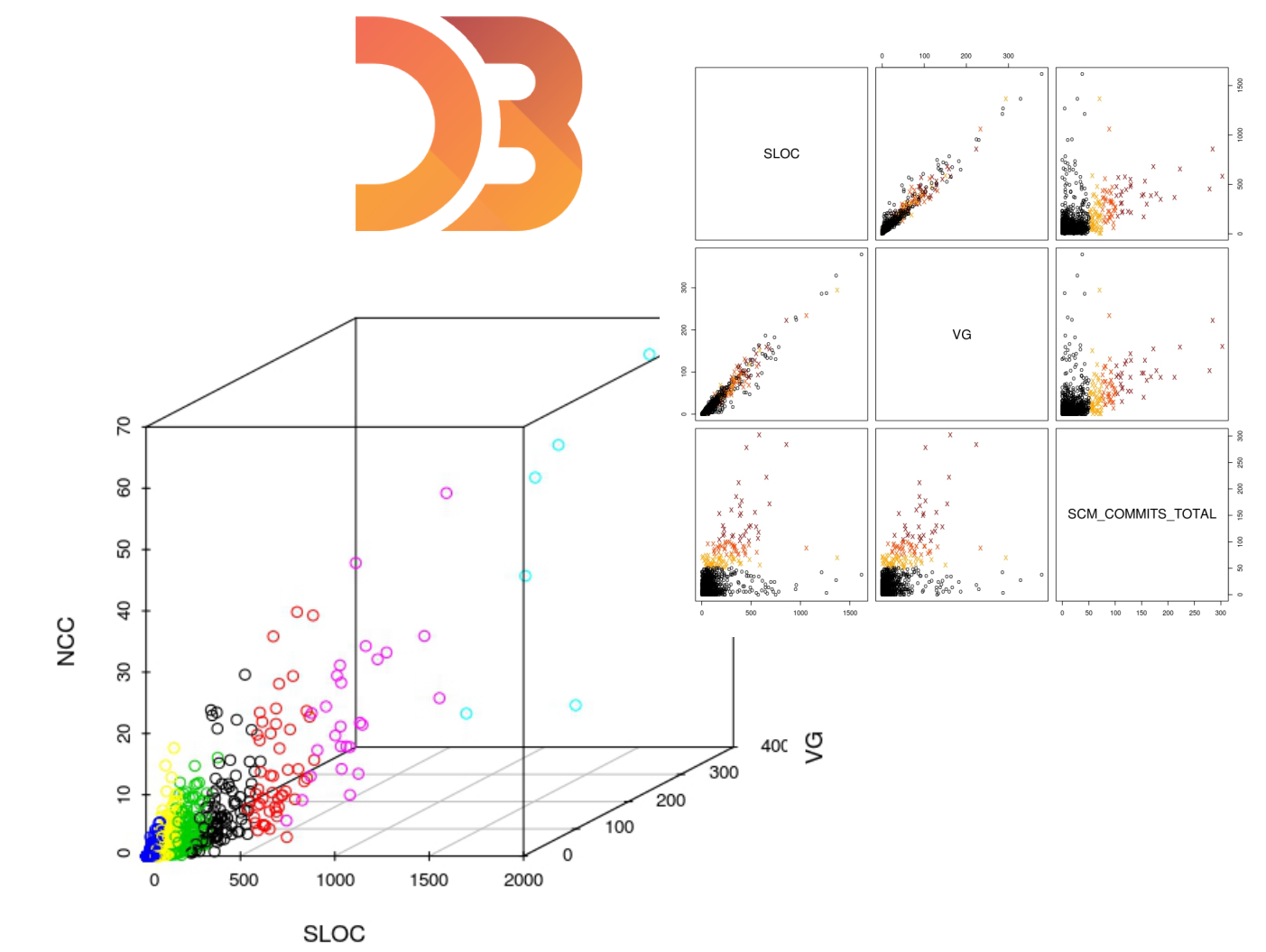
- Weekly on the project lifetime,
- Source releases,
- Single SVN extracts

From a variety of **open-source** software **projects** ...



You can use them for ...

... **Visualisation** ...



... **Analysis** ...

regression analysis



clustering of data

... and reproducible research.

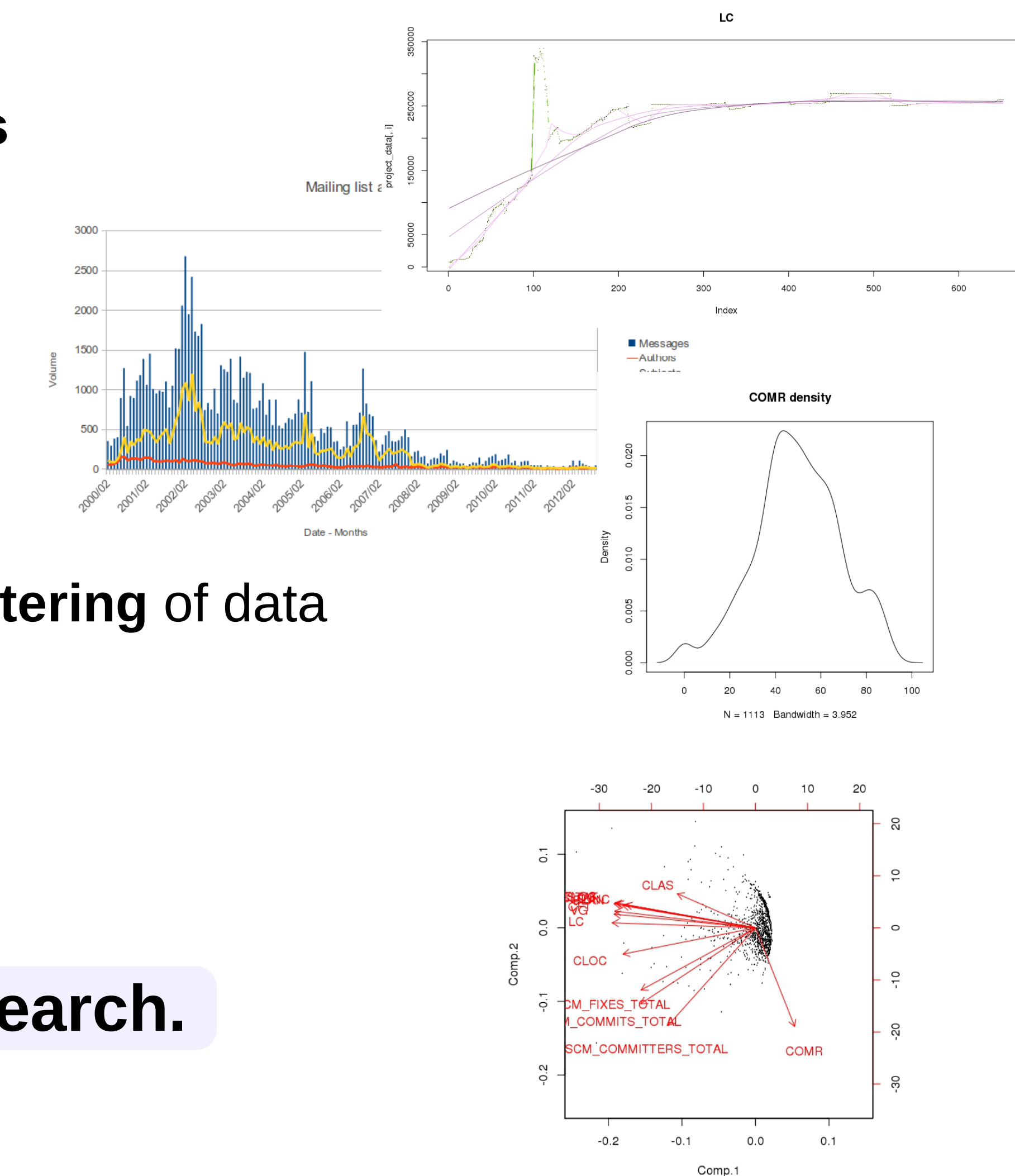


Figure 17: File metrics correlation matrix

Learn more on our web site : downloads, documentation, examples of usage...

<http://castalia.solutions/data-sets>