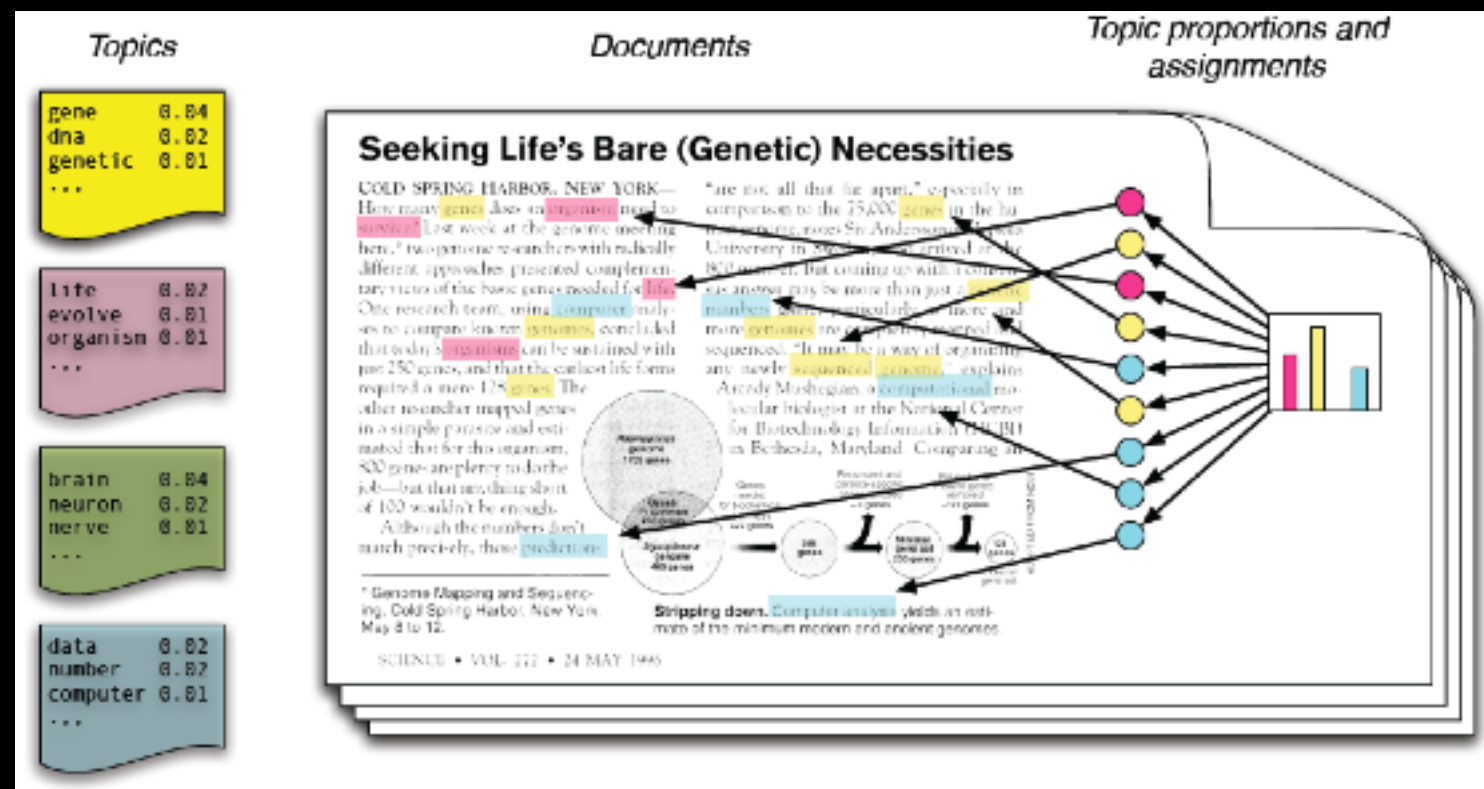


THE LDA TOPIC MODEL



- **Topic modeling** is the process of identifying **topics** (or clusters of terms that co-occur together) in a corpus (i.e., a set of documents).
- **Latent Dirichlet Allocation** (LDA) is an example of topic model and is used to classify text in a document to a particular topic.
- **LDA** is a form of **machine learning** with text data that views documents as **bags of words**; in other words, the order of the **n-grams** does not matter.

