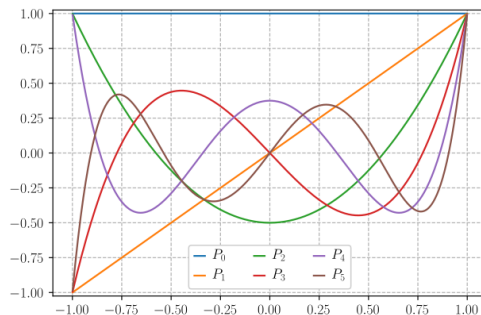


DOS04 PolySim – Similarity Detection Based on Polynomials

Background

Similarity detection plays an important role for information retrieval to detect similar documents. In open science, we not always own the right to share and process documents. Hence, we aim to mask the contents of input documents in the similarity detection process to keep the document plaintext hidden and protected.



Goal

- Develop a Python program to calculate the similarity between input documents based on polynomials

Tasks

- Transform document features and their positions into coordinates
- Approximate polynomials which are unique to each document
- Calculate the similarity between these polynomials

Cornelius Ihle
ihle@gipplab.org



Moritz Schubotz
Moritz.Schubotz@
fiz-karlsruhe.de



DOS05 Literature Review on Privacy-enhancing Tools for IPFS

Background

IPFS is an open network for sharing data. However, privacy is not a built-in feature of the network. Therefore, users typically rely on third-party tools to encrypt their data and anonymization tools like VPNs and TOR to protect their privacy. It is your task to provide a scoping overview on the currently existing tools to improve user privacy in IPFS.



Goal

- Develop a Python program to calculate the similarity between input documents based on polynomials

Tasks

- Systematically search public source code repositories for privacy enhancing projects that are suitable for IPFS
- Review articles for methods to improve privacy in public distributed hash tables.

Cornelius Ihle
ihle@gipplab.org



Moritz Schubotz
Moritz.Schubotz@
fiz-karlsruhe.de

