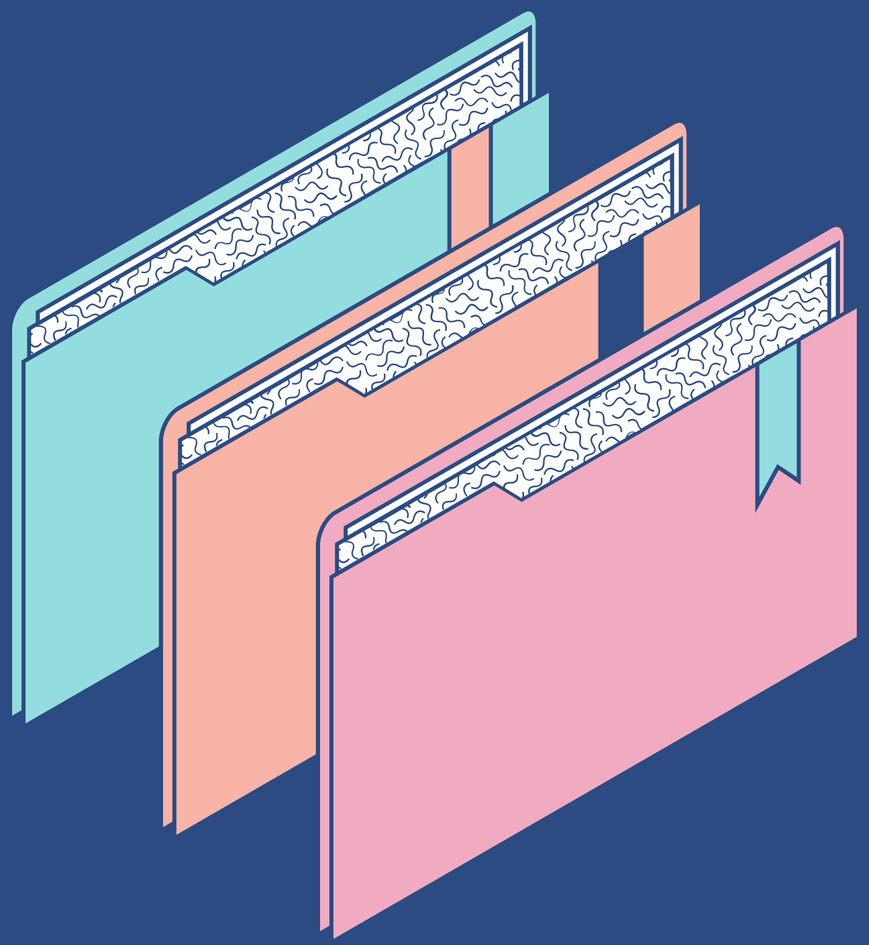
M.EIC PRI 2022/2023 G55

arXiv - Scientific Papers

Scientific Articles Search System

Beatriz Santos, up201905680
Sérgio Estêvão, up201905680
Sérgio da Gama, up201905680



Dataset

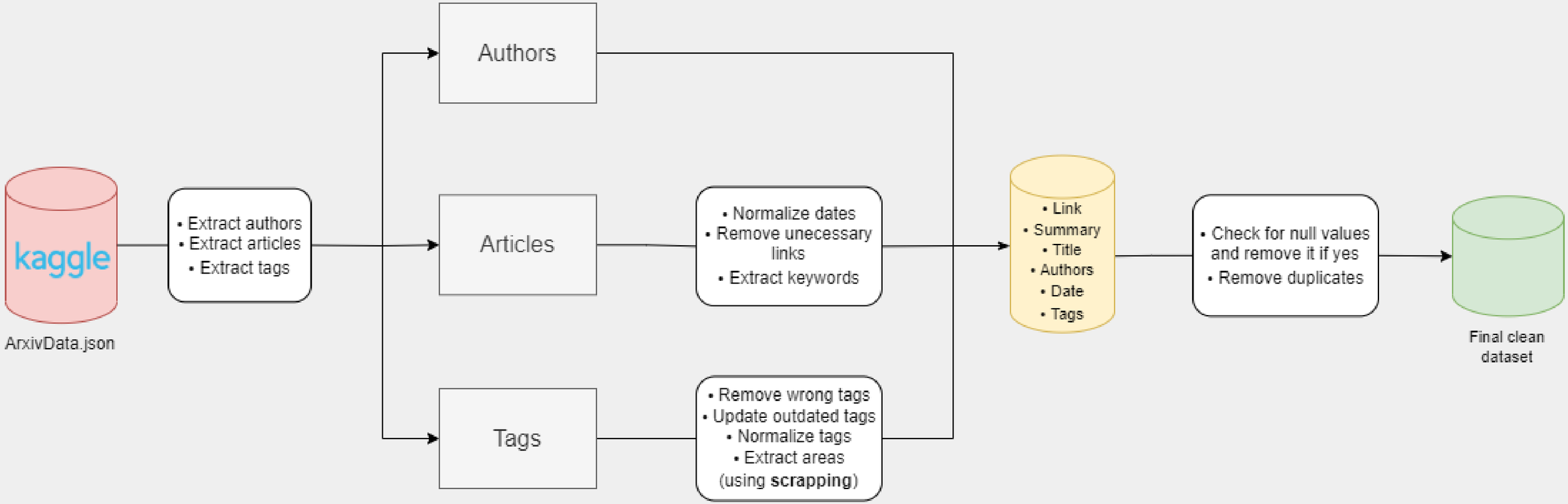
kaggle

Collection of scientific papers and their corresponding information from the website ARXIV

Data Refinement

Cleansing and shaping of the given dataset

Data Refinement



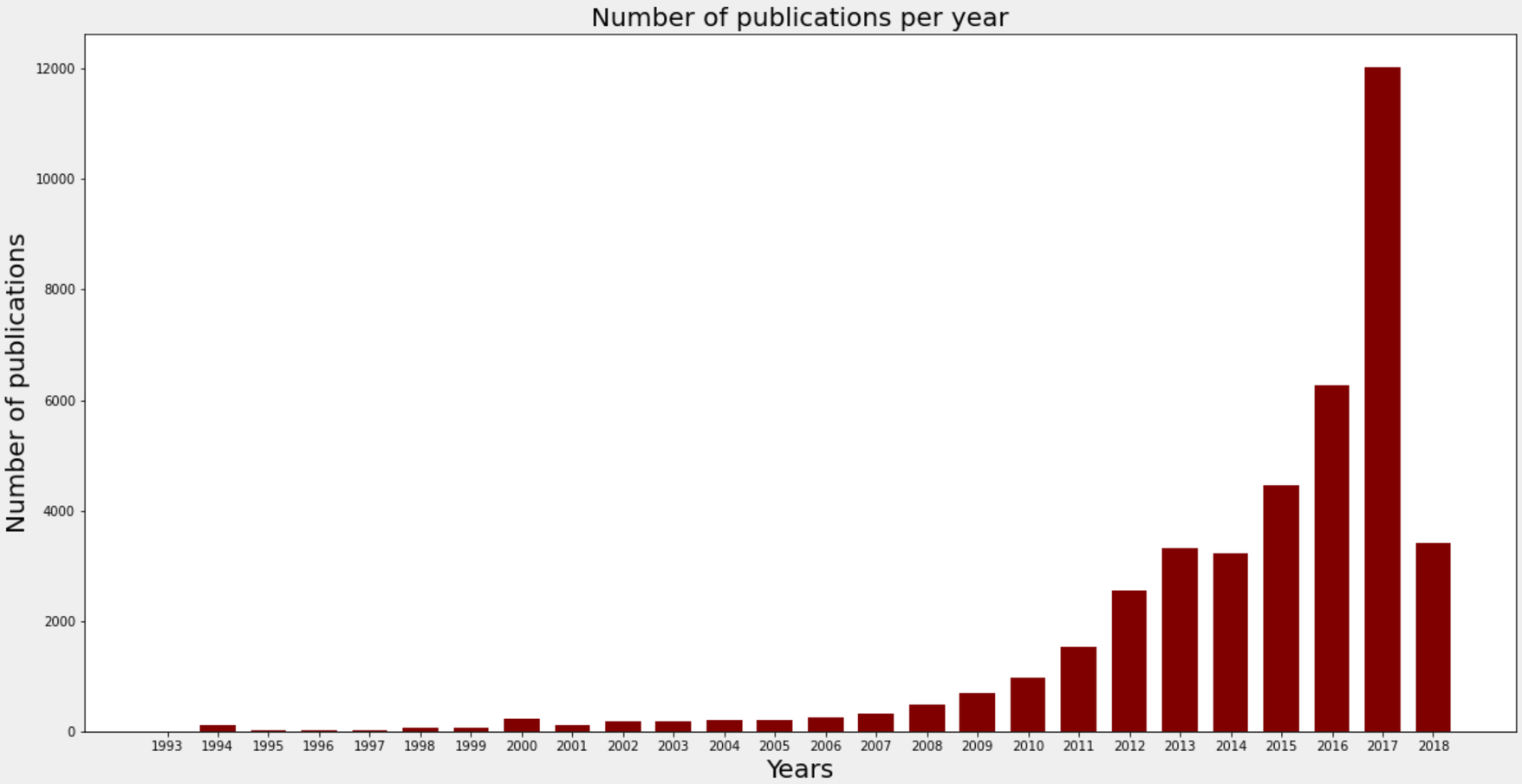
Original dataset


```
{
    "author": "[{'name': 'Iulian Vlad Serban'}, {'name': 'Tim Klinger'}, {'name': 'Gerald Tesauro'}, {'name': 'Kartik Talamadupula'}, {'name': 'Bowen Zhou'}, {'name': 'Yoshua Bengio'}, {'name': 'Aaron C
    "day": 2,
    "id": "1606.00776v2",
    "link": "[{'rel': 'alternate', 'href': 'http://arxiv.org/abs/1606.00776v2', 'type': 'text/html'}, {'rel': 'related', 'href': 'http://arxiv.org/pdf/1606.00776v2', 'type': 'application/pdf', 'title':
    "month": 6,
    "summary": "We introduce the multiresolution recurrent neural network, which extends the\nsequence-to-sequence framework to model natural language generation as two\nparallel discrete stochastic pro
    "tag": "[{'term': 'cs.CL', 'scheme': 'http://arxiv.org/schemas/atom', 'label': None}, {'term': 'cs.AI', 'scheme': 'http://arxiv.org/schemas/atom', 'label': None}, {'term': 'cs.LG', 'scheme': 'http:/
    "title": "Multiresolution Recurrent Neural Networks: An Application to Dialogue\n  Response Generation",
    "year": 2016
},
```

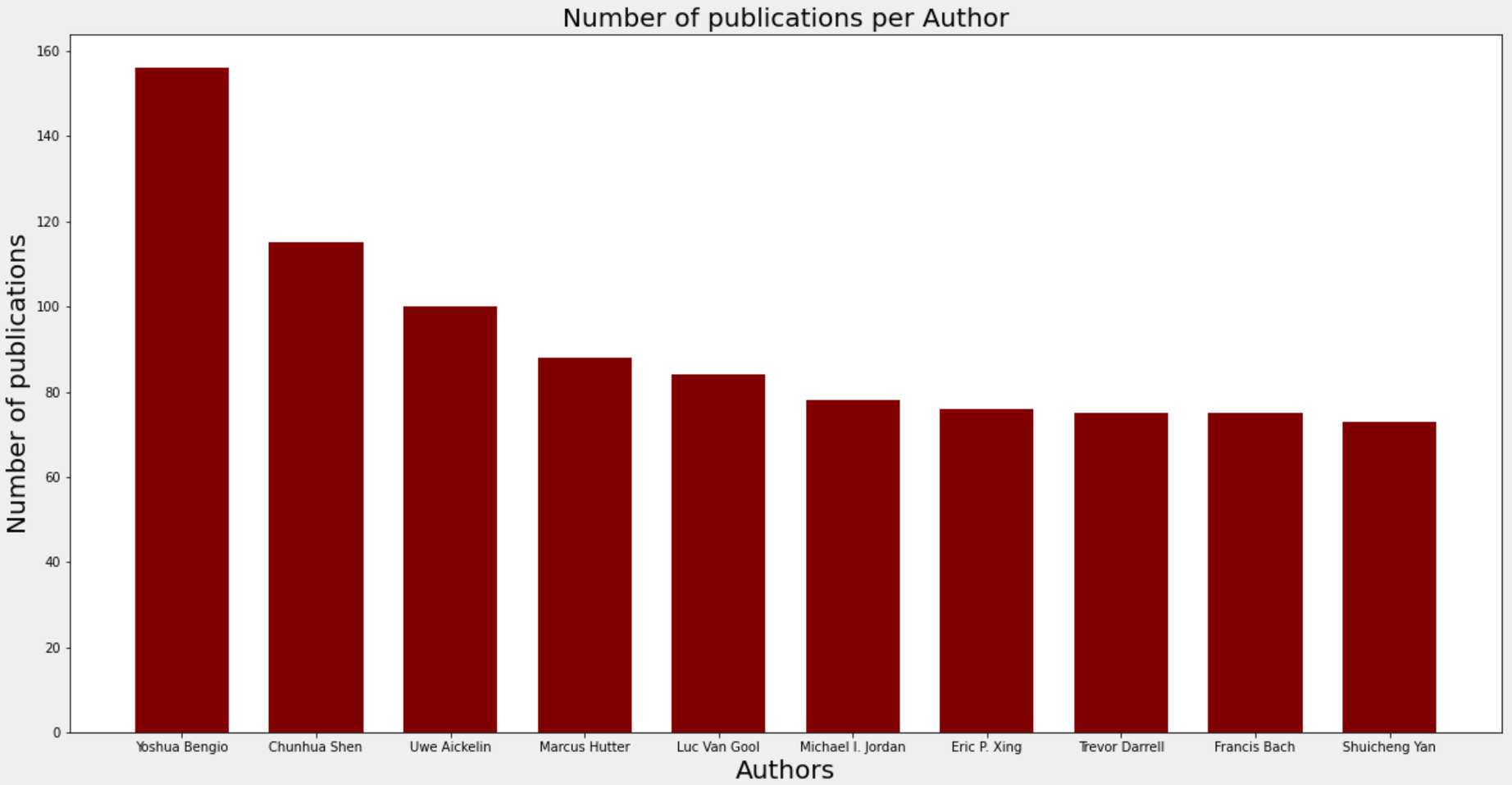
Refined dataset


```
"link": "http://arxiv.org/abs/1604.00289v3",
"summary": "Recent progress in artificial intelligence (AI) has renewed interest in\nbuilding systems
"title": "Building Machines That Learn and Think Like People",
"authors": [
  "Brenden M. Lake",
  "Tomer D. Ullman",
  "Joshua B. Tenenbaum",
  "Samuel J. Gershman"
],
"date": "2016-4-1",
"tags": {
  "Computer Science": {
    "Computer Science": [
      "Artificial Intelligence",
      "Computer Vision and Pattern Recognition",
      "Machine Learning",
      "Neural and Evolutionary Computing"
    ]
  },
  "Statistics": {
    "Statistics": [
      "Machine Learning"
    ]
  }
}
```

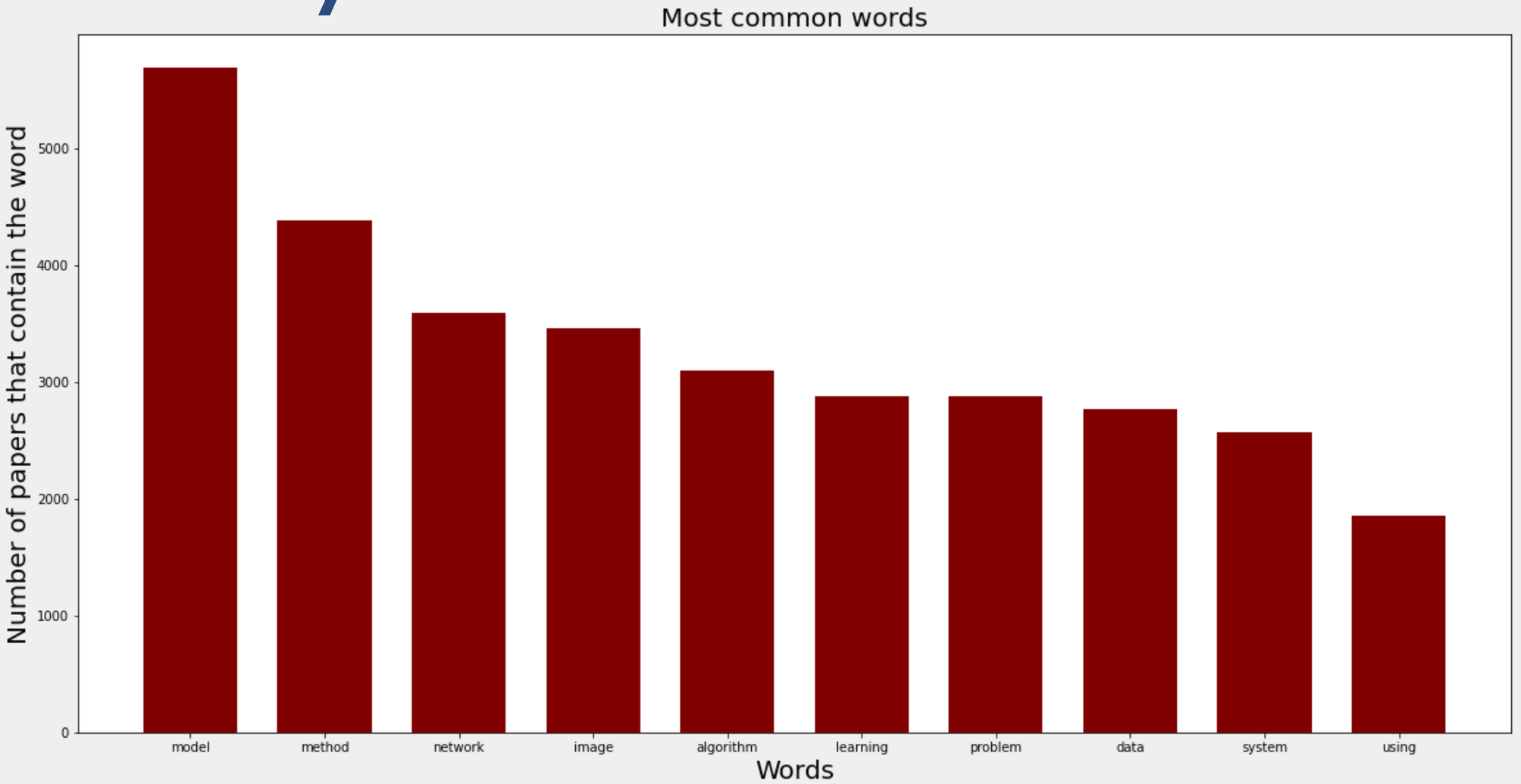
Data Analysis



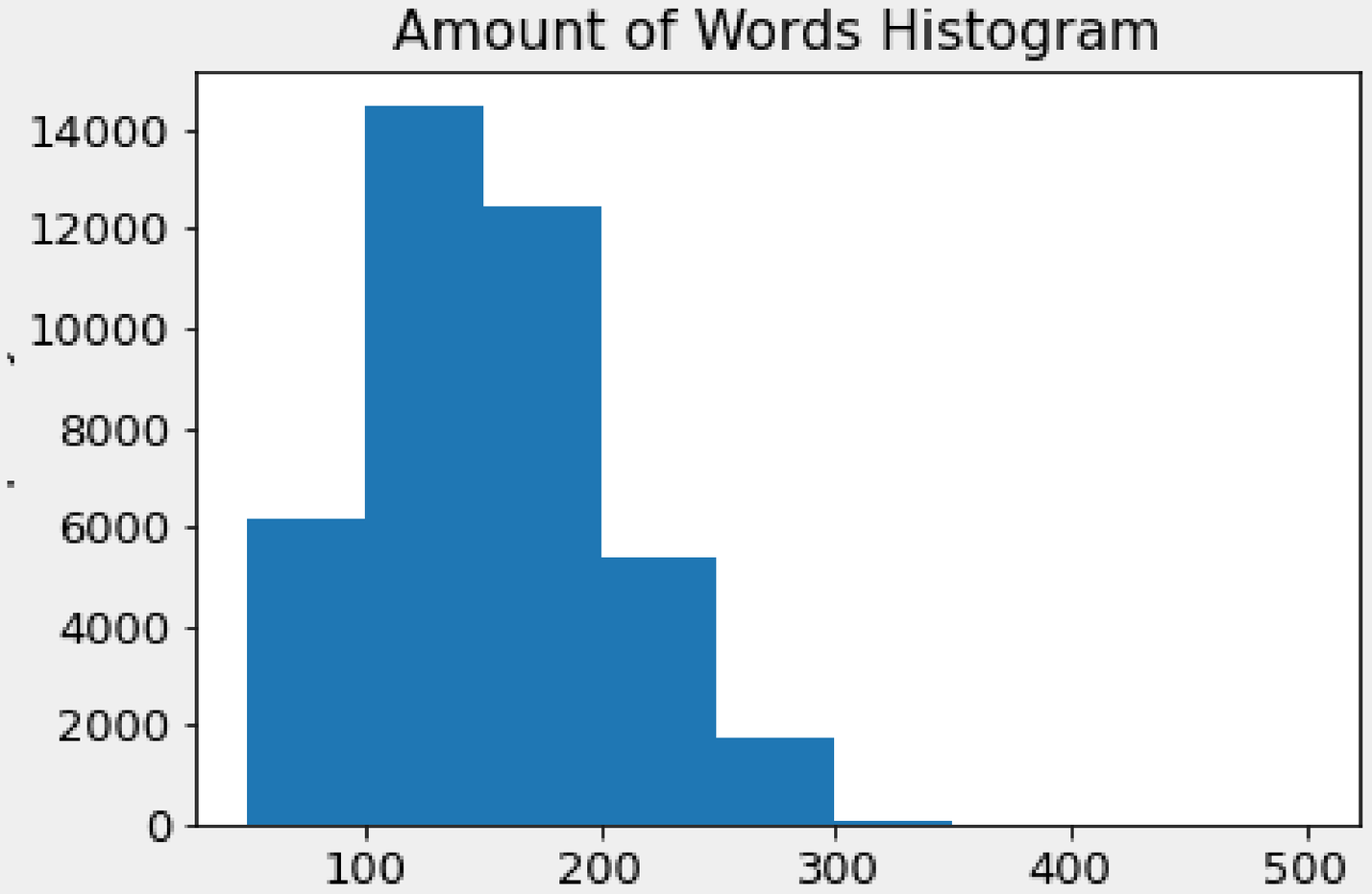
Data Analysis



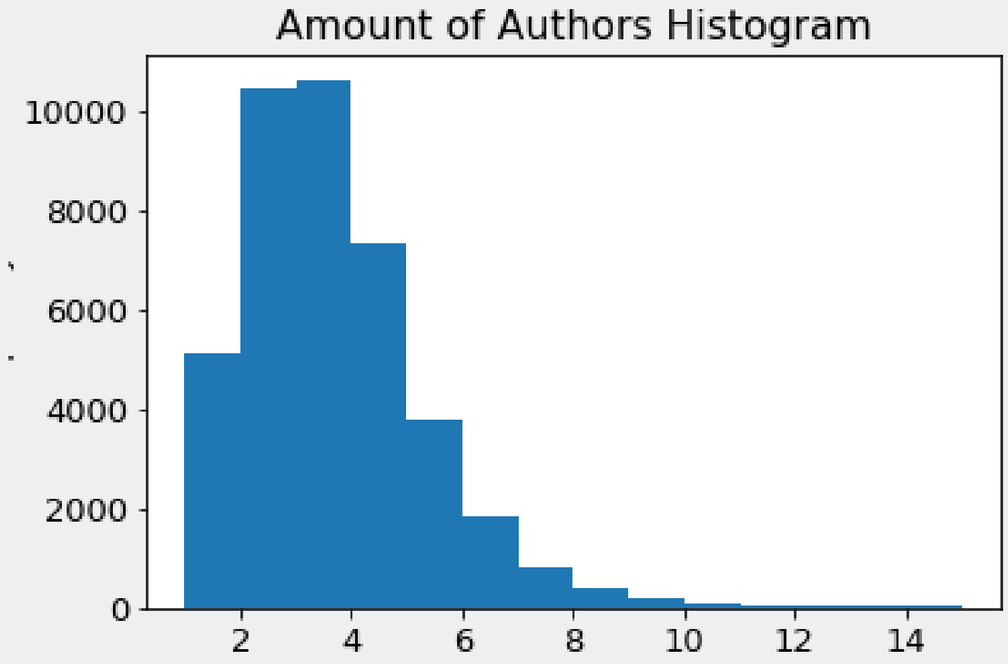
Data Analysis

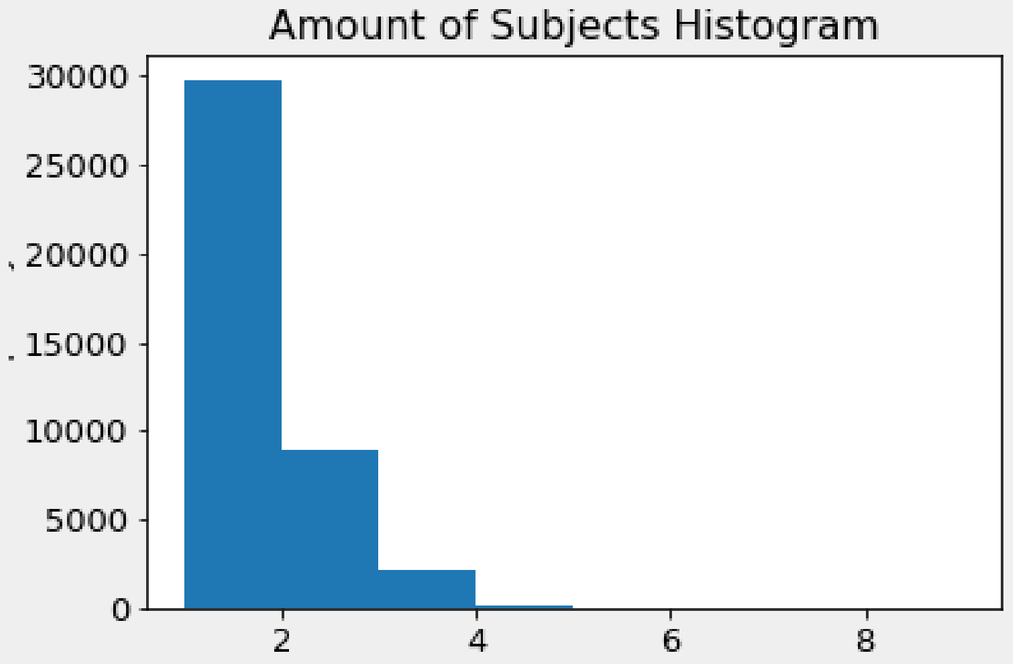


Data Analysis

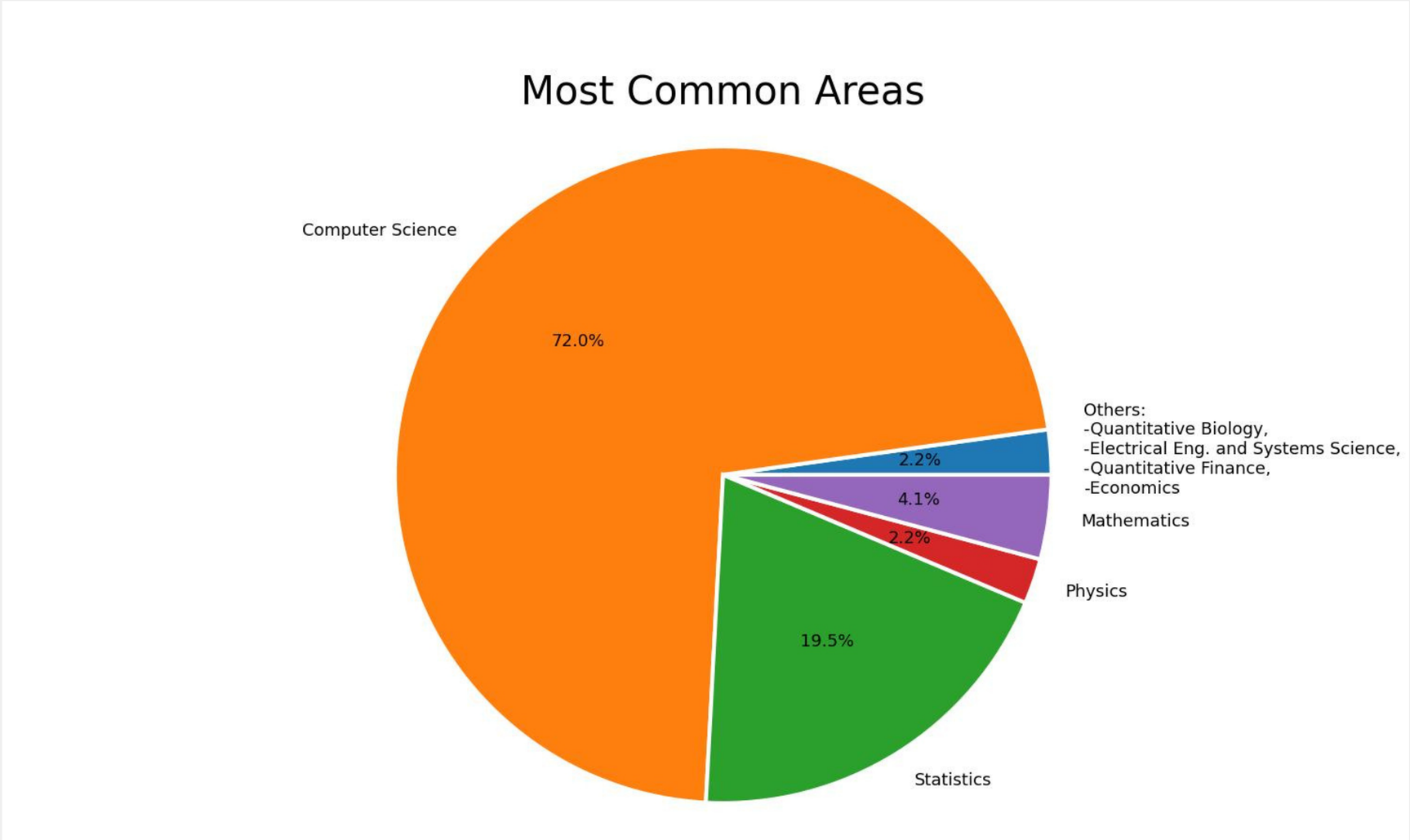


Data Analysis

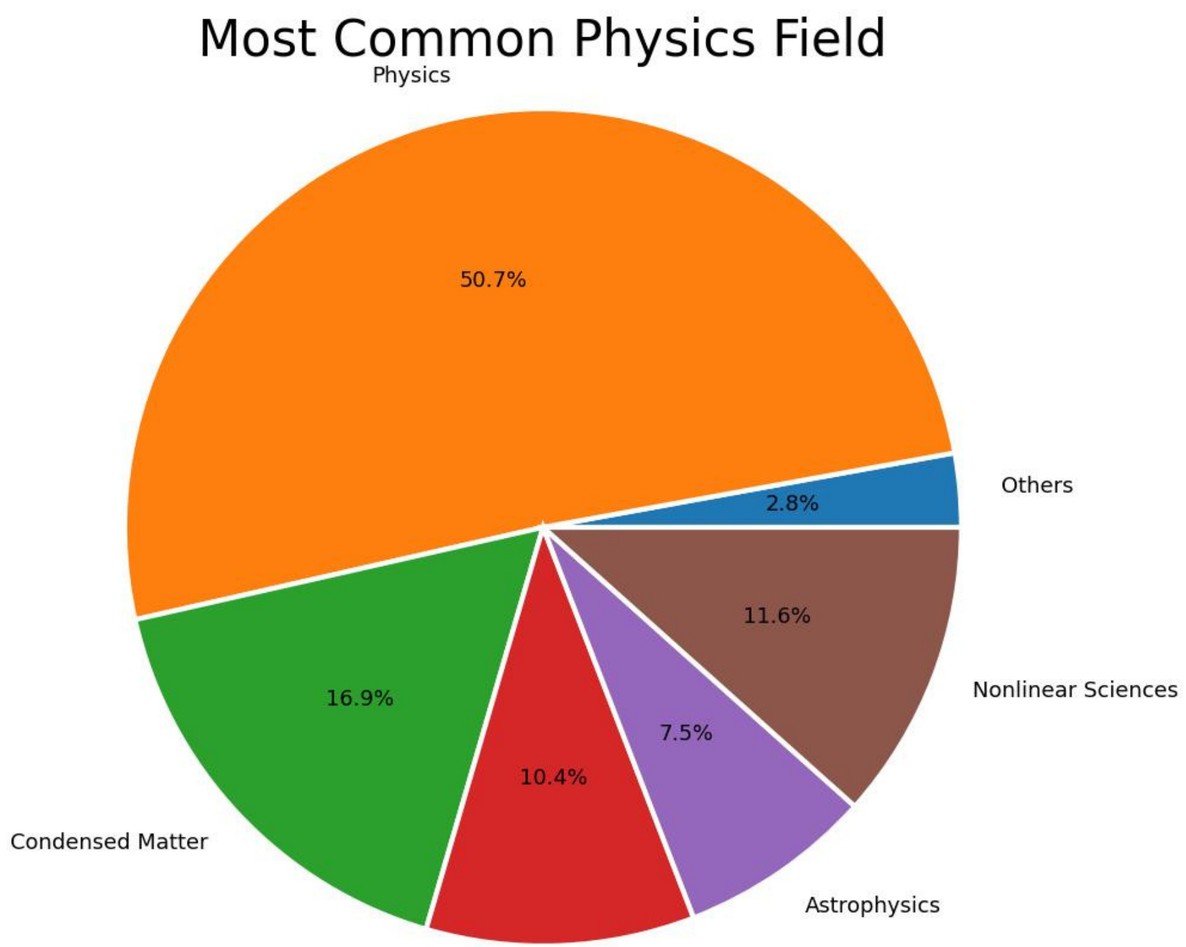




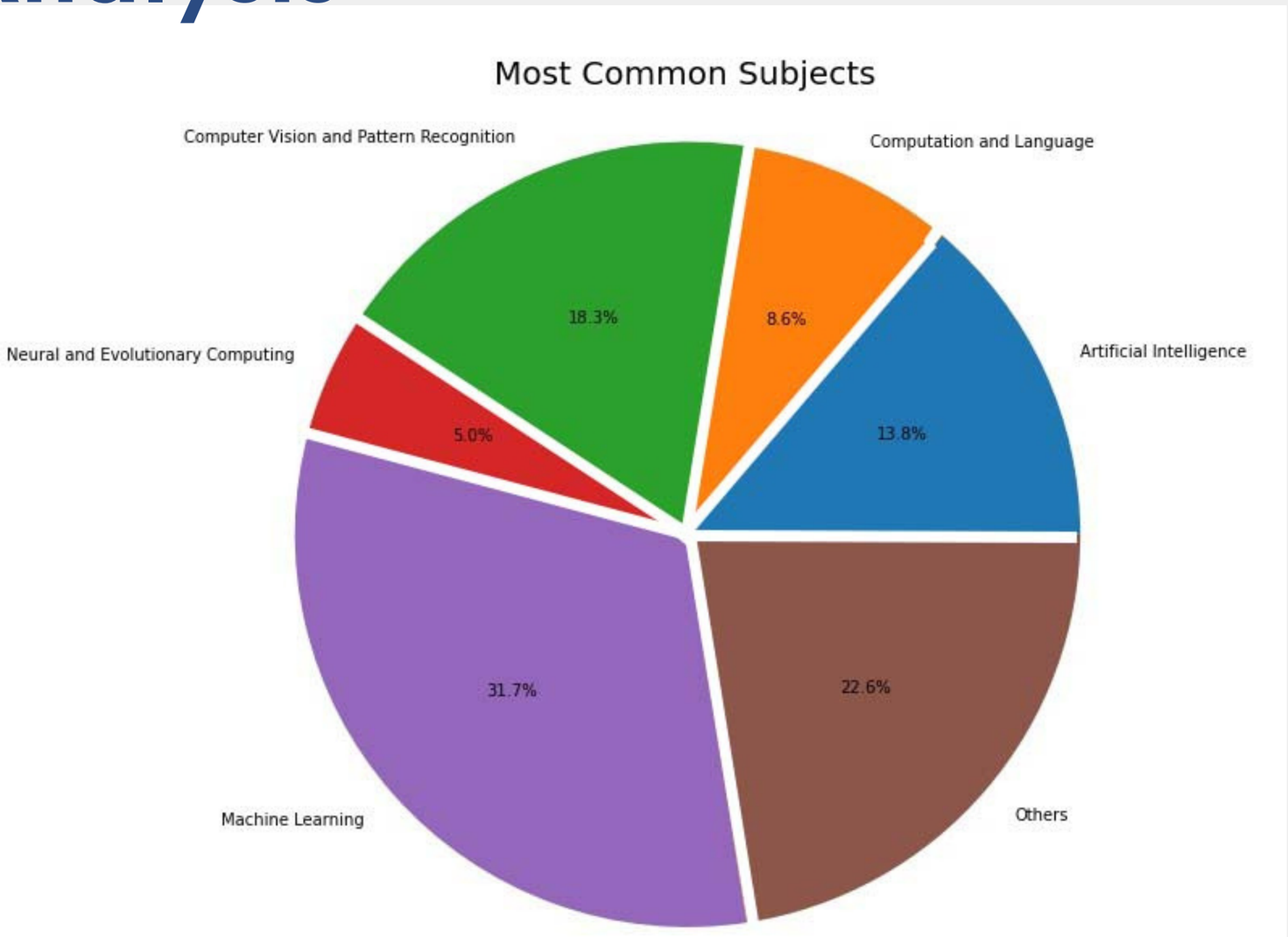
Data Analysis



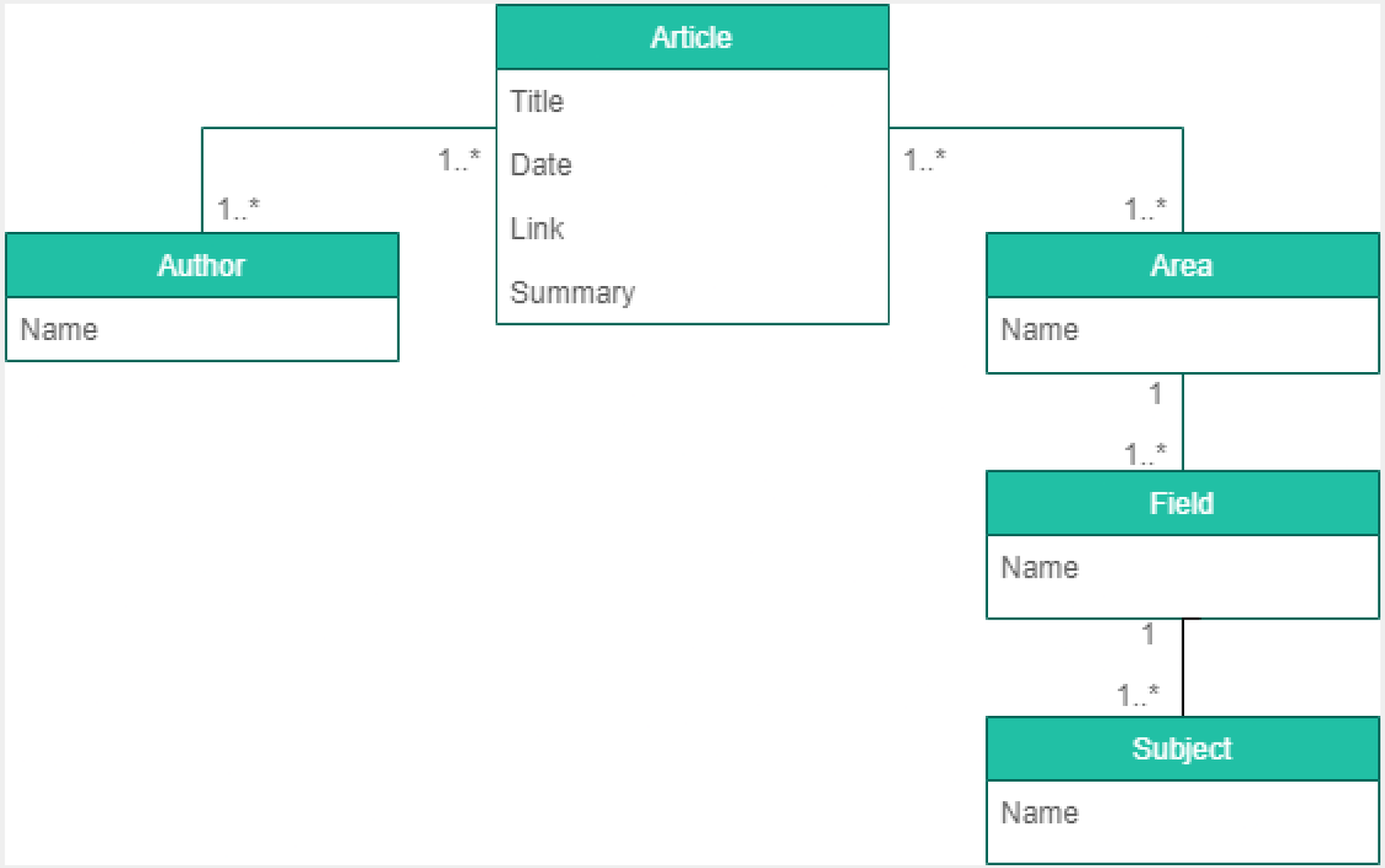
Data Analysis



Data Analysis



Conceptual Model



Prospetive search task

Information processing

Found 500 search results

Randomized channel-satet duality

Authors: Bin Yan, Nikolai A. Sinitsyn

Summary: Channel-state duality is a central result in quantum information science. It refers to the correspondence between a dynamical process (quantum channel) and a static quantum state in an enlarged Hilbert... more

Publication date: 7 October, 2022

Integration of Skyline Queries into Spark SQL

Authors: Lukas Grasmann, Reinhard Pichler, Alexander Selzer