# Network Analysis of Programming Languages and Their Fields



**Taksh Girdhar, 27012699**
**Kussh Satija, 80384878**

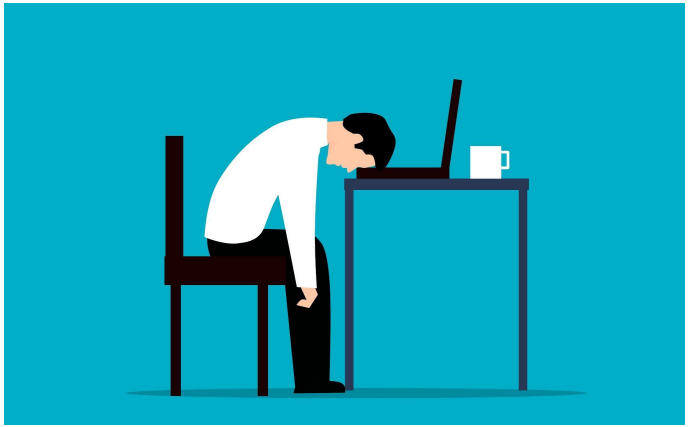
**COSC 421 Project Presentation**

# Table of contents

# Introduction

Back to Table of Contents

- **Overview**:
  - Computer science industry's rapid growth creates challenges for students.
- **Problem**:
  - Students often find themselves overwhelmed by the sheer variety of programming languages, and frameworks available.[1]
- **Challenge**:
  - Determining which skills are necessary for flexibility in the specialization choices causes uncertainty in where to place efforts.[2]
- **Personal Motivation**



1. (Akdur,2022)
2. (Rabai et. al, 2015)

# Problem Statement

“The overwhelming abundance of resources in computer science presents a challenge for computer science students, making it difficult to identify and prioritize what skills and knowledge to learn to build a strong foundation in the field and ultimately, leading to uncertainty about where to focus their efforts.”



Back to Table of Contents

# Research Questions

1. What programming languages are most central in their versatility across different fields?

2. Which programming languages are optimal for beginners while facilitating further specialization in various computer science fields?

3. Which programming languages tend to co-occur in the same job descriptions?



Back to Table of Contents

# Literature Review

## Research Insights

1. Charting Educational Pathways with Network Science
2. Exploring Skill Proximity and Compatibility Dynamics
3. Identifying Dominant Fields through CS Trends
4. Bridging Academia and Industry Practice Gaps
5. Guiding Skill Prioritization and Adaptable Learning

## Research Gaps

1. Lack of focus on programming languages in course-prerequisite networks
2. Omission of programming languages in skill proximity analyses
3. Underexplored role of programming languages in interdisciplinary growth
4. Disconnect between academia and industry in programming language adoption
5. Limited insights into the professional relevance of programming languages



Back to Table of Contents

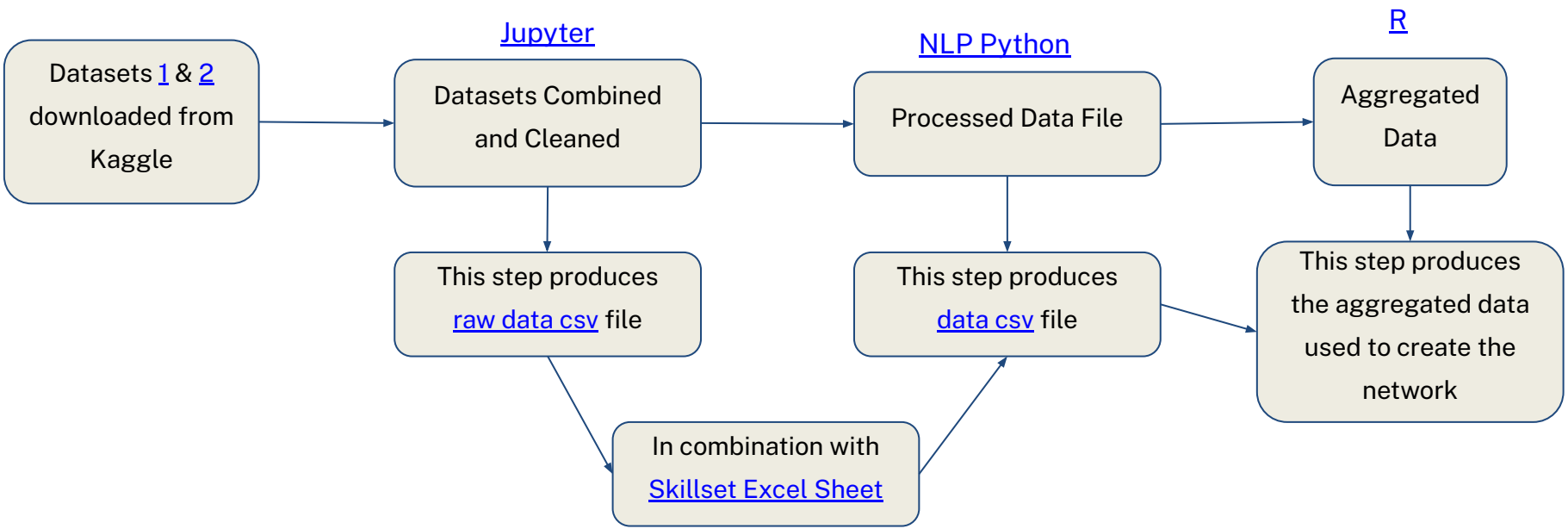
# Methodology Overview

- **Objective**: Explore trends in the employment landscape, specifically, in the demands of the skill-sets of incoming computer science students
- **Data Sources**:
  a. 2 Kaggle Datasets with a combined 12,000+ job postings from 2019.
  b. Supplementary skill sets dataset with curated list of top 20 programming languages and frameworks per CS field
- **Goal**: Provide data driven insights that address the uncertainty many computer science students face when determining their learning paths

Back to Table of Contents

# Data Collection & Pre-Processing

Back to Table of Contents



# Kaggle Dataset 1

| | Job Title | Job Description |
|---|---|---|
| 0 | Flutter Developer | We are looking for hire experts flutter developer. So you are eligible this post then apply your resume. Job Types: Full-time, Part-time Salary: ₹20,000.00 - ₹40,000.00 per month Benefits: Flexible schedule Food allowance Sche |
| 1 | Django Developer | PYTHON/DJANGO (Developer/Lead) - Job Code(PDJ - 04) Strong Python experience in API development (REST/RPC). Experience working with API Frameworks (Django/flask). Experience evaluating and improving the efficiency |
| 2 | Machine Learning | Data Scientist (Contractor) Bangalore, IN Responsibilities We are looking for a capable data scientist to join the Analytics team, reporting locally in India Bangalore. This person's responsibilities include research, design and dev |
| 3 | iOS Developer | JOB DESCRIPTION: Strong framework outside of iOS is always a plus iOS experience and generalist engineers with backgrounds in related technologies is a plus A disciplined approach to development, documentation and file st |
| 4 | Full Stack Developer | job responsibility full stack engineer – react role make impact petsmart transforming engineering team meet need rapidly changing retail environment role foundational helping build craft petsmart ' prowess making react j nativ |
| 5 | Java Developer | Software Developer - Integration* Immediate Opening!* A dynamic Akron / Cleveland area company is looking for an experienced software developer. The Integration Developer designs, develops, tests, maintains and enhance |
| 6 | Full Stack Developer | senior full stack developer \- 1800026h cwt looking senior full stack developer proven back-end skill well strong front-end skill role ' use cutting-edge technology help develop next generation business travel solution million use |
| 7 | JavaScript Developer | Job Description: ReactJS + NodeJs, Azure Functions, and GraphQL capability Strong hands on experience on javascript frameworks like reactjs Having hands on experience on nodejs Hands on experience on Azure function and |
| 8 | DevOps Engineer | Main Responsibilities and Deliverables: Manage/support the rollout, scalability and execution of the automation as it is consumed by the various teams Ensure automation services scales to handle rapid growth Using the automa |
| 9 | Software Engineer | Overview Based in Silicon Valley, Tintri is a wholly owned subsidiary of DataDirect Networks (DDN.com), the data-at-scale powerhouse and world's largest privately held storage company. Tintri solutions serve the needs of ente |
| 10 | Database Administrator | richardsontexasunited state bausp4 boeing company currently seeking database administrator join 's richardson texas location responsibility include limited following partner customer gather document requirement produce pro |

# Kaggle Dataset 2

| ID | Query | Job Title | Description |
|---|---|---|---|
| 1 | Data Scientist | Junior Data Scientist Apprenticeship | Job Description As a Junior Data Scientist at IBM, you will work as part of a team to solve busi |
| 2 | Data Scientist | HBO Data Scientist, Content Science | OVERALL SUMMARY As a Data Scientist on the Data Science Solutions team, this individual wil |
| 3 | Data Scientist | Junior Data Scientist | The Team: The Data science team is a newly formed applied research team within S&P; Global |
| 4 | Data Scientist | Jr Data Scientist | We now have a need for junior Data Scientist(s) in the NY area (or remote). The successful ca |
| 5 | Data Scientist | Data Scientist, Premium Content | Do you want to help guide the core business of Spotify using insights from analyses and data? |
| 6 | Data Scientist | Data Scientist | The Mayor's Office of Data Analytics (MODA) is New York City's ideas incubator for operational |
| 7 | Data Scientist | Customer Data Scientist | Smartly.io is a fast-growing company aiming to be the best and nicest place for the most talen |
| 8 | Data Scientist | Data Analyst / Data Scientist | We have a client that is looking for a data scientist to add to its team. This position is for a For |
| 9 | Data Scientist | Data Scientist | Marketing Statement MetroPlus Health Plan provides the highest quality healthcare services to |
| 10 | Data Scientist | Junior Data Scientist | Job Description: Junior Data Scientist, Membership Analytics Job Description: We are looking f |
| 11 | Data Scientist | Data Scientist | About Narrativ Narrativ [https://narrativ.com/] is a NYC-based, NEA-backed tech startup that i |
| 12 | Data Scientist | Geospatial Data Scientist | NYC Department of Finance (DOF) is responsible for administering the tax revenue laws of the |
| 13 | Data Scientist | Associate Data Scientist | Overview: Alliant is seeking an Associate Data Scientist to join its Data Science team to build p |
| 14 | Data Scientist | Data Scientist | _Position Mission/Summary:_ Looking for opportunities to use cutting edge technologies analy |
| 15 | Data Scientist | Data Scientist | Overview: Alliant is seeking a Data Scientist to join its Data Science team to build predictive m |

# Raw Data

| | Job Title | Description |
|---|---|---|
| 0 | Junior Data Scientist Apprenticeship | Job Description As a Junior Data Scientist at IBM, you will work as part of a team to solve business-related problems using da |
| 1 | HBO Data Scientist, Content Science | OVERALL SUMMARY As a Data Scientist on the Data Science Solutions team, this individual will be responsible for building adv |
| 2 | Junior Data Scientist | The Team: The Data science team is a newly formed applied research team within S&P; Global Ratings that will be responsible |
| 3 | Jr Data Scientist | We now have a need for junior Data Scientist(s) in the NY area (or remote). The successful candidate will join a team who is r |
| 4 | Data Scientist, Premium Content | Do you want to help guide the core business of Spotify using insights from analyses and data? The mission of our Premium An |
| 5 | Data Scientist | The Mayor's Office of Data Analytics (MODA) is New York City's ideas incubator for operational analytics. MODA collaborates w |
| 6 | Customer Data Scientist | Smartly.io is a fast-growing company aiming to be the best and nicest place for the most talented people to work and learn. W |
| 7 | Data Analyst / Data Scientist | We have a client that is looking for a data scientist to add to its team. This position is for a Fortune 50 company that is based |
| 8 | Data Scientist | Marketing Statement MetroPlus Health Plan provides the highest quality healthcare services to residents of Bronx, Brooklyn, N |
| 9 | Junior Data Scientist | Job Description: Junior Data Scientist, Membership Analytics Job Description: We are looking for a junior data scientist to join |

# Skillset Excel Sheet

| Artificial Intelligence (AI | Data Science | Cybersecurity | Computer Networks | Databases | Cloud Computing | Quantum Computing | Bioinformatics | Scientific Computing | Bloc |
|---|---|---|---|---|---|---|---|---|---|
| Python | Python | Python | Python | SQL | Python | Qiskit (Python-based) | Perl | MATLAB | Solidity |
| R | R | C | C | Python | Java | Cirq | Biopython | NumPy | Vyper |
| Java | Julia | C++ | Java | Java | JavaScript | Strawberry Fields | R | SciPy | Chainco |
| C++ | SQL | Java | JavaScript | C# | Go | PennyLane | Julia | Fortran | Rust |
| Julia | SAS | Ruby | Go | R | C++ | Quantum Development Kit (Q#) | Python | Cython | Python ( |
| Scala | Scala | PowerShell | C++ | PHP | Ruby | PyQuil | Scala | IDL | JavaScr |
| MATLAB | MATLAB | Perl | Swift | Kotlin | Rust | OpenQASM | Java | ROOT | C# |
| Prolog | Spark | JavaScript | Rust | Scala | C# | QuTiP | C++ | Julia | Kotlin |
| Lisp | TensorFlow | Swift | Bash/Shell | MATLAB | Perl | D-Wave Ocean SDK | PHP | OpenFOAM | Rholang |
| Haskell | NumPy | Rust | Perl | Ruby | PHP | QCL (Quantum Computation Language) | Swift | Wolfram Language | Go |
| TensorFlow | Pandas | Go | Haskell | Swift | Scala | Rigetti Forest | C | Octave | LLL (Lov |
| PyTorch | PyTorch | PHP | Kotlin | Go | Bash/Shell | Yao.jl (Julia-based) | Haskell | COMSOL Script | PHP |
| Swift | Hadoop | MATLAB | SQL | Lua | Kotlin | Tequila | Lua | Maple | Java |
| Rust | H2O.ai | Lisp | Lua | TensorFlow | MATLAB | Qiskit Metal | Rust | TensorFlow (scientific apps) | Python |
| Go | Dask | Scala | Julia | PyTorch | Scala | Aqua (IBM Quantum) | MATLAB | Mathematica | Erlang |

## Processed Data

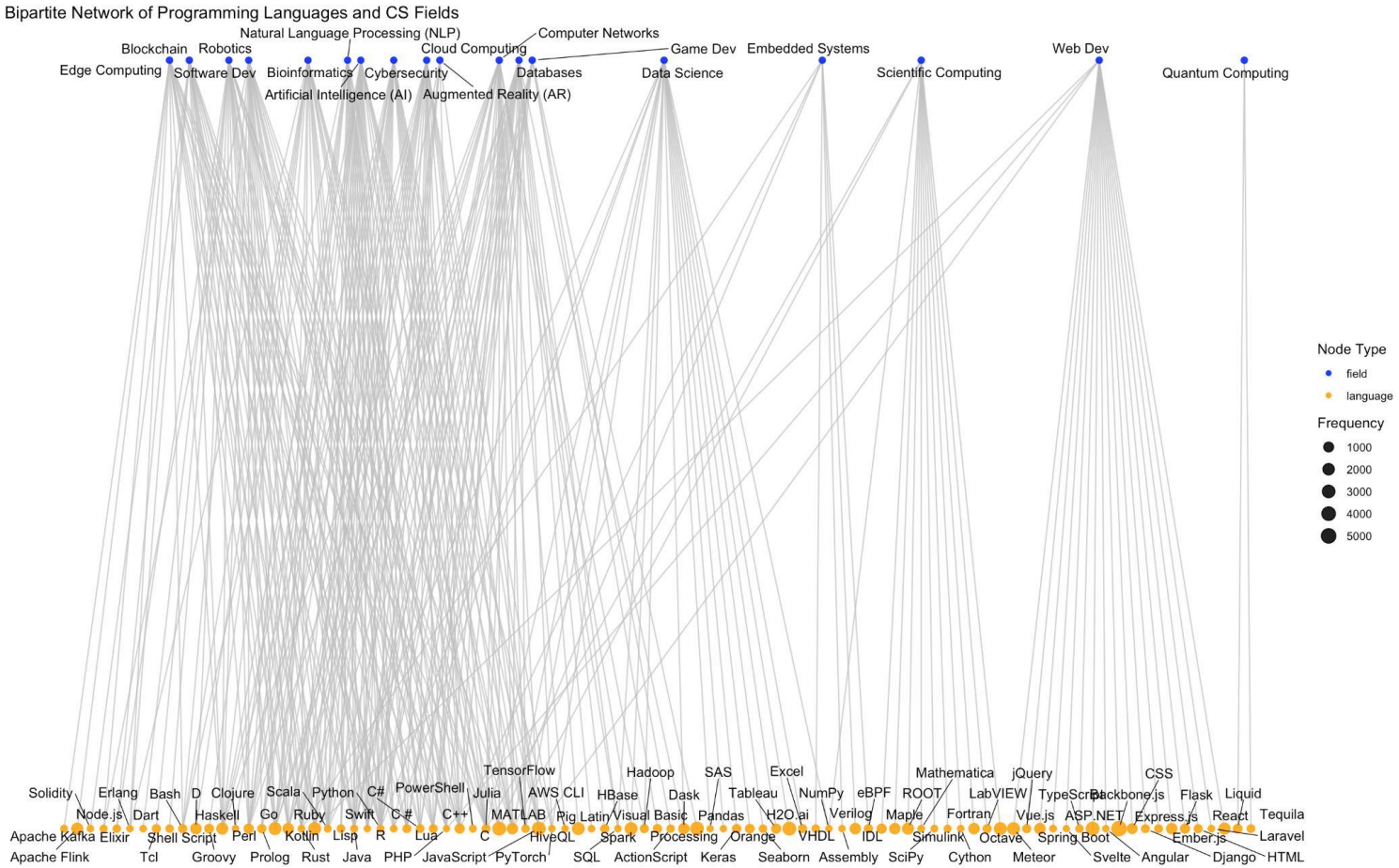
| CS Field | Language | Count |
|---|---|---|
| Bioinformatics | Python | 3 |
| Artificial Intelligence (AI) | Python | 3 |
| Game Dev | Python | 3 |
| Computer Networks | Python | 3 |
| Augmented Reality (AR) | Python | 3 |
| Cybersecurity | Python | 3 |
| Software Dev | Python | 3 |
| Data Science | Python | 3 |
| Robotics | Python | 3 |
| Databases | Python | 3 |
| Blockchain | Python | 3 |
| Edge Computing | Python | 3 |
| Cloud Computing | Python | 3 |
| Natural Language Processing (NLP) | Python | 3 |
| Data Science | Hadoop | 1 |
| Computer Networks | Hadoop | 1 |
| Data Science | Spark | 1 |
| Computer Networks | Spark | 1 |
| Bioinformatics | Java | 1 |

## Aggregated Data

| | CS.Field | Language | Count |
|---|---|---|---|
| 1 | Artificial Intelligence (AI) | C # | 2 |
| 2 | Artificial Intelligence (AI) | C# | 116 |
| 3 | Artificial Intelligence (AI) | C++ | 653 |
| 4 | Artificial Intelligence (AI) | Go | 269 |
| 5 | Artificial Intelligence (AI) | Haskell | 13 |
| 6 | Artificial Intelligence (AI) | Java | 2337 |
| 7 | Artificial Intelligence (AI) | JavaScript | 1693 |
| 8 | Artificial Intelligence (AI) | Julia | 26 |
| 9 | Artificial Intelligence (AI) | Lisp | 1 |
| 10 | Artificial Intelligence (AI) | MATLAB | 112 |
| 11 | Artificial Intelligence (AI) | Perl | 268 |
| 12 | Artificial Intelligence (AI) | Prolog | 2 |
| 13 | Artificial Intelligence (AI) | PyTorch | 89 |
| 14 | Artificial Intelligence (AI) | Python | 3165 |

# Network Structure

Back to Table of Contents



# Results Overview

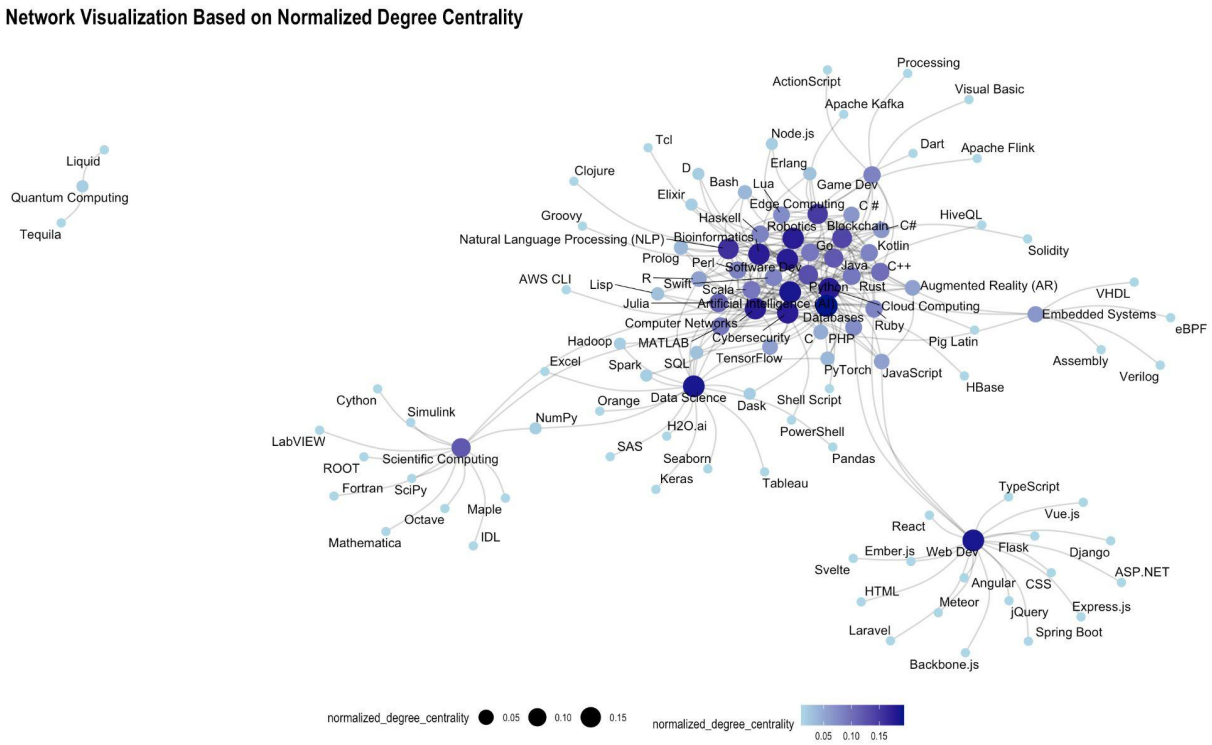
- **Normalized Degree Centrality**: Measures node connectivity and influence.
- **Betweenness Centrality**: Measures node importance as a bridge between other nodes.
- **Leiden Coefficient**: Measures the strength of community structure in the network.
- **Analytics**: Conducted in R, we used the libraries dyplr, igraph, ggraph, tidygraph , ggrepel, networkD3 and simplewidgets



Back to Table of Contents

# Normalized Degree Centrality

Back to Table of Contents



| Programming Language | Normalized Degree Centrality |
|---|---|
| Python | 0.1296 |
| Java | 0.1204 |
| Julia | 0.1111 |
| C++ | 0.1019 |
| Go | 0.0926 |
| MATLAB | 0.0926 |
| Rust | 0.0926 |
| Scala | 0.0926 |
| Haskell | 0.0833 |
| Perl | 0.8333 |

**Table 1.** Top 10 Programming Languages and Frameworks from normalized degree centrality

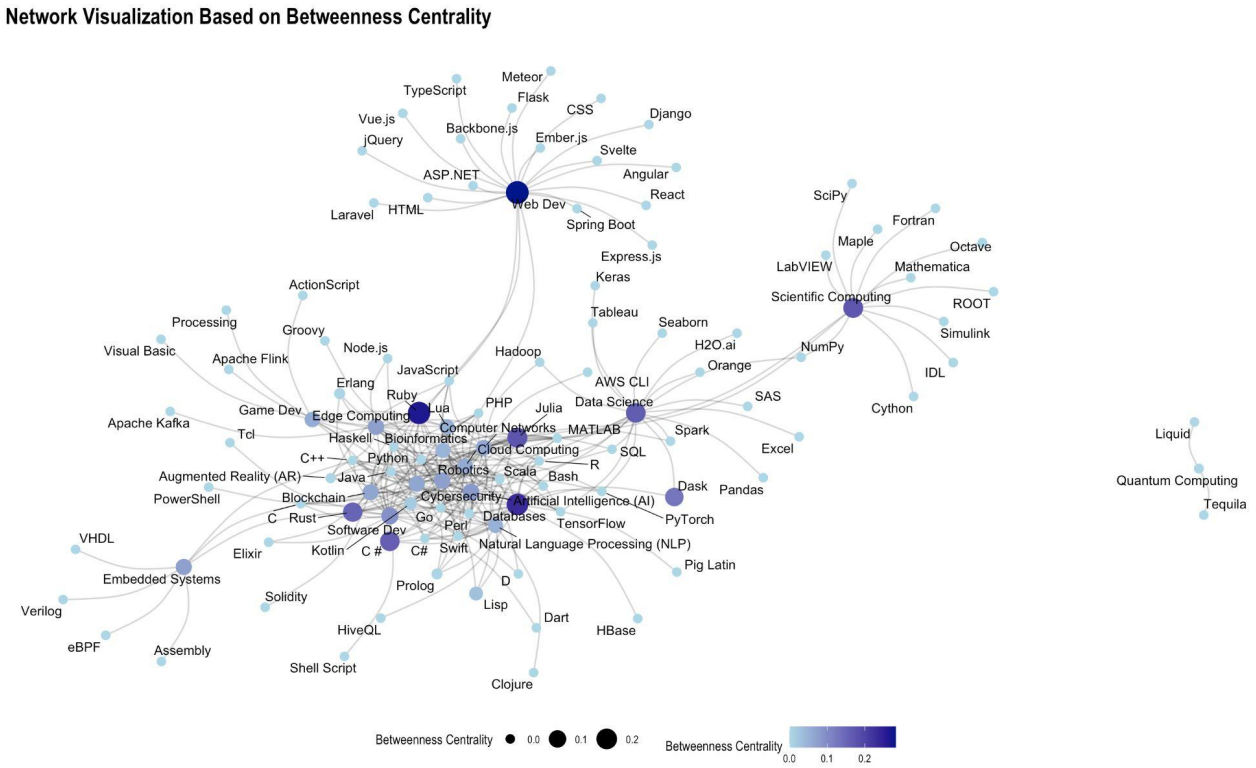
# Degree Centrality Discussion

**Findings for RQ. 1: What programming languages are most central in their versatility across different fields?**

- **Vailshery (2024)**:
  - Reports Python and Java as widely used among 60,000 developers.
  - Our study aligns with Python and Java as central but identifies Python as the most popular, contrary to JavaScript being ranked highest by Vailshery.
- **TIOBE Index (Jansen, 2022)**:
  - Supports our findings with Python ranked as the most popular, followed by C++ and Java.
  - Reinforces the centrality of Python, Java, and C++ in our analysis.
- **Key Insights**
  - Our study focuses on employer demand, providing a more current reflection of industry needs compared to prior surveys based on skillsets of past developers.
  - Languages like Julia, Haskell, and Perl are deemed insignificant but might prove useful in valuable contexts
  - Example: Julia's design, combining the speed of C, generality of Python, and statistical power of R. Julia’s straightforward syntax makes it ideal for beginners and advanced specializations. Libraries like GameZero (game development) and Dash.jl (web applications) broaden its applicability.

# Betweenness Centrality

Back to Table of Contents



| Programming Language | Betweenness Centrality |
|---|---|
| Ruby | 0.2648 |
| Julia | 0.1751 |
| C# | 0.1611 |
| Rust | 0.1551 |
| Dask | 0.1294 |
| Lua | 0.0507 |
| Lisp | 0.0298 |
| Kotlin | 0.0171 |
| Erlang | 0.0035 |
| Prolog | 0.0033 |

**Table 2.** Top 10 Programming Languages and Frameworks based on betweenness centrality

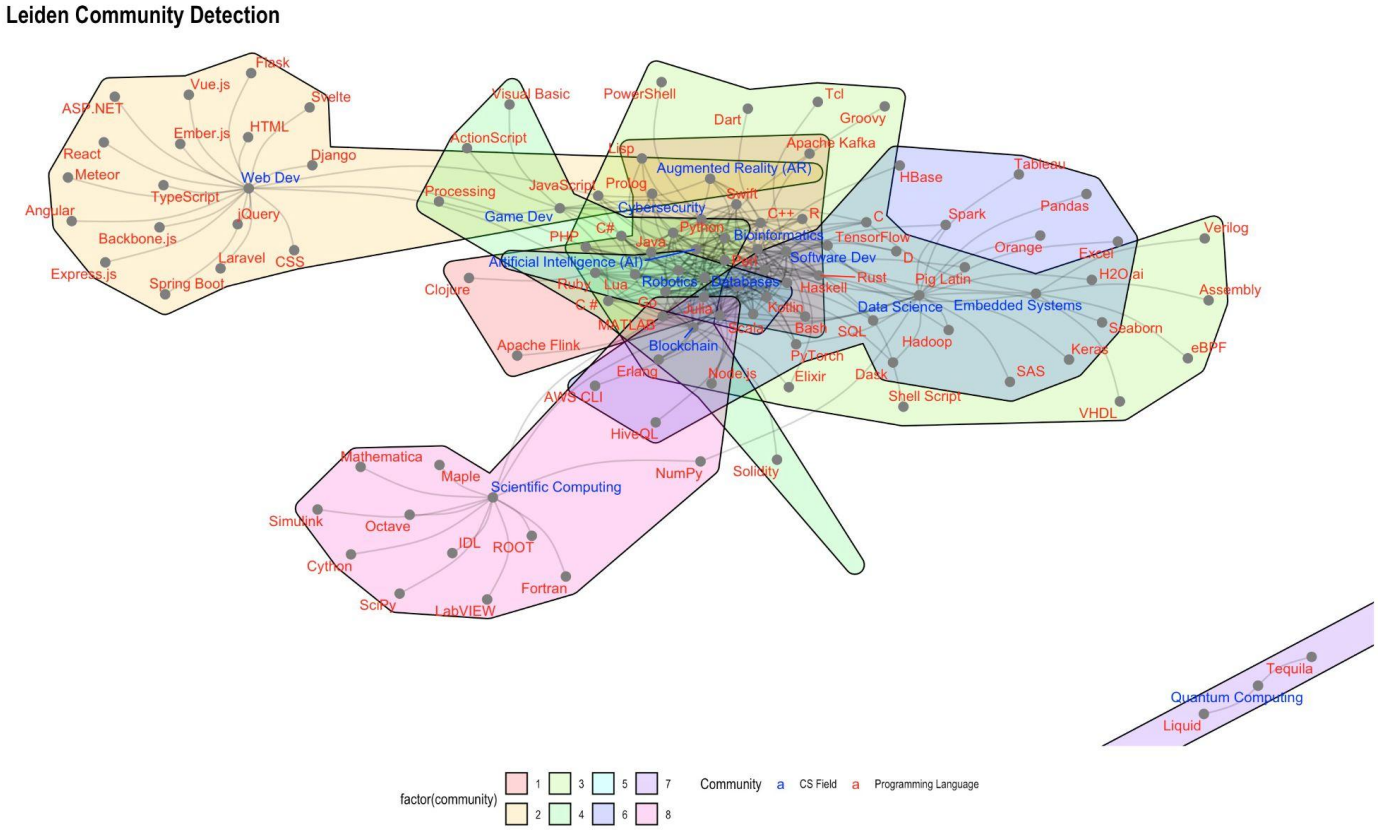
# Betweenness Centrality Discussion

**Findings for RQ. 2: Which programming languages are optimal for beginners while facilitating further specialization in various computer science fields?**

- **Hewner et al. (2011)**:
  - Discusses factors influencing CS students' specialization choices:
    1. Enjoyment of specific classes.
    2. Lack of adequate research.
    3. Reliance on curriculum for guidance.
    4. Absence of a clear career vision.
  - Our findings support Ruby, Julia, and C++ as transitional languages bridging foundational learning and specialization.
- **Danao (2023)**:
  - Ruby's simple syntax boosts productivity, making it a flexible language for beginners.
  - Frameworks like Ruby on Rails power major platforms (Airbnb, GitHub).
- **Seaton (2022)**:
  - Ruby's broad applications include software engineering, bioinformatics, and robotics.
- **Key Insight**:
  - Ruby's simplicity aids learning core concepts while transitioning to advanced languages like Python for specialized tasks.

# Leiden Community Detection

Back to Table of Contents



| Community | Size | Density | Most Central Programming Language Node | Most Central CS Field Node | CS Field Nodes Present |
|---|---|---|---|---|---|
| 1 | 11 | 0.2727 | Java | Artificial Intelligence | Artificial Intelligence<br>Edge Computing<br>Natural Language Processing |
| 2 | 20 | 0.1000 | JavaScript | Web Dev | Augmented Reality<br>Web Dev |
| 3 | 23 | 0.1462 | C++ | Bioinformatics | Bioinformatics<br>Cybersecurity<br>Embedded Systems,<br>Software Dev |
| 4 | 11 | 0.2545 | Python | Robotics | Blockchain,<br>Game Dev,<br>Robotics |
| 5 | 6 | 0.3333 | Go | Cloud Computing | Cloud Computing |
| 6 | 21 | 0.1238 | TensorFlow | Databases | Computer Networks<br>Data Science<br>Databases |
| 7 | 3 | 0.6667 | Liquid | Quantum Computing | Quantum Computing |
| 8 | 14 | 0.1429 | Julia | Scientific Computing | Scientific Computing |

**Table 3.** Statistics of Leiden communities within the Programming Languages and CS Fields Network

# Leiden Communities Discussion

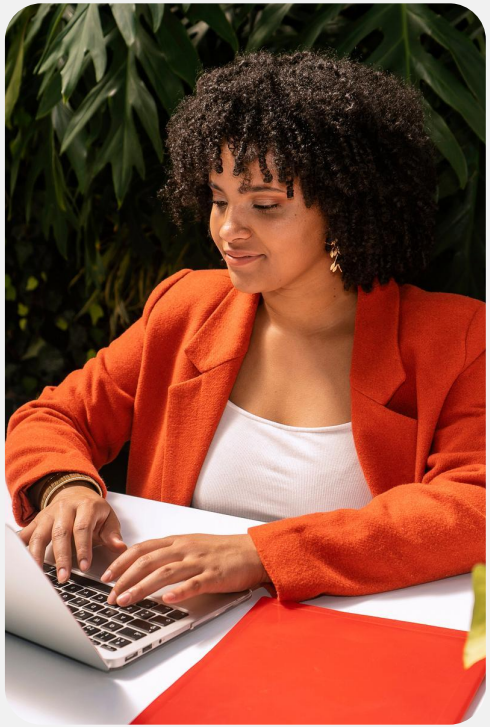
**Findings for RQ. 3: Which programming languages tend to co-occur in the same job descriptions?**

- **Dada et al. (2022)**:
  - Identifies C++ as most used across IT roles like mobile development and system administration, aligning with our finding of its centrality in Community 3.
  - Python, while central to data science in their study, is also highlighted by us in fields like blockchain and game development.
- **Mishra (2024)**:
  - Python's extensive libraries make it ideal for handling complex tasks in data science, robotics, and more.
- **Banoula (2024)**:
  - TensorFlow's suitability for large-scale computations supports its centrality in fields like databases, data science, and computer networks.
- **Key Insight**:
  - Our analysis focuses on interrelationships between fields and languages, offering insights into foundational and interconnected skills for CS students.

# Difficulties faced:

- **Challenge**:
  Our web scraper was unable to extract real-time job descriptions from LinkedIn due to restricted API access and LinkedIn's guidelines.
  **Solution**:
  Identified and utilized a large dataset containing job postings and descriptions.
- **Challenge**:
  Difficulty in locating a recent and comprehensive dataset of job postings.
  **Solution**:
  Merged two datasets from February 2019, containing job descriptions and related details.
- **Challenge**:
  Lack of an existing cumulative skill sheet linking fields with their associated programming languages and frameworks.
  **Solution**:
  Conducted research to develop a custom skill sheet, listing the top 20 programming languages and frameworks for each field.

Back to Table of Contents



# End Product

Link

Back to Table of Contents

# Conclusion

- **Study Objectives**:
  - Explored programming language networks and identified versatile, central languages.
  - Focused on transitional languages aiding specialization and co-occurrence in job descriptions.
- **Key Findings**:
  - Central languages/frameworks are widely applicable across CS fields (Table 1).
  - Transitional languages help beginners build foundational knowledge (Table 2).
  - Languages bridging multiple CS fields offer diverse career paths (Table 3).
- **Relevance**:
  - Provides students with a roadmap to focus on key skills and navigate learning paths.
  - Interactive tool aids understanding of study findings.
- **Broader Impact**:
  - Guides curriculum design for academia and emphasizes continuous learning for professionals.
- **Limitations & Future Work**:
  - Dataset may not reflect current trends; newer languages underrepresented.
  - Future studies can expand scope, include qualitative data, and explore additional metrics.



Back to Table of Contents

# Thank you!

# Contributions

1. Taksh Girdhar: Methodology, Results and Discussion, Conclusion

2. Kussh Satija: Abstract, Introduction, Literature Review, Methodology