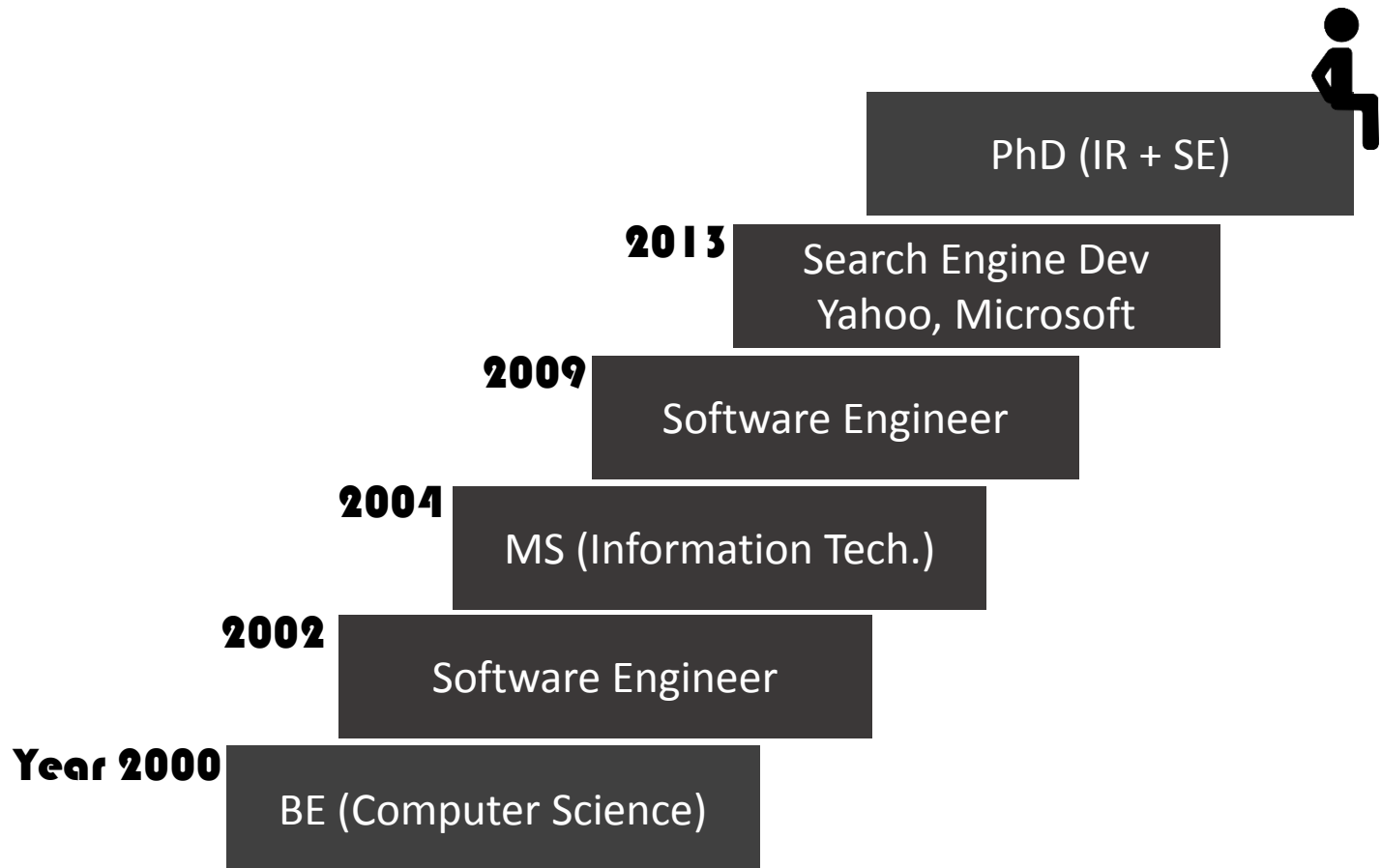


Me and My Research

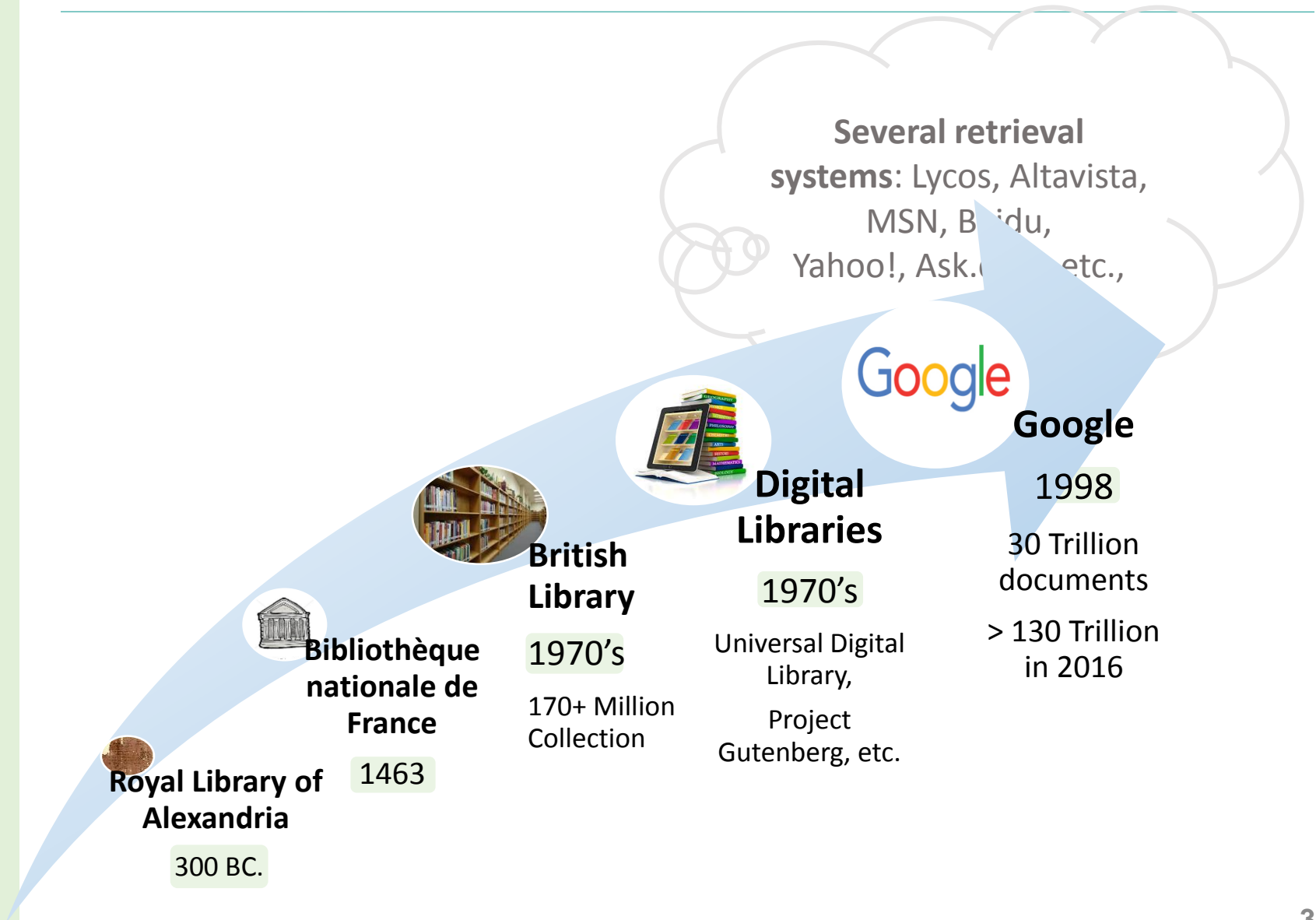
Venkatesh Vinayakarao

For prospective students who wish to do Honors program with me.

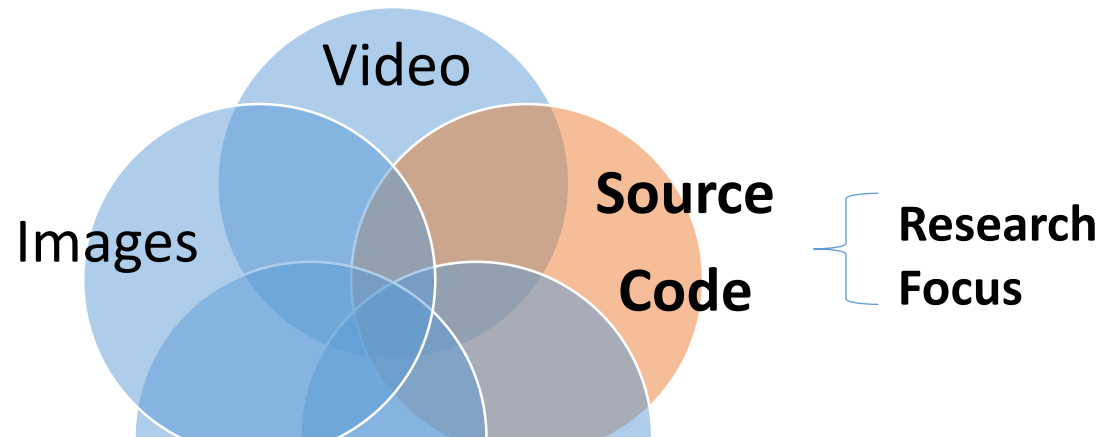
About Me



Growing Information & Information Needs



Content Types



Why Code Search?

Classical program analysis techniques

1

- do not scale well to large set of programs.
- have limitations while working on partial programs.

2

Limitations exist in IDE features for code search and web-scale code search engines, especially, with natural language queries.

3

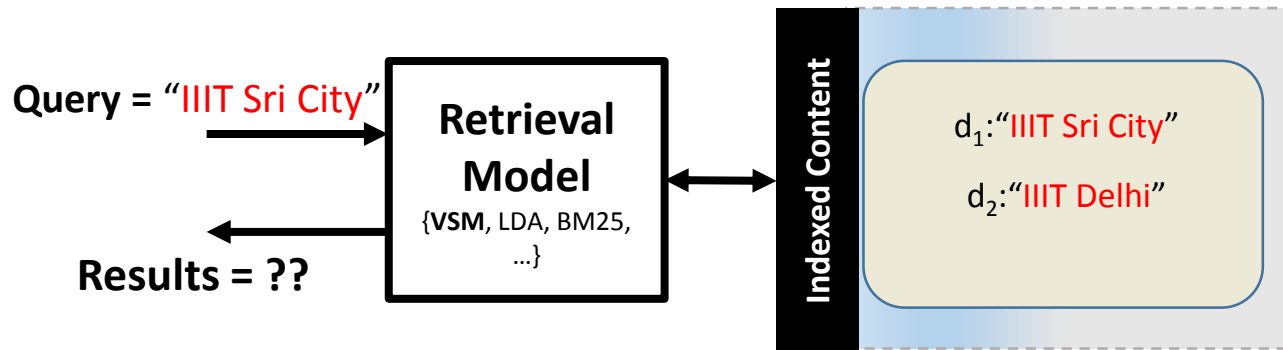
Several software engineering problems which are otherwise difficult to solve can be modeled as code search problems.

My Research

How to Build (Code) Search Engines?

**Information Retrieval
(aka. Search) + Program Analysis + Software Engineering**

Which Document to Retrieve?



Example

Let query $q = \text{“IIIT Sri City”}$.

Let document, $d_1 = \text{“IIIT Sri City”}$ and $d_2 = \text{“IIIT Delhi”}$.

	IIIT	Sri	City	Delhi
q	1	1	1	0
d_1	1	1	1	0
d_2	1	0	0	1

In our VSM, $q = (1,1,1,0)$, $d_1 = (1,1,1,0)$ and $d_2 = (1,0,0,1)$

$$\text{similarity}(d_1, q) = \frac{d_1 \cdot q}{\|d_1\| \|q\|} = \frac{1.1 + 1.1 + 1.1}{\sqrt{1^2+1^2+1^2}\sqrt{1^2+1^2+1^2}} = 1.$$

$$\text{similarity}(d_2, q) = \frac{d_2 \cdot q}{\|d_2\| \|q\|} = \frac{1.1}{\sqrt{1^2+1^2}\sqrt{1^2+1^2+1^2}} = 0.41.$$

My Focus

How to build the retrieval models for queries that are source code (Java, C, etc.)?

Venkatesh Vinayakarao

- Areas of research:
 - Building (Tools for) Search Engines for Source Code.
 - Partial Program Analysis – Working with code snippets that do not compile well.
 - Automating Software Engineering Tasks.
- Prerequisites/constraints:
 - Mandatory:
 - Programming in Java.
 - Should be willing to work on all types of research tasks such as data collection, analysis, implementation, evaluation and writing.
 - Good to have:
 - Taken a course on Information Retrieval, Program Analysis, Software Engineering and/or Compiler Design.
- Number of Honor projects/students:
 - 3