

What does this position entail?

- Do you want to build quantitative models millions of people will use, based on data from the world's largest online laboratory? Are you passionate about formulating relevant questions and producing solutions to initially ill-defined problems? Do the challenges and opportunities of terabytes of data excite you? Can you think abstractly and apply your ideas to the real world? Can you contribute to the big picture and are not afraid to handle the details?
- We are looking for people with the right blend of vision, intellectual curiosity, and hands-on skills, who want to be part of a highly visible, entrepreneurial team

<http://www.ph.tn.tudelft.nl/PRInfo/jobs/msg00185.html>

What is this about?

- Ideal candidates will have a track record of creating innovative solutions, and typically a Ph.D. in computer science, physics, statistics, or electrical engineering. Significant research experience is desired in fields including active learning, probabilistic graphical models and Bayesian networks, data mining and visualization, Web search and information retrieval, judgment and decision making, consumer modeling, and behavioral economics.
- What is data mining? What is machine learning?

What is Computing? Informatics?

- **What is computer science, what is its potential?**
 - What can we do with computers in our lives?
 - What can we do with computing for society?
 - Will networks transform thinking/knowing/doing?
 - Society affecting and affected by computing?
 - Changes in science: biology, physics, chemistry, ...
 - Changes in humanity: access, revolution (?), ...
- **Privileges and opportunities available if you know code**
 - Writing and reading code, understanding algorithms
 - Majestic, magical, mathematical, mysterious, ...

What is the Internet?

- The Internet was originally designed as an "overlay" network running on top of existing phone and other networks. It is based on a small set of software protocols that direct routers inside the network to forward data from source to destination, while applications run on the Internet to rapidly scale into a critical global service. However, this success now makes it difficult to create and test new ways of protecting it from abuses, or from implementing innovative applications and services.

<http://www.intel.com/labs/features/idf09041.htm>

How does the Internet work?

- Differences between the Internet and phone networks
 - Dedicated circuits/routes
 - Distributed, end-to-end
- Where is the intelligence?
 - Not in the network, per se, in the design and the ends
 - End-to-end Arguments in System Design
- Success of email, web, etc., relies on not building intelligence into the network
 - What about overlay networks?
 - What about PlanetLab?

Find all unique/different words in a file

Across different languages: do these languages have
the same power?

Unique Words in Python

```
#! /usr/bin/env python

def main():
    f = open('/data/melville.txt', 'r')
    words = f.read().strip().split()
    allWords = set()
    for w in words:
        allWords.add(w)
    for word in sorted(allWords):
        print word

if __name__ == "__main__":
    main()
```

Unique words in Java

```
import java.util.*;
import java.io.*;
public class Unique {
    public static void main(String[] args)
        throws IOException{
        Scanner scan =
            new Scanner(new File("/data/melville.txt"));
        TreeSet<String> set = new TreeSet<String>();
        while (scan.hasNext()) {
            String str = scan.next();
            set.add(str);
        }
        for(String s : set) {
            System.out.println(s);
        }
    }
}
```

Unique words in C++

```
#include <iostream>
#include <fstream>
#include <set>
using namespace std;

int main() {
    ifstream input("/data/melville.txt");
    set<string> unique;
    string word;
    while (input >> word) {
        unique.insert(word);
    }
    set<string>::iterator it = unique.begin();
    for(; it != unique.end(); it++) {
        cout << *it << endl;
    }
    return 0;
}
```

Unique words in PHP

```
<?php

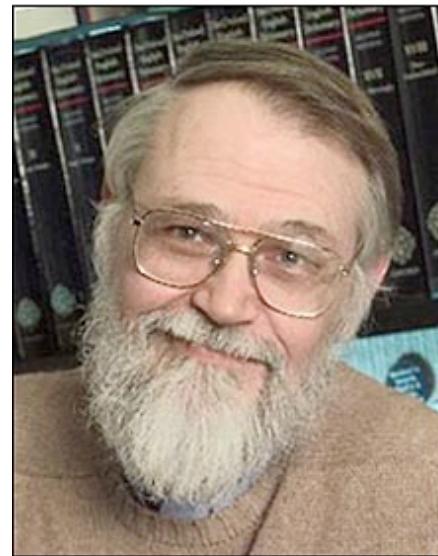
$wholething = file_get_contents("file:///data/melville.txt");
$wholething = trim($wholething);

$array = preg_split("/\s+/", $wholething);
$uni = array_unique($array);
sort($uni);
foreach ($uni as $word) {
    echo $word."<br>";
}

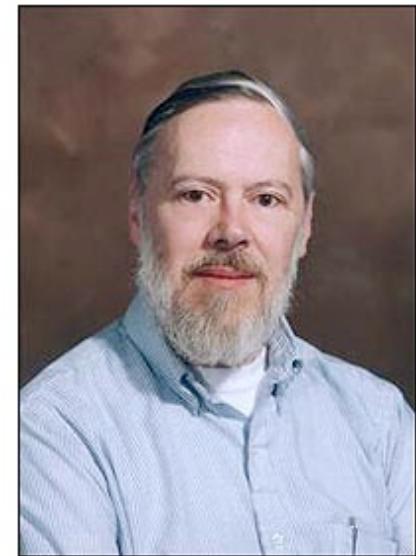
?>
```

Kernighan and Ritchie

- First C book, 1978
- First ‘hello world’
- Ritchie: Unix too!
 - Turing award 1983
- Kernighan: tools
 - Strunk and White



Brian Kernighan



Dennis Ritchie

- Everyone knows that debugging is twice as hard as writing a program in the first place. So if you are as clever as you can be when you write it, how will you ever debug it?

Brian Kernighan