

# Introduction

---

PERSONAL INTRODUCTION: [HTTP://DATA.SCIENCE.UOIT.CA/](http://data.science.uoit.ca/)

DATABASE SYSTEMS: HISTORY – PRESENT – FUTURE

DATABASE SYSTEMS: WHAT IT MEANS FOR THIS COURSE.

DATABASE SYSTEMS: A GLIMPSE

# CSCI 3030U

## Database Systems and Concepts

---

**Q: What is this course about?**

A:  
\$ psql uoit  
SELECT description  
FROM ac\_course  
WHERE course\_code = 'CSCI 3030U';

**Q: Is it fun?**

A: Definitely

**Q: Is it useful?**

A: Definitely

The aim of the course is to provide students with an overview of database management system architectures and environments, an understanding of database design and implementation techniques, and practical experience of designing and building a relational database.

# Things we cover:

---

Elements of the relational data model

Theory of relational query languages

The Structured Query Language (SQL)

Relational database management using Postgres: performance tuning and access control.

Application programming interface of RDBMS for host languages: Java and Python

# Things we cover:

---

Elements of the semi-structured data model



XML



Xquery based query engine



JSON



Application Programming Interface for host languages: Python (and Java)

# Things we cover:

---

Advanced topics:

Data mining

Data visualization

Database driven Web application programming

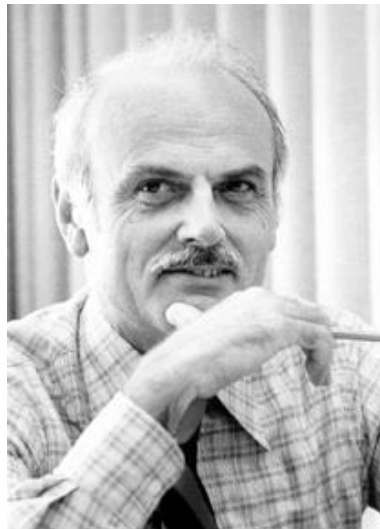
# History

---

IBM was founded in 1896 as TMC (Tabulating Machine Company) by Herman Hollerith.



Edgar Codd invents the Relational Data Model, and its first order theory. IBM team implements System R. 1970



Larry Ellison implements Oracle from the System R paper, and markets Oracle. 1978



# History

---

Google implements its own Big Table to store the entire WWW. Big Table was designed and implemented by Jeffery Dean and Sanjay Ghemawat. 2000



Facebook and eBay deploys a radically different family of data storage engines, known as NoSQL. CouchDB is implemented by Damien Katz (former IBM engineer), 2005



Relational database engines are scaled down to be embedded in mobile devices: Android and iPhone. SQLite is used by both smart phone OS. 2008  
SQLite was implemented by Richard Hipp, for on board data management of Navy missile system.



# Universality of Database Management

---

iPhone calendar	30,000 bytes
UOIT course database	30,000,000 bytes
Audio collection of Beethoven	30,000,000,000 bytes
Printed collection of US Library of Congress	30,000,000,000,000 bytes
Data processed by Google (big table) per day	30,000,000,000,000,000 bytes
Total global Internet traffic per month	30,000,000,000,000,000,000 bytes

---

Almost all (all except the last one) are stored in databases, and queried using a common language **!**

<http://en.wikipedia.org/wiki/Petabyte>

<http://en.wikipedia.org/wiki/Exabyte>

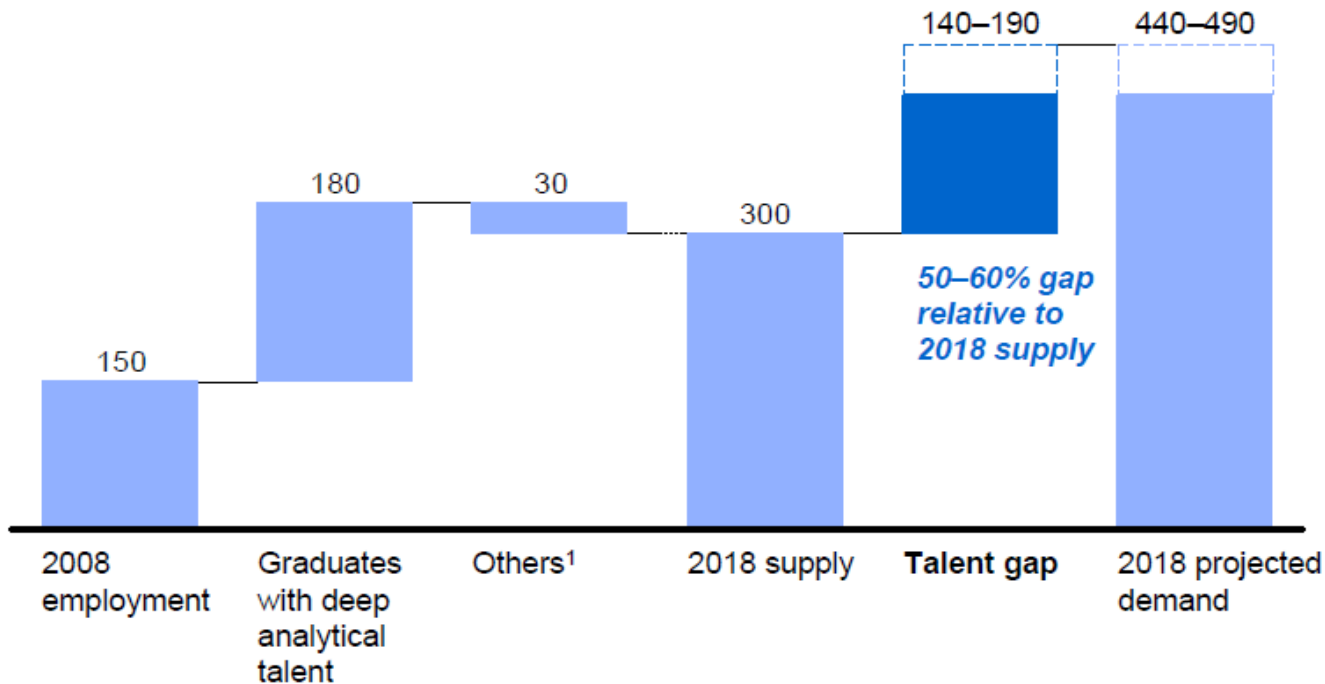
[http://www.jamesshuggins.com/h/tek1/how\\_big.htm](http://www.jamesshuggins.com/h/tek1/how_big.htm)



# Good News: Demand for Data Science!

**Demand for deep analytical talent in the United States could be 50 to 60 percent greater than its projected supply by 2018**

Supply and demand of deep analytical talent by 2018  
Thousand people



<sup>1</sup> Other supply drivers include attrition (-), immigration (+), and reemploying previously unemployed deep analytical talent (+).

SOURCE: US Bureau of Labor Statistics; US Census; Dun & Bradstreet; company interviews; McKinsey Global Institute analysis

# Facebook Country...

---

1. China (1.339 billion)
2. India (1.218 billion)
3. Facebook (1 billion)

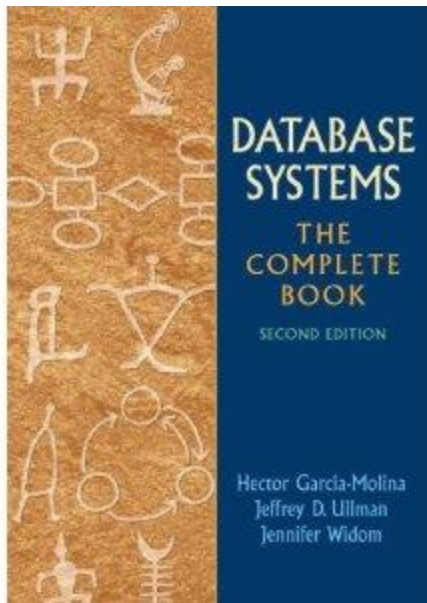


# Text book:

---

Database Systems, The Complete Book

Hector Garcia-Molina, Jeffery Ullman and Jennifer Widom

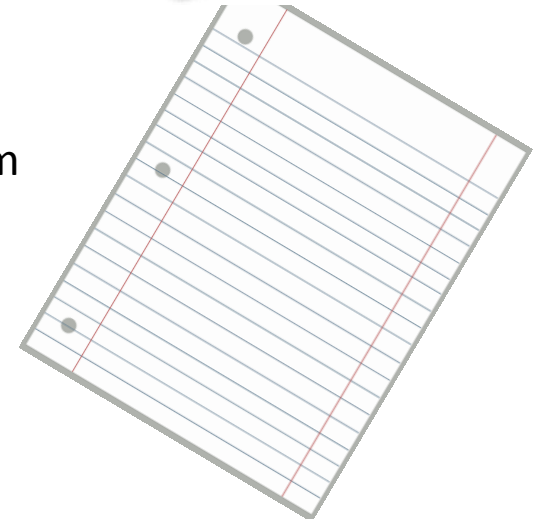
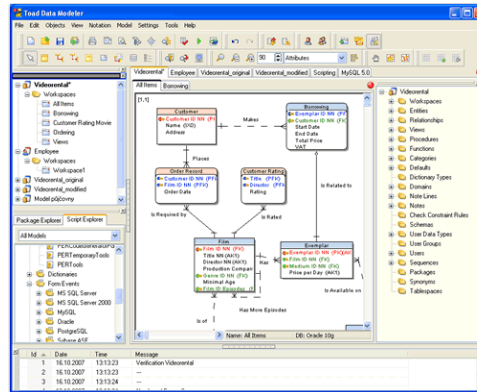
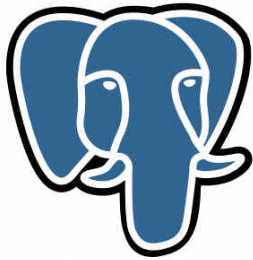


It is definite source of reference if your future work touches on database management.

The book is a *text book* not a reference manual. So, you won't find indepth reference on PHP, ODBC (though introductory materials are included in the book).

# Equipments:

PostgreSQL



PostgreSQL relational datababase management system (RDBMS) running on your laptop.

# Structure of the course

---

Lectures: Twice a week, refer to [www.uoit.ca/mycampus](http://www.uoit.ca/mycampus)

Labs: Once a week (**First lab in the week 16th-20th of Jan.**)

Midterm: One (1+1)

Final: One

---

## Marking:

Labs and Project 30% (10% Labs + 20% Project)

Midterm 20%

Final 40%

Participation & Presentation 10%

# Correspondence

---

Use Blackboard and Slack to discuss things with other students

- Ask questions and exchange knowledge!

Take time when composing a message - think of it as a professional message to a co-worker.

- There will not be space for SMS-speak in your work life.

Use e-mail for correspondence: [jaroslaw.szlichta@uoit.ca](mailto:jaroslaw.szlichta@uoit.ca)

TA: Alexander Keller: [alexander.keller@uoit.net](mailto:alexander.keller@uoit.net)

# A glimpse of the course...

---

## Movie database.

- Designing movie database (Toad Data Modeler)
- Storage (Postgresql)
- Querying (SQL, ...)

# A glimpse of the course: It's all about getting answers...

---

Which movie has the best rating?

Who is the director of „Beatiful Mind” movie?

Which movie is the longest?

...

In which movies was Brad Pitt playing?



# Your Action Items

---

Get a textbook!