Inverse methods in the era of machine learning and deep learning (Part I)

Prof. Sang-Mook Lee

Seoul National University

smlee@snu.ac.kr

The Abdus Salam International Centre for Theoretical Physics

International Union of Geodesy and Geophysics (IUGG)

SNUMGGL

Seoul National University Marine Geophysics & Geodynamics Lab

Outline

Part 1

- A bit about myself (before and after 2006 accident)
  - Awaken & new outlook on life
- Quality of Life Technology and Computational Sciences Programs at SNU
  - Data Science and Numerical Analysis
- How programming has changed as an old school
  - Stack, Heap, Pointers, Objects
- Computer is God's gift for people with disability (voice coding)
- JEDI and raising skilled program developers among people with disability

Part 2

- Earth Science is about data and observation (new era of global open ocean research)
- Geophysical Inverse Theory in the era of ML and DL
  - New perspectives but are limited in their application
  - Neural Network is like a black box and we never have enough data
- Planet A as part of SNU's 10-10 Initiative
  - Using global data hubs to address societally important issues
- SK E&S (producing startups in global environment and energy sectors)
- Data Science Hackathon and Advanced Instrumentation Competition

Before and after the Accident

: 삶의 매순간은 신성하다

나는 항상 이런 생각을 한다.
사고를 통해 장애를 입었지만, 다시 재기해 활동하는 데 필요한 최소의 부분은 하늘이 가져가지 않았다고.
나는 언제나 운이 좋았다. 지금도 예전과 마찬가지로 나는 하늘이 내린 행운을 누리고 있다.

0.1그램의 희망

| 이상묵 · 강인식 지음 |

〈뉴욕 타임스〉 등 세계적 언론이 주목한 인간 승리의 드라마

한국의 '스티븐 호킹' 이상묵 서울대 교수
전신마비의 장애를 딛고 삶의 희망을 쏘아 올리다!

: : 이 책의 판매로 발생하는 저자의 수익금 전액은 서울대 이혜정 장학금에 기부됩니다.

랜덤하우스

Having earned PhD from US, I was appalled by the use of the state-of-the-art research vessel on survey of Mn nodules in the Pacific as opposed to exiting basic science (e.g. mid-ocean ridges and deepsea discoveries).

Things have changed over 20 years but one that remains is understanding of the importance of science by public and government.

Science Policy & Policy for Science

InterRidge Update

Rifts. Rifted Margins.
Spreading Ridges.
Planning Committee

Prof. Sang-Mook Lee

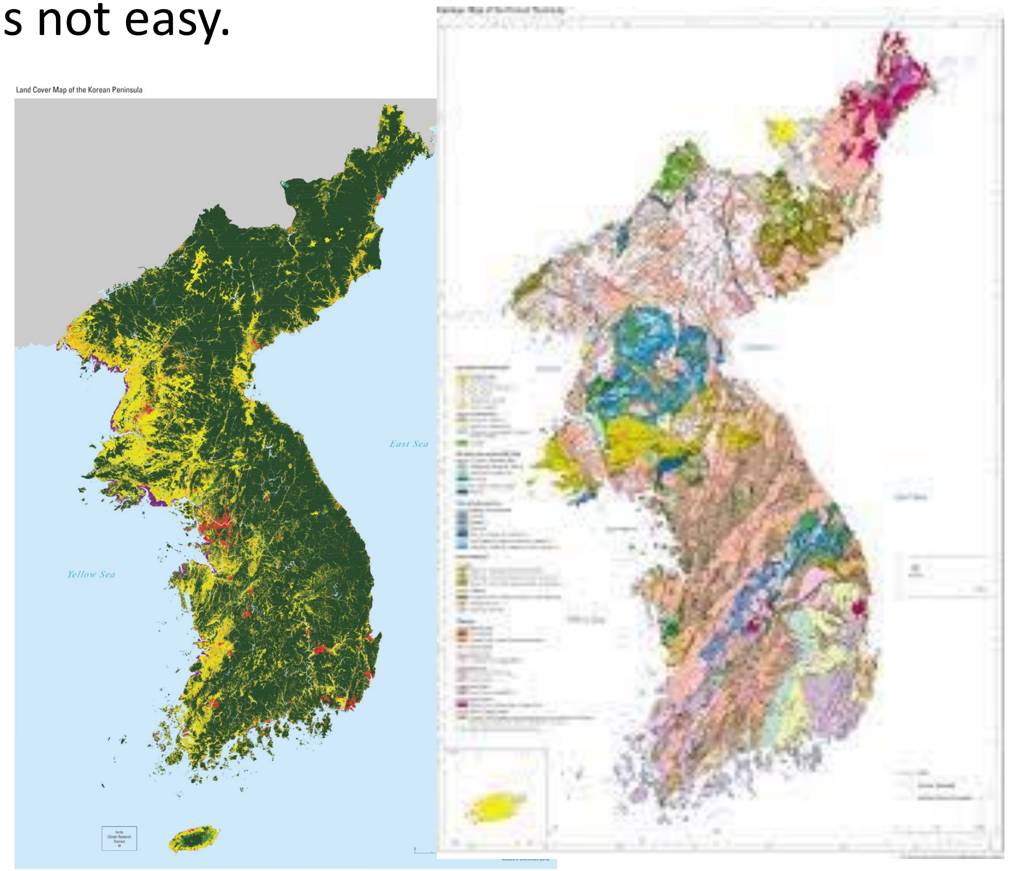
Seoul National University

InterRidge Chair, Geology/Geophysics (2020 -)

www.interridge.org

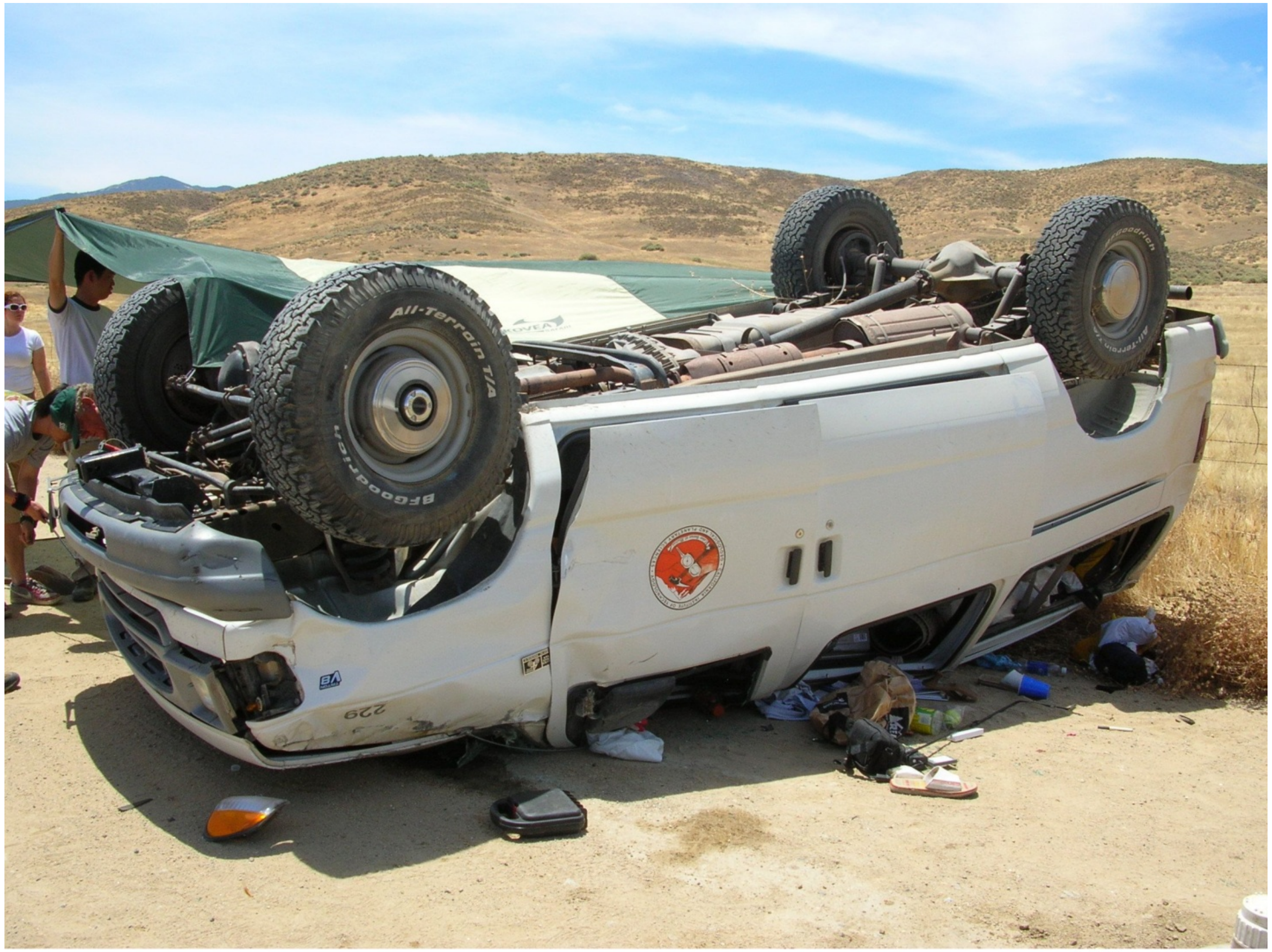
After 12 years abroad and 7 years at Korea Ocean Research and Development Institute running global ocean open operations, I returned to my alma mater (Seoul National University) in late 2003.

But teaching active geology in Korea was not easy.



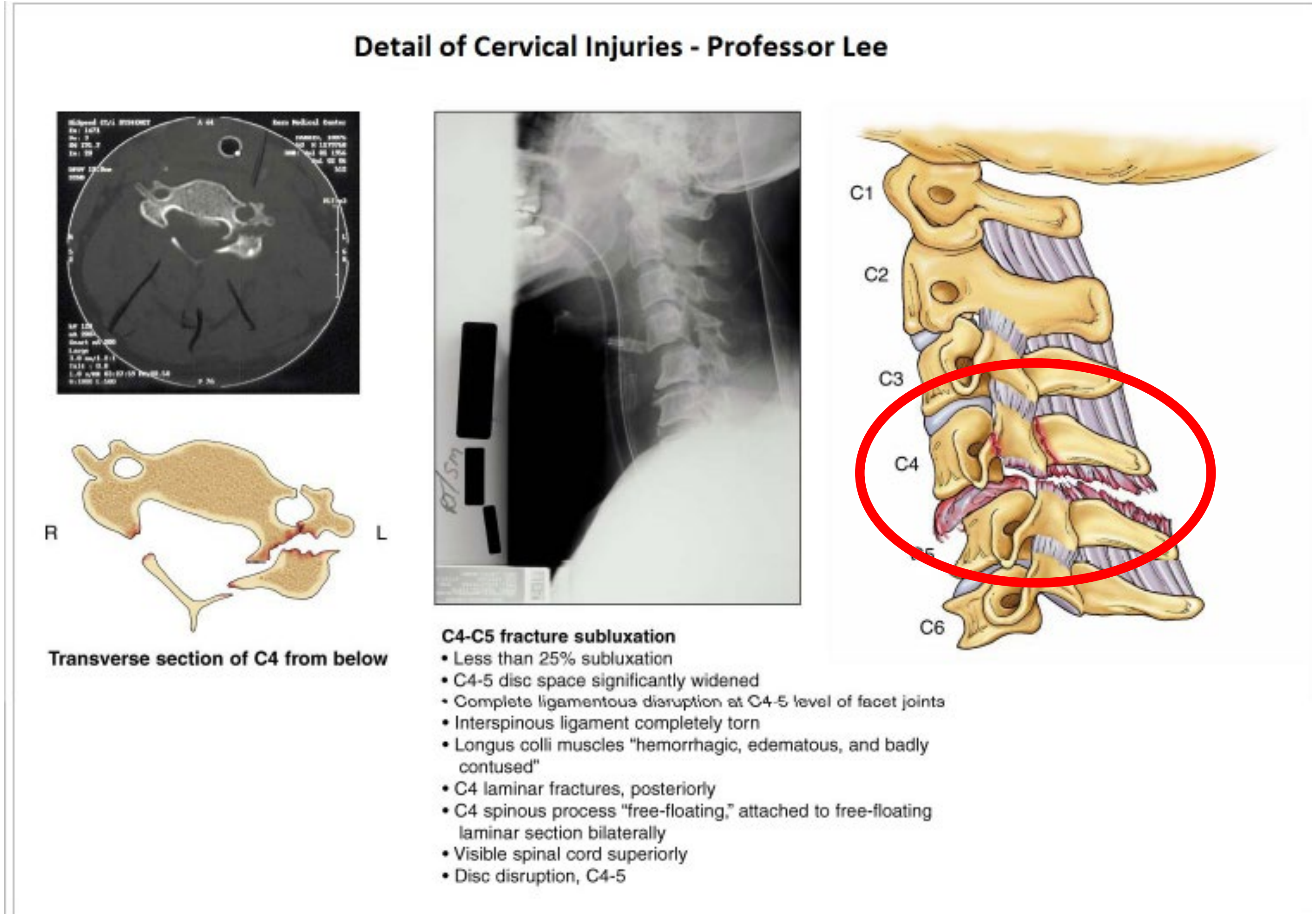
2006 6 25
WARNING

During geological field trip in California, 2006



EXH.0005-A

SCI “Neck fusion”



Detail of Cervical Injuries - Professor Lee

Transverse section of C4 from below

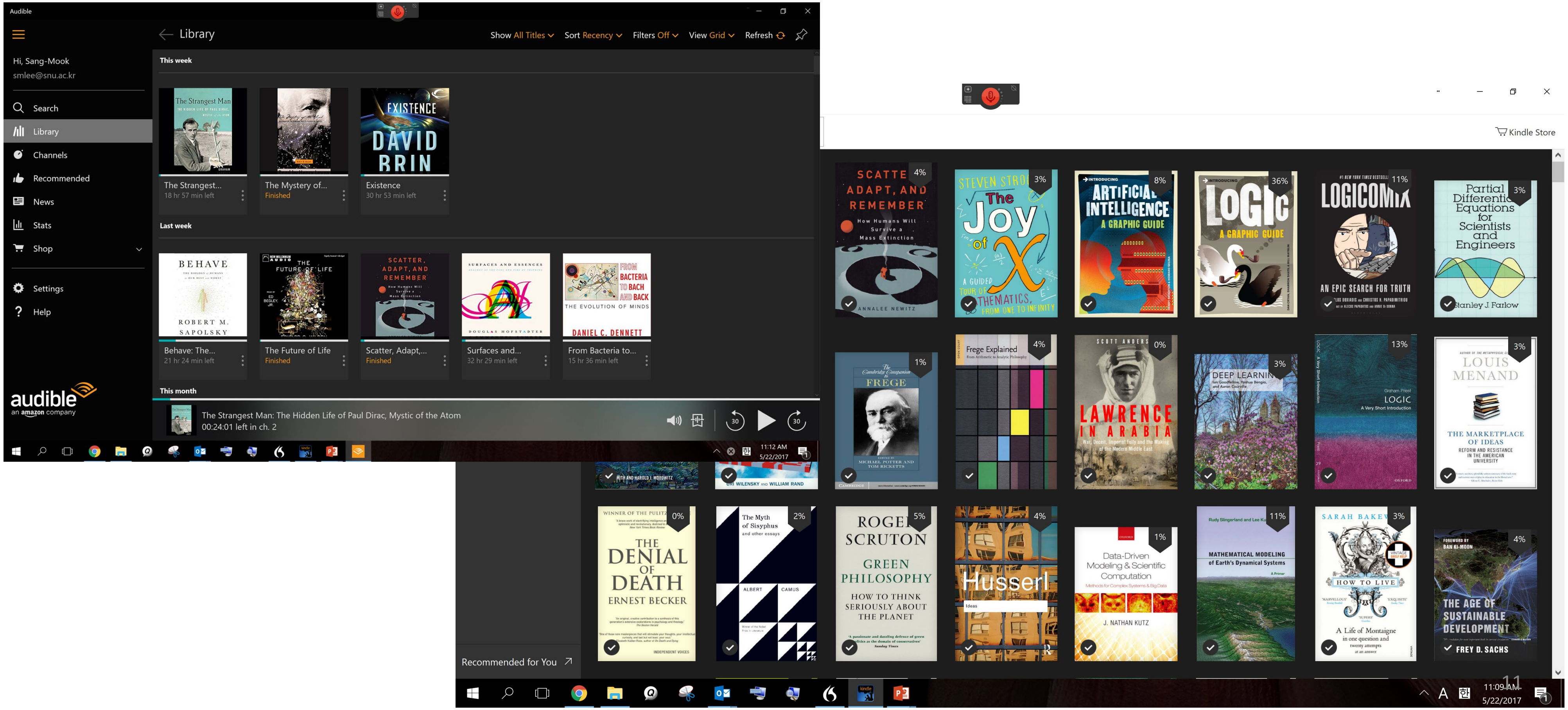
C4-C5 fracture subluxation

- Less than 25% subluxation
- C4-5 disc space significantly widened
- Complete ligamentous disruption at C4-5 level of facet joints
- Interspinous ligament completely torn
- Longus colli muscles “hemorrhagic, edematous, and badly contused”
- C4 laminar fractures, posteriorly
- C4 spinous process “free-floating,” attached to free-floating laminar section bilaterally
- Visible spinal cord superiorly
- Disc disruption, C4-5

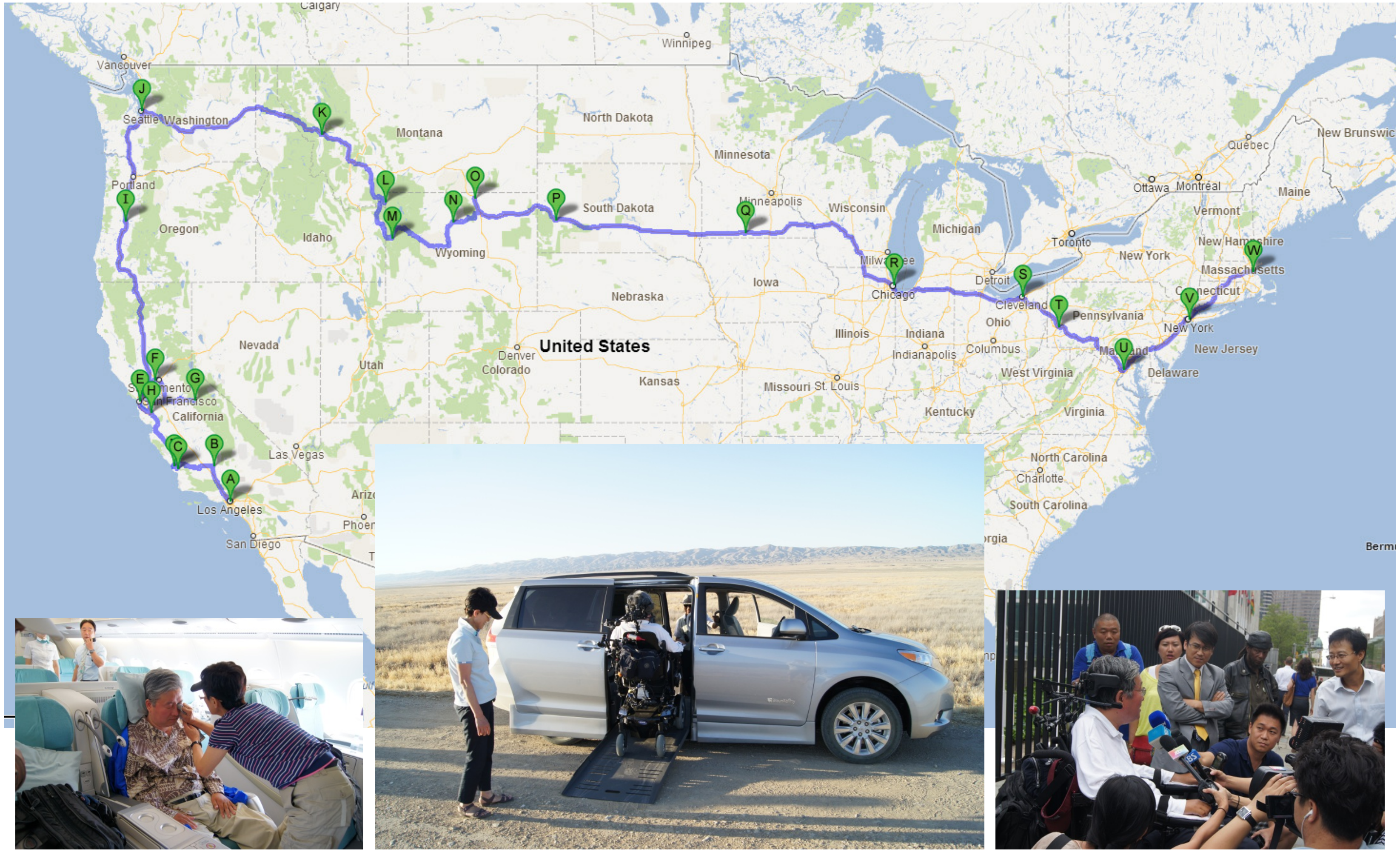
What would I have missed by dying early?

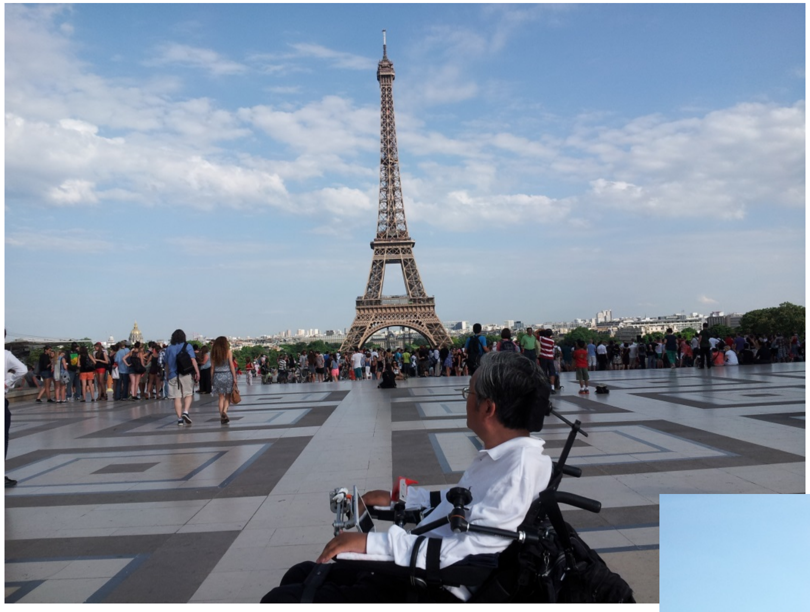
I don't want to believe. I want to know.
- Carl Sagan

I cannot flip a page but with special assistive device I can read books and listen to audio books. You have all the knowledge you want there. It is in life amazing?



40 days across United States – Now or Never

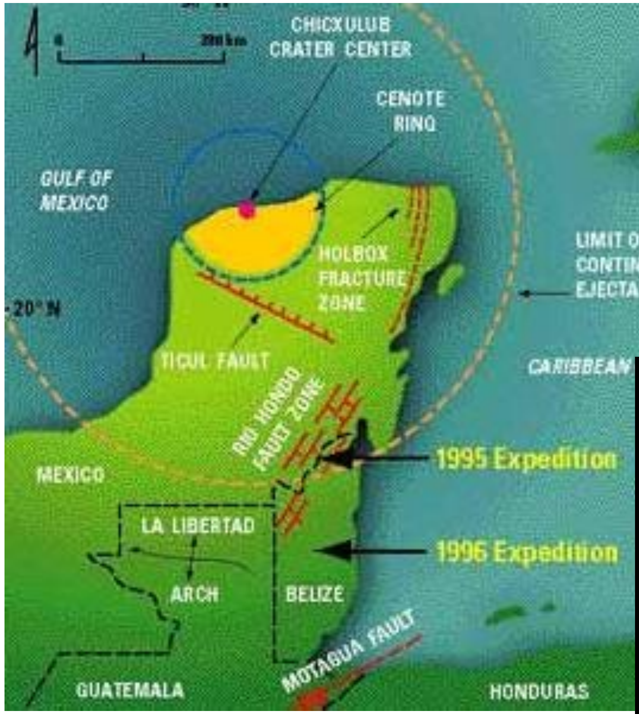


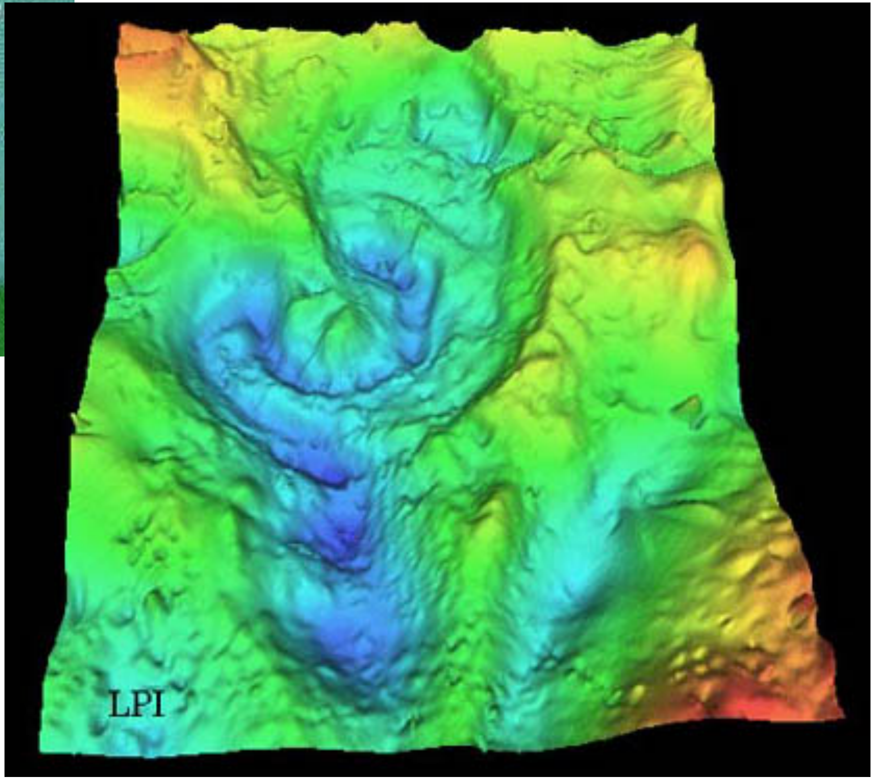


Paris

Berlin

Chicxulub, Mexico





2013 Merida Mexico

IUGG

IAGA

International Association of Geomagnetism and Aeronomy

Mayan
Heritage

Developing assistive technology apps for people with disability sponsored by LG Electronics and Her Highness Jameela Al Qasimi

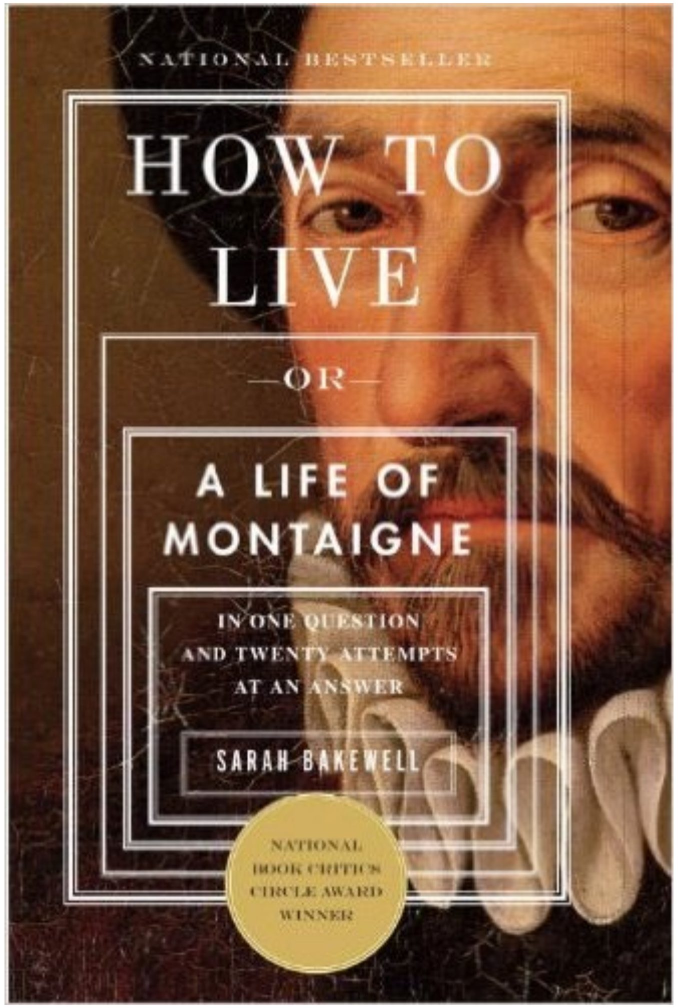
AT EDUCOM 2017 at Sharjah UAE

PyeongChang Forum

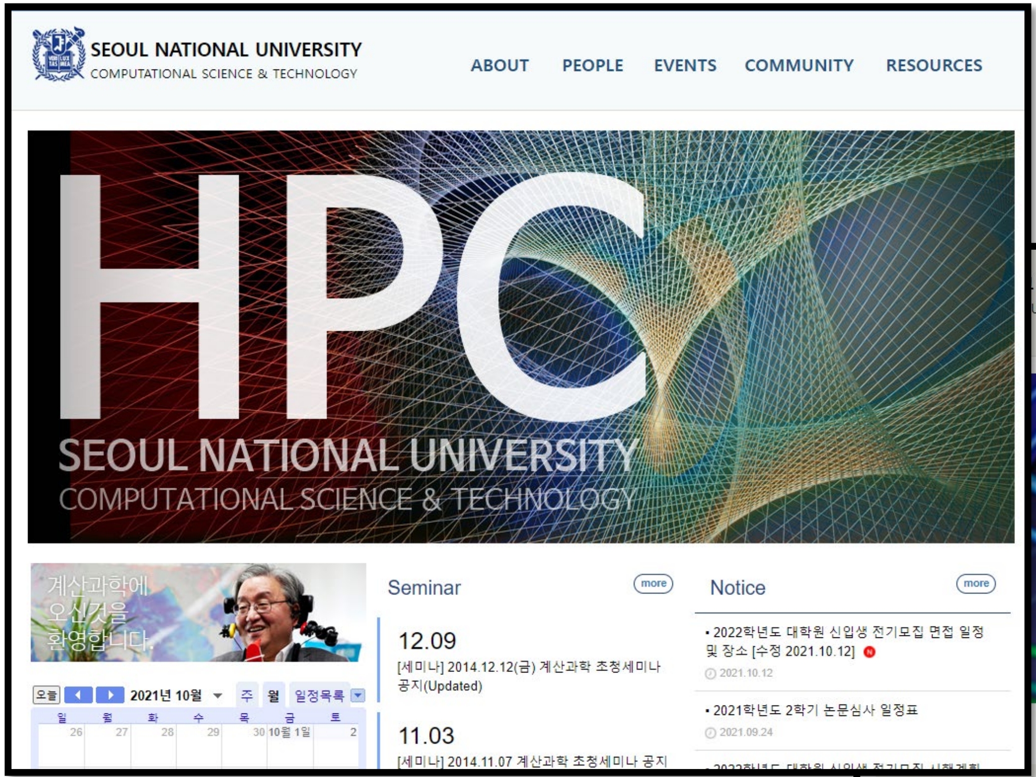
February 13-15, 2019

At the Limit of Our Knowledge, Staring into the Future

포럼 2018
A New Solidarity for the Earth and Its Stewardship
GWCVB 강원컨벤션뷰로
Gangwon Convention & Visitors Bureau
한국관광공사
KOREA TOURISM ORGANIZATION


NATIONAL BESTSELLER
HOW TO LIVE
—OR—
A LIFE OF MONTAIGNE
IN ONE QUESTION AND TWENTY ATTEMPTS AT AN ANSWER
SARAH BAKEWELL
NATIONAL BOOK CRITICS CIRCLE AWARD WINNER

Interdisciplinary Program of Computational Science and Technology (Graduate)



SEOUL NATIONAL UNIVERSITY
COMPUTATIONAL SCIENCE & TECHNOLOGY

오늘 2021년 10월 주 월 일정목록

일	월	화	수	목	금	토
26	27	28	29	30	10월 1일	2

Seminar more

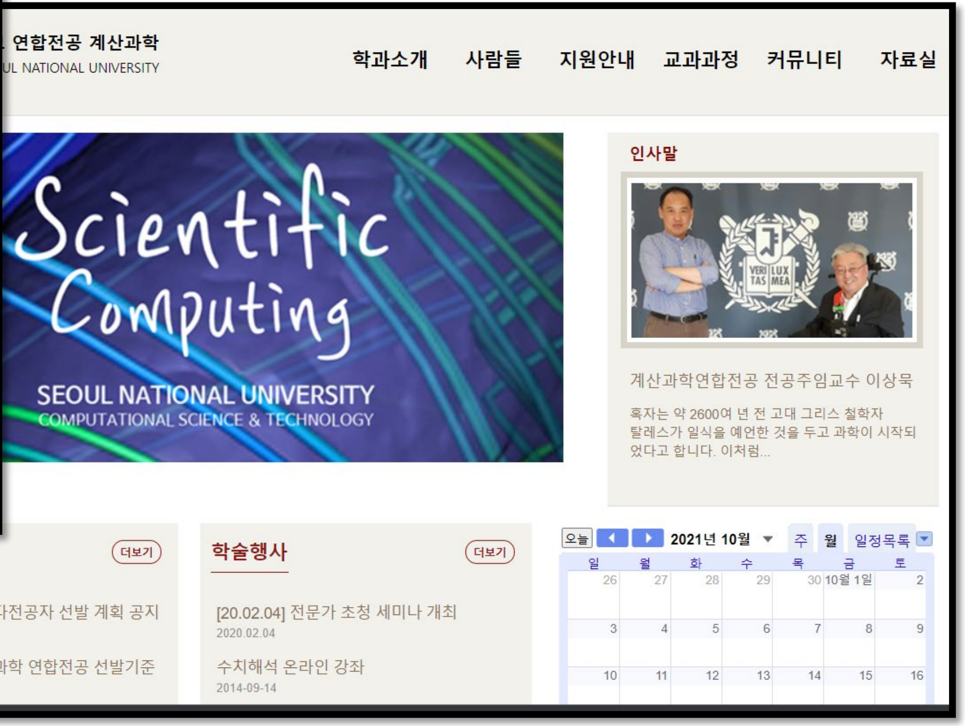
12.09
[세미나] 2014.12.12(금) 계산과학 초청세미나 공지(Updated)

11.03
[세미나] 2014.11.07 계산과학 초청세미나 공지

Notice more

- 2022학년도 대학원 신입생 전기모집 면접 일정 및 장소 [수정 2021.10.12]
  2021.10.12
- 2021학년도 2학기 논문심사 일정표
  2021.09.24

Undergraduate Interdisciplinary Program in Computational Sciences



연합전공 계산과학
UL NATIONAL UNIVERSITY

인사말

계산과학연합전공 전공주임교수 이상묵

혹자는 약 2600여 년 전 고대 그리스 철학자 탈레스가 일식을 예언한 것을 두고 과학이 시작되었다고 합니다. 이처럼...

공지사항 더보기

2022학년도 1학기 다전공자 선발 계획 공지
2021.09.14

2022년 1학기 계산과학 연합전공 선발기준
2021.09.14

학술행사 더보기

[20.02.04] 전문가 초청 세미나 개최
2020.02.04

수치해석 온라인 강좌
2014-09-14

오늘 2021년 10월 주 월 일정목록

일	월	화	수	목	금	토
26	27	28	29	30	10월 1일	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16

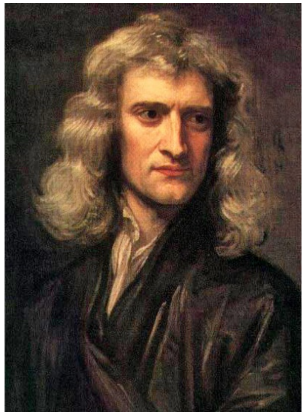
Quality of Life Technology - US$35 million project funded by the Ministry of Knowledge Economy for people with disability

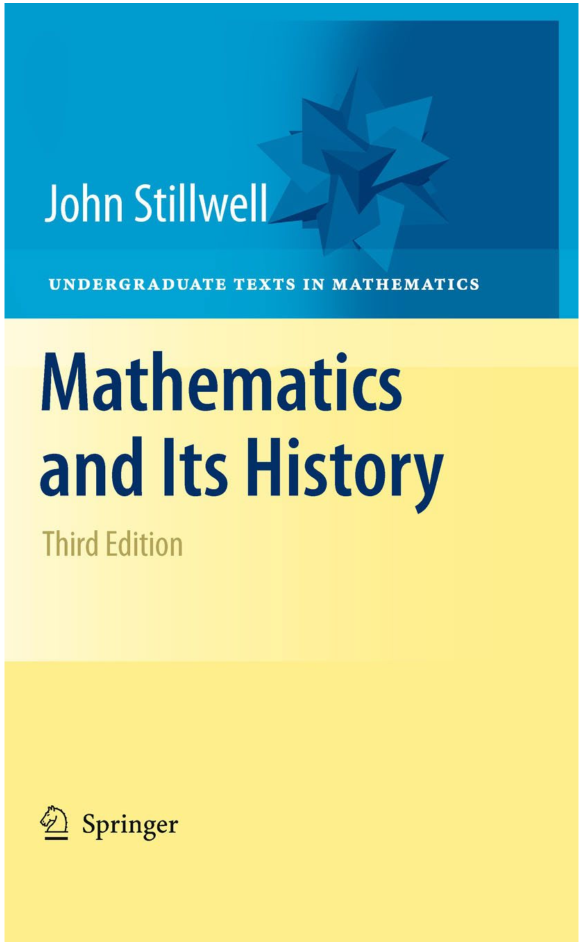
Undergraduate Curriculum

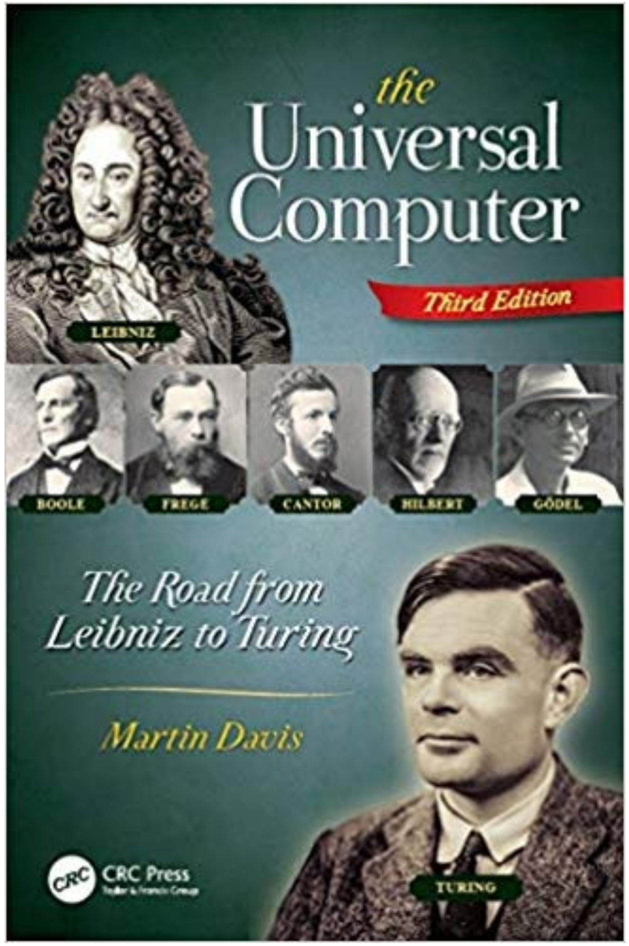
학년	1학기 교과목 및 학점	2학기 교과목 및 학점
2학년	• 계산과학의 이해 (Understanding Computational Sciences) • 계산과학의 기초와 역사 (History and Foundation of Computational Sciences) • 계산과학이론 및 실습 1 (Theory and Practices of Computational Sciences 1)	• 계산과학이론 및 실습 2 (Theory and Practices of Computational Sciences 1)
3학년	• 데이터과학 (Data Sciences)	• 과학계산개론 (Introduction to Scientific Computing)
4학년		• 계산과학 주제연구 (Special Topics On Computational Sciences) • 계산과학 종합설계 (Capstone Research on Computational Sciences)

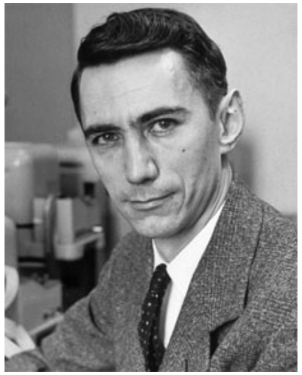
6 Prerequisites and 2 Electives

Foundation and History of Computational Sciences




John Stillwell
UNDERGRADUATE TEXTS IN MATHEMATICS
Mathematics and Its History
Third Edition
Springer





Leibniz
Boole
Frege
Cantor
Russell
Hilbert
Gödel
Turing
Shannon

Numerical Analysis

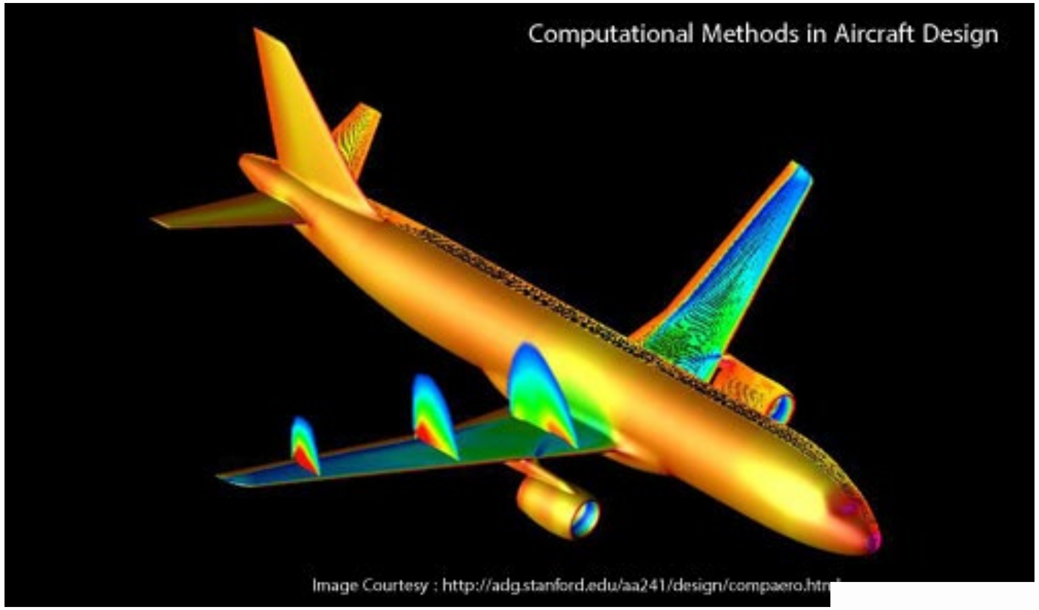
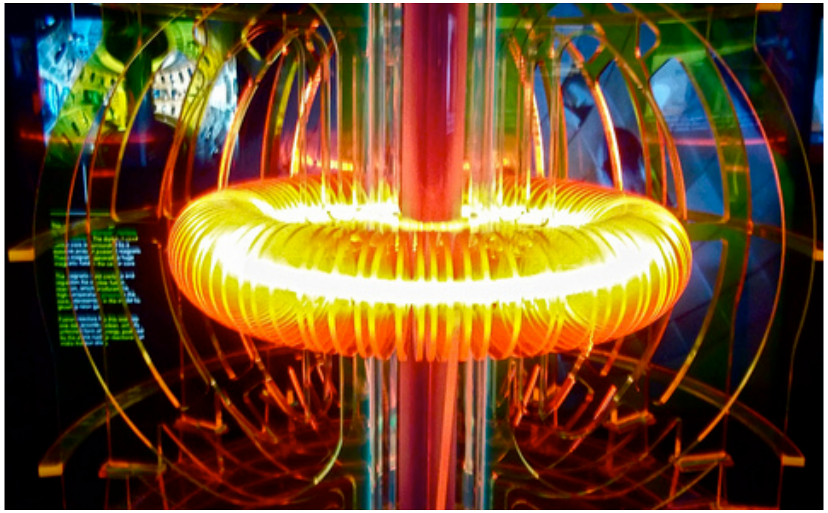
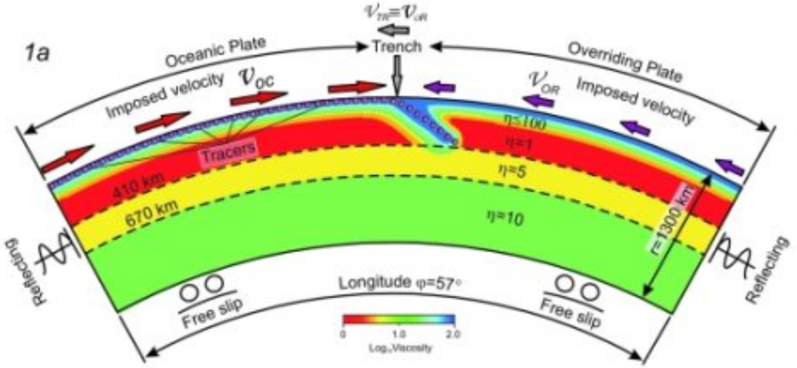
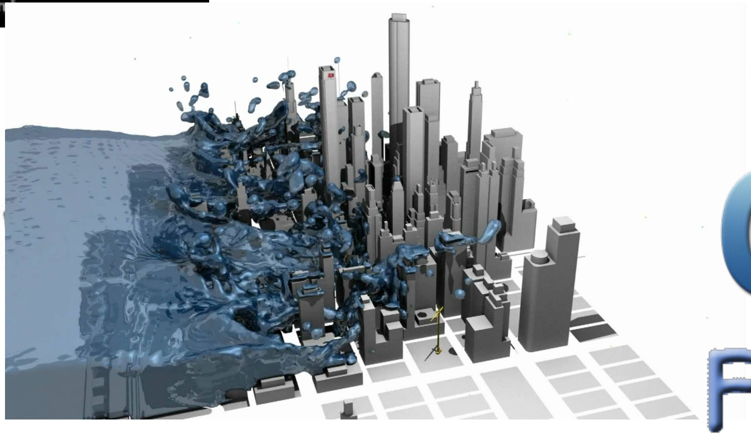


Image Courtesy : http://adg.stanford.edu/aa241/design/compaero.html





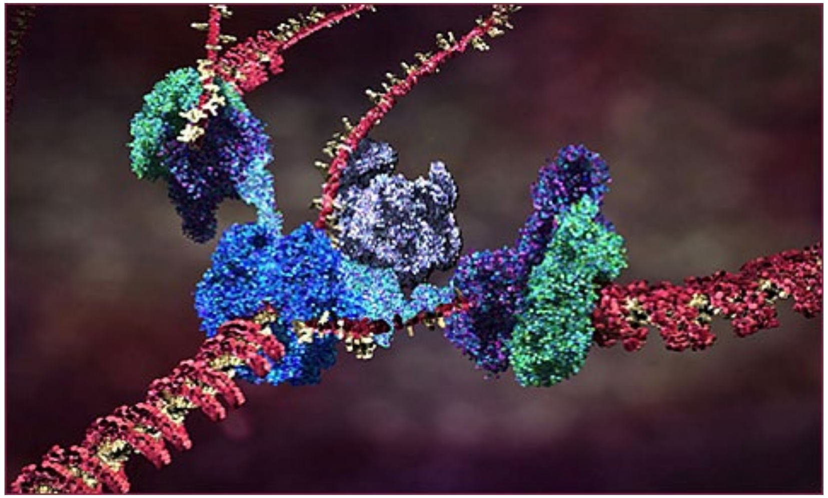


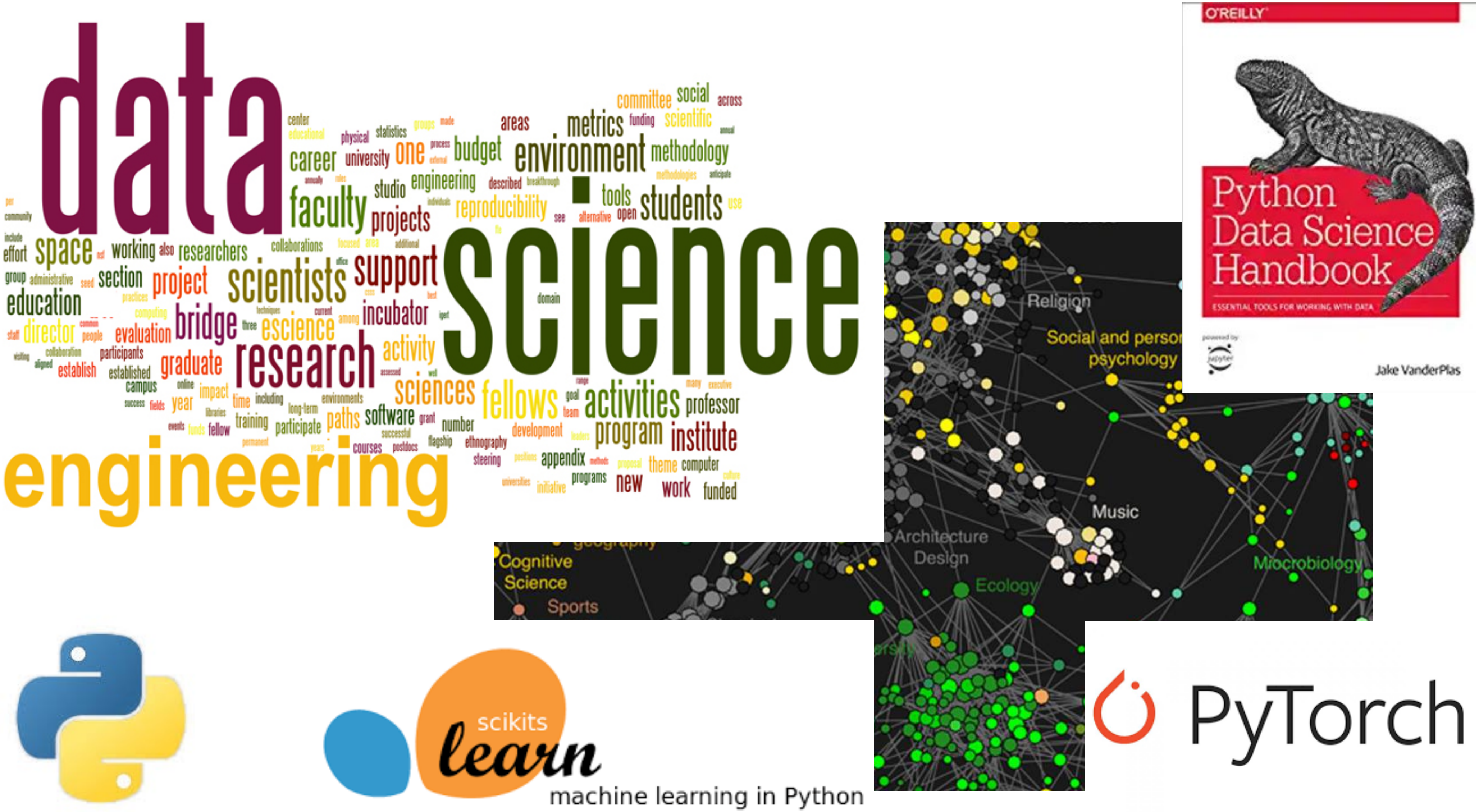
C/C++

Fortran 95

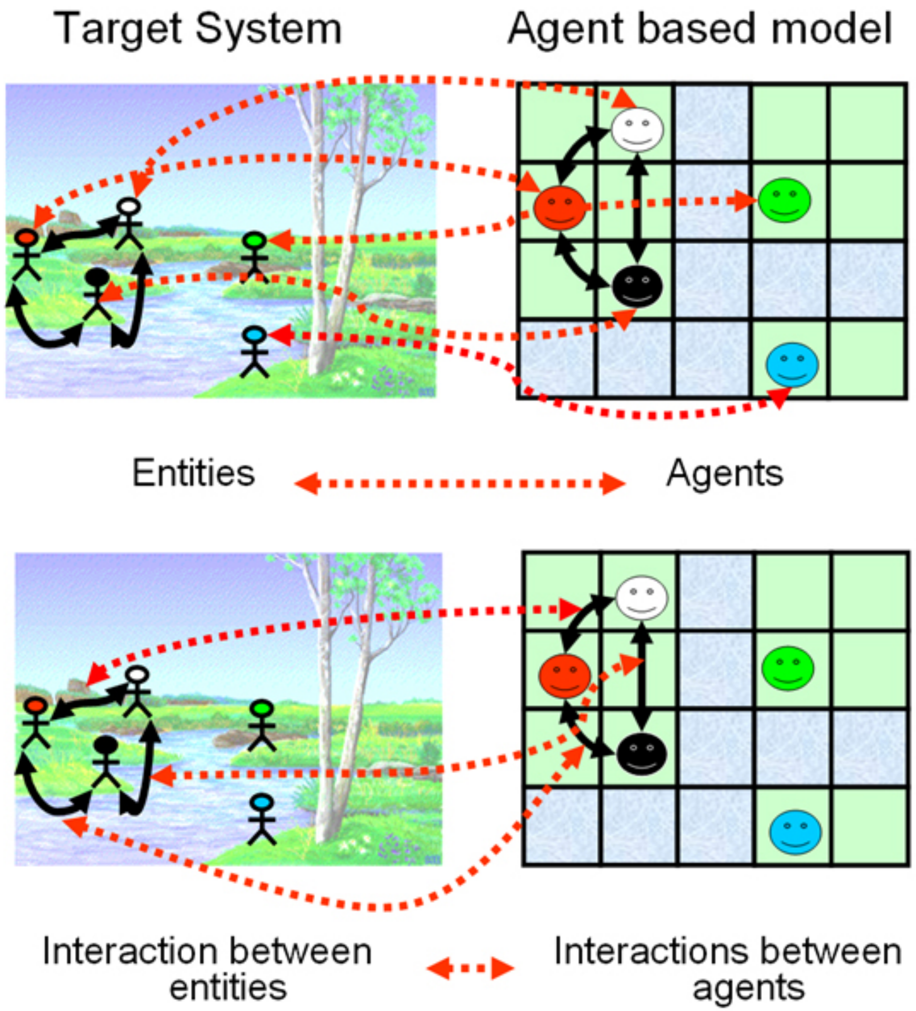
Scientific Visualization

2013년 7월,
대한민국
슈퍼스타가
온다!
G
MR. GO
〈국가대표〉 김용화 감독
미스터 고
Mr.GO 3D
김용화 감독 / 링링 성동일 서교 /
제공/배급 쇼박스㈜미디어플렉스 제작 덱스터스튜디오 공동제작 ㈜광대무변 공동제공 ㈜화이브라더스




data
science
engineering
research
scientists
support
faculty
students
environment
metrics
methodology
reproducibility
education
space
bridge
project
incubator
activity
sciences
fellows
activities
program
institute
professor
escience
graduate
director
evaluation
researchers
career
university
one
budget
engineering
projects
studio
tools
committee
social
across
scientific
areas
physical
statistics
center
collaborations
working
section
effort
group
participants
establish
established
campus
year
impact
time
training
participate
paths
software
number
ethnography
development
appendix
programs
new
work
funded
theme
computer
courses
postdocs
steering
initiative
Religion
Social and persor
psychology
Music
Architecture
Design
Ecology
Microbiology
geography
Cognitive
Science
Sports
O'REILLY
Python
Data Science
Handbook
ESSENTIAL TOOLS FOR WORKING WITH DATA
powered by
jupyter
Jake VanderPlas
scikits
learn
machine learning in Python
PyTorch

Complex phenomena

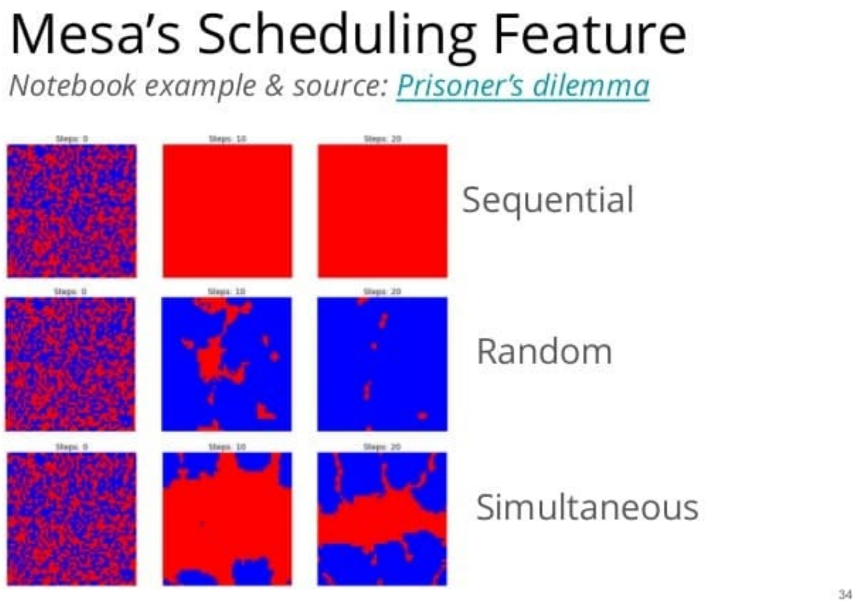


Target System

Agent based model

Entities ↔ Agents

Interaction between entities ↔ Interactions between agents



Mesa's Scheduling Feature

Notebook example & source: Prisoner's dilemma

Sequential

Random

Simultaneous

34

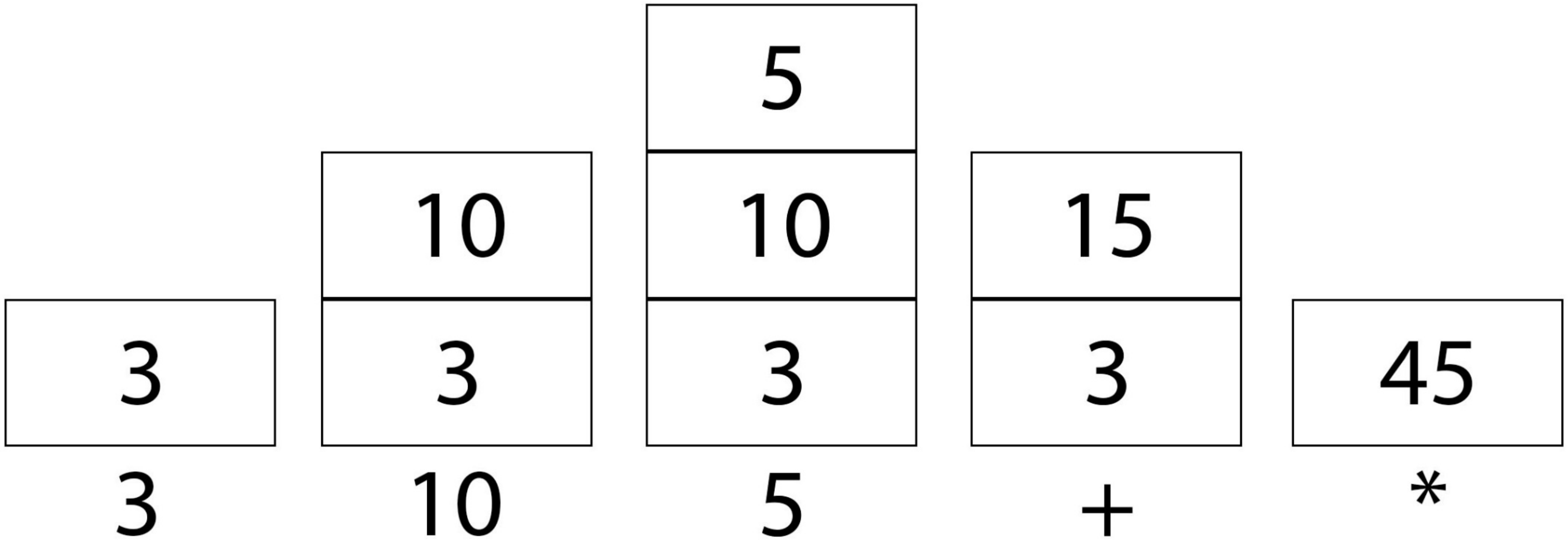
SANTA FE INSTITUTE

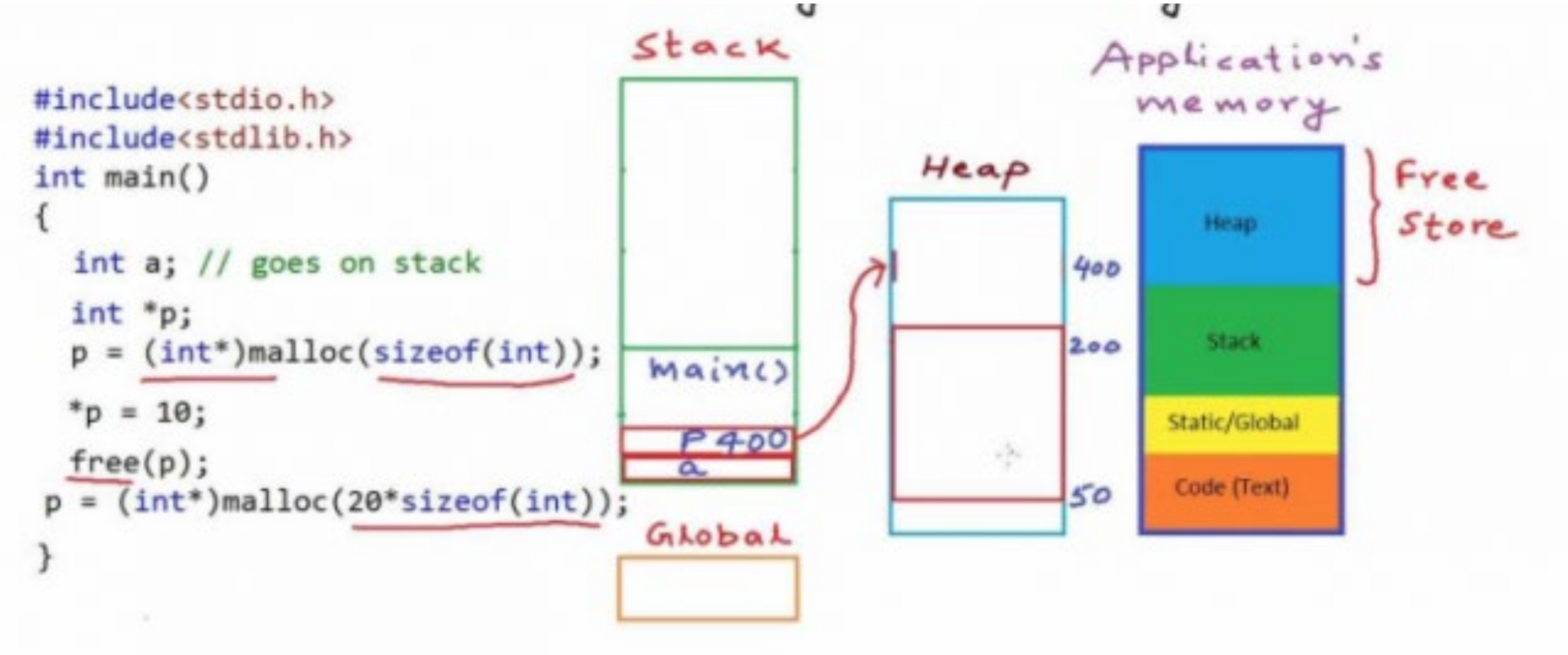
I'm an old-school and had to reinvent myself.

Stack and Heap

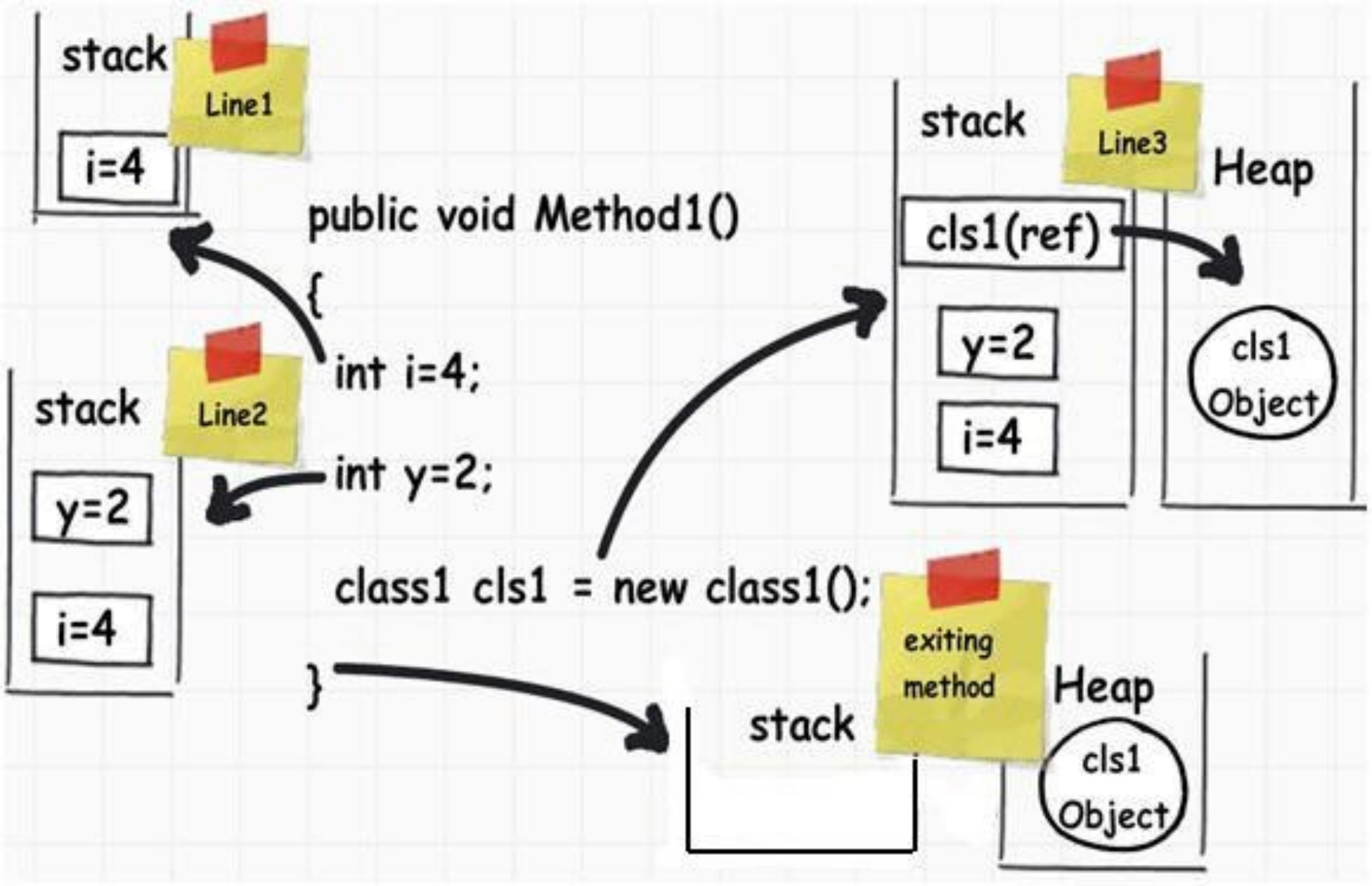
Pointer and Objects (custom defined data type)

Equation: 3 10 5 + *



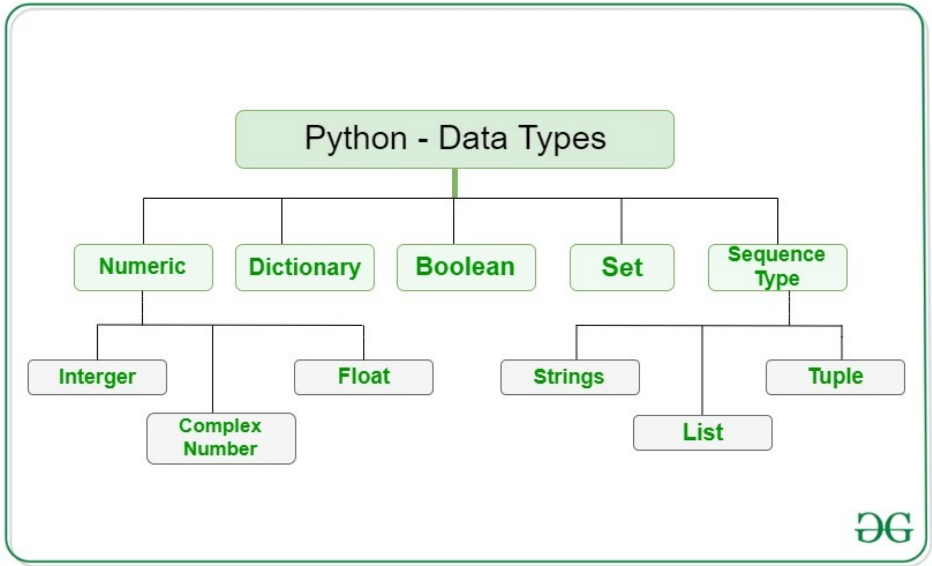


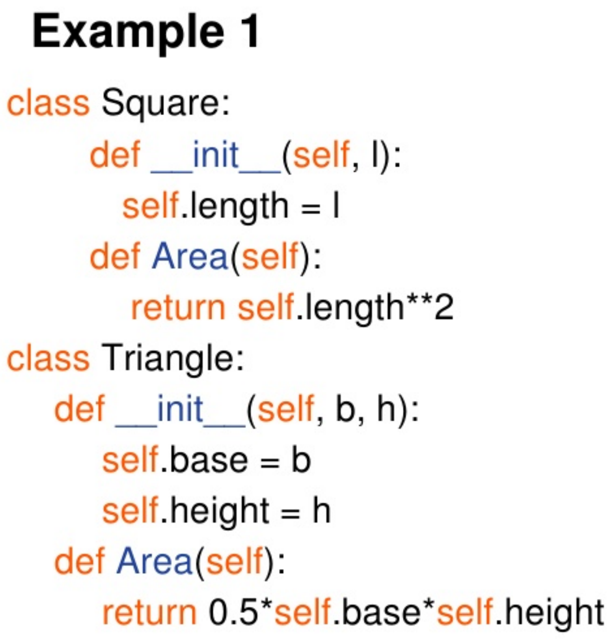
```
#include<stdio.h>
#include<stdlib.h>
int main()
{
  int a; // goes on stack
  int *p;
  p = (int*)malloc(sizeof(int));
  *p = 10;
  free(p);
 p = (int*)malloc(20*sizeof(int));
}
```



```
public void Method1()
{
int i=4;
int y=2;
class1 cls1 = new class1();
}
```

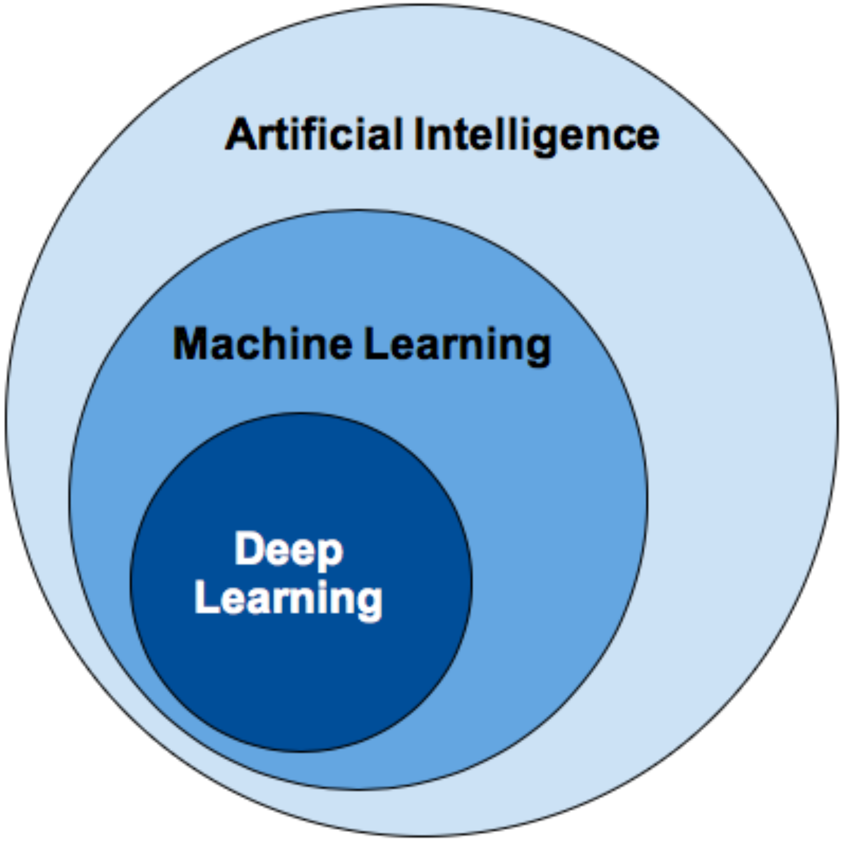
Computers did not appear suddenly; there is a rich long history.





Example 1

```
class Square:
        def __init__(self, l):
            self.length = l
        def Area(self):
            return self.length**2
class Triangle:
    def __init__(self, b, h):
        self.base = b
        self.height = h
    def Area(self):
        return 0.5*self.base*self.height
```

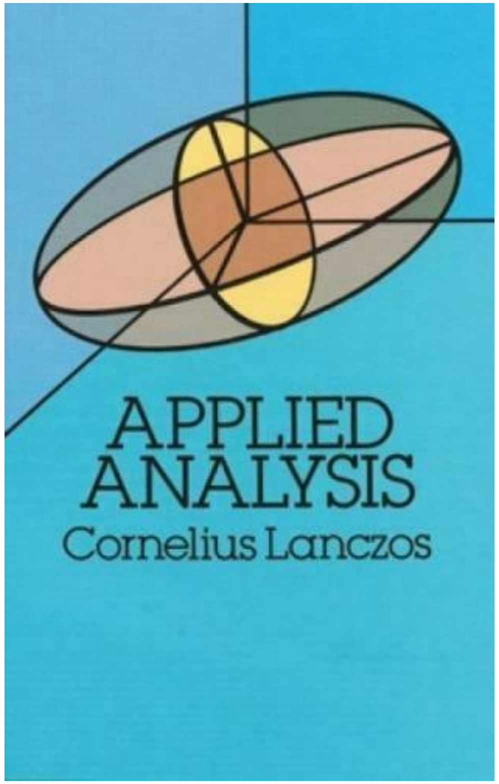


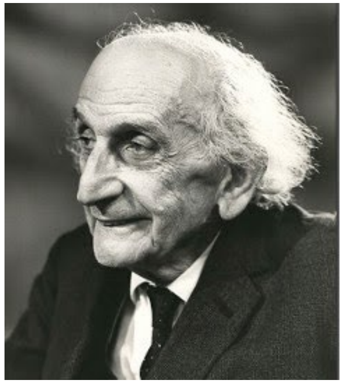
Enterprise Computing versus Scientific Computing

What about Inverse Theory?

- In Science, there is not enough Data.
- Well it works on images but how about other form of data?

If you understand Singular Value Decomposition(SVD), then you understand Invert Theory and all the linear algebra behind it.

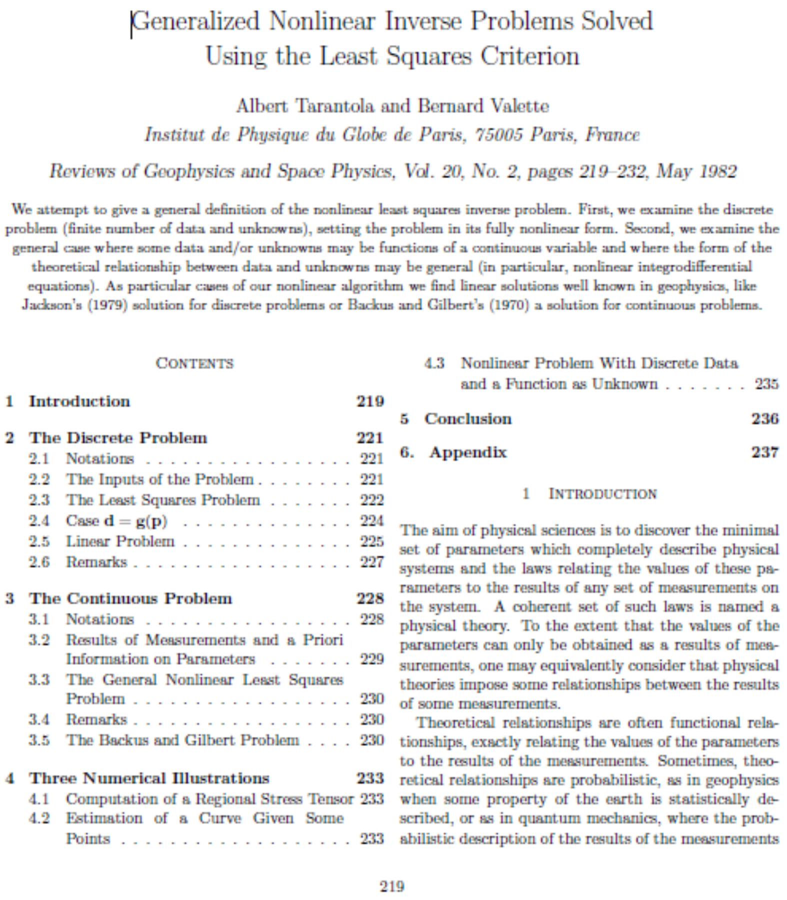




First published in 1956

Over determined problems
Under determined problems
Ill pose problems

How does one update once information



Generalized Nonlinear Inverse Problems Solved Using the Least Squares Criterion

Albert Tarantola and Bernard Valette

Institut de Physique du Globe de Paris, 75005 Paris, France

Reviews of Geophysics and Space Physics, Vol. 20, No. 2, pages 219–232, May 1982

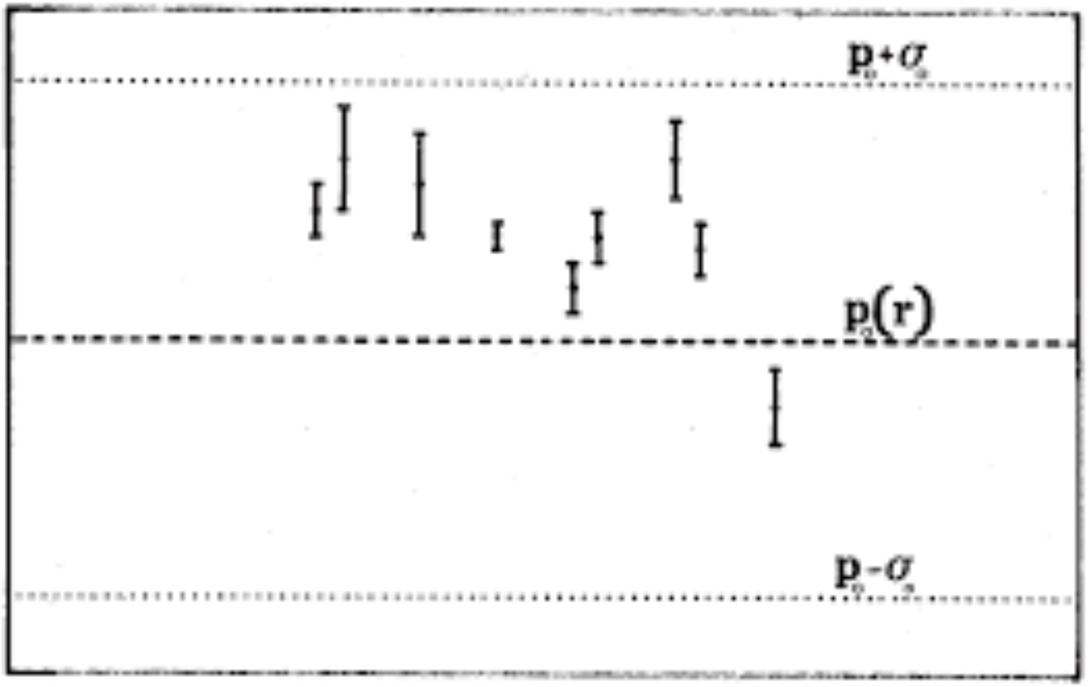
We attempt to give a general definition of the nonlinear least squares inverse problem. First, we examine the discrete problem (finite number of data and unknowns), setting the problem in its fully nonlinear form. Second, we examine the general case where some data and/or unknowns may be functions of a continuous variable and where the form of the theoretical relationship between data and unknowns may be general (in particular, nonlinear integrodifferential equations). As particular cases of our nonlinear algorithm we find linear solutions well known in geophysics, like Jackson's (1979) solution for discrete problems or Backus and Gilbert's (1970) a solution for continuous problems.

Contents

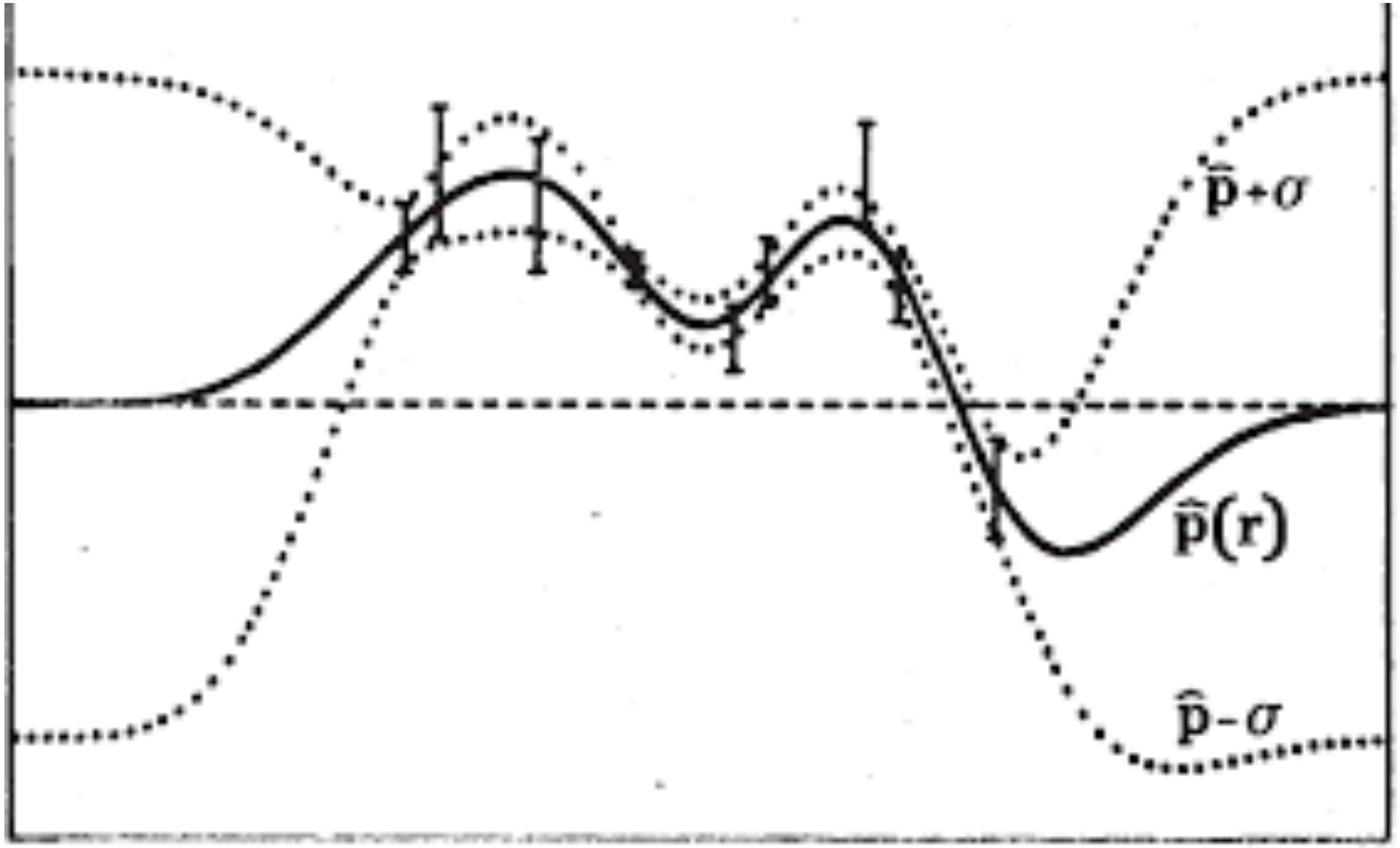
1 Introduction

The aim of physical sciences is to discover the minimal set of parameters which completely describe physical systems and the laws relating the values of these parameters to the results of any set of measurements on the system. A coherent set of such laws is named a physical theory. To the extent that the values of the parameters can only be obtained as a results of measurements, one may equivalently consider that physical theories impose some relationships between the results of some measurements.

Theoretical relationships are often functional relationships, exactly relating the values of the parameters to the results of the measurements. Sometimes, theoretical relationships are probabilistic, as in geophysics when some property of the earth is statistically described, or as in quantum mechanics, where the probabilistic description of the results of the measurements



The solid line will be the updated your information and everything has uncertainty.



“Computer Is God's Gift to People with Disability”

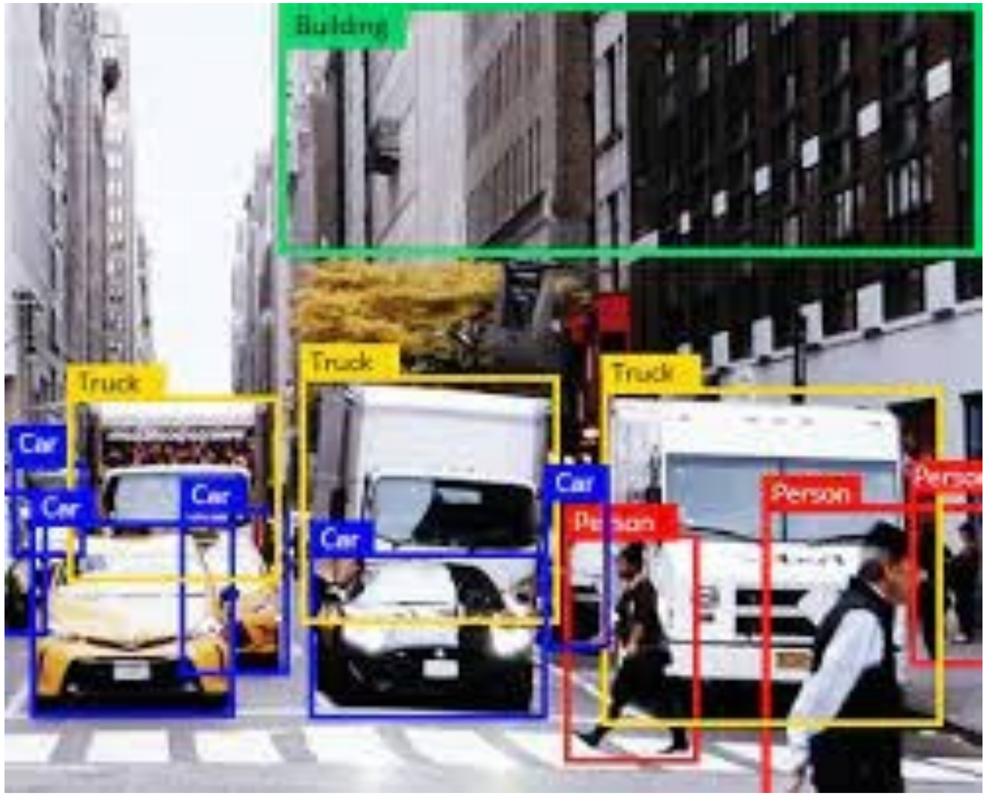
- The three difficulties: economic, social and family.

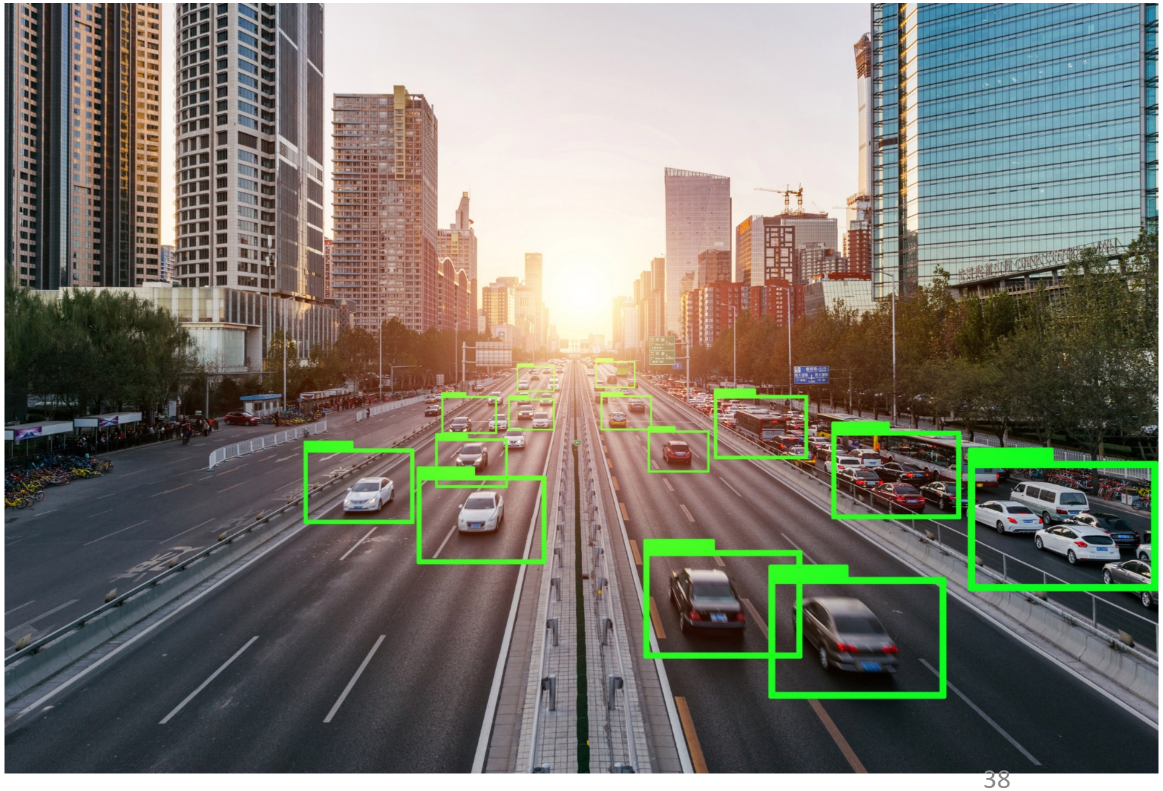
- By providing decent job, the three difficulties can be mitigated.
- But to do so, training and education is the most fundamental task.

- JEDI (justice, equality diversity and inclusion)

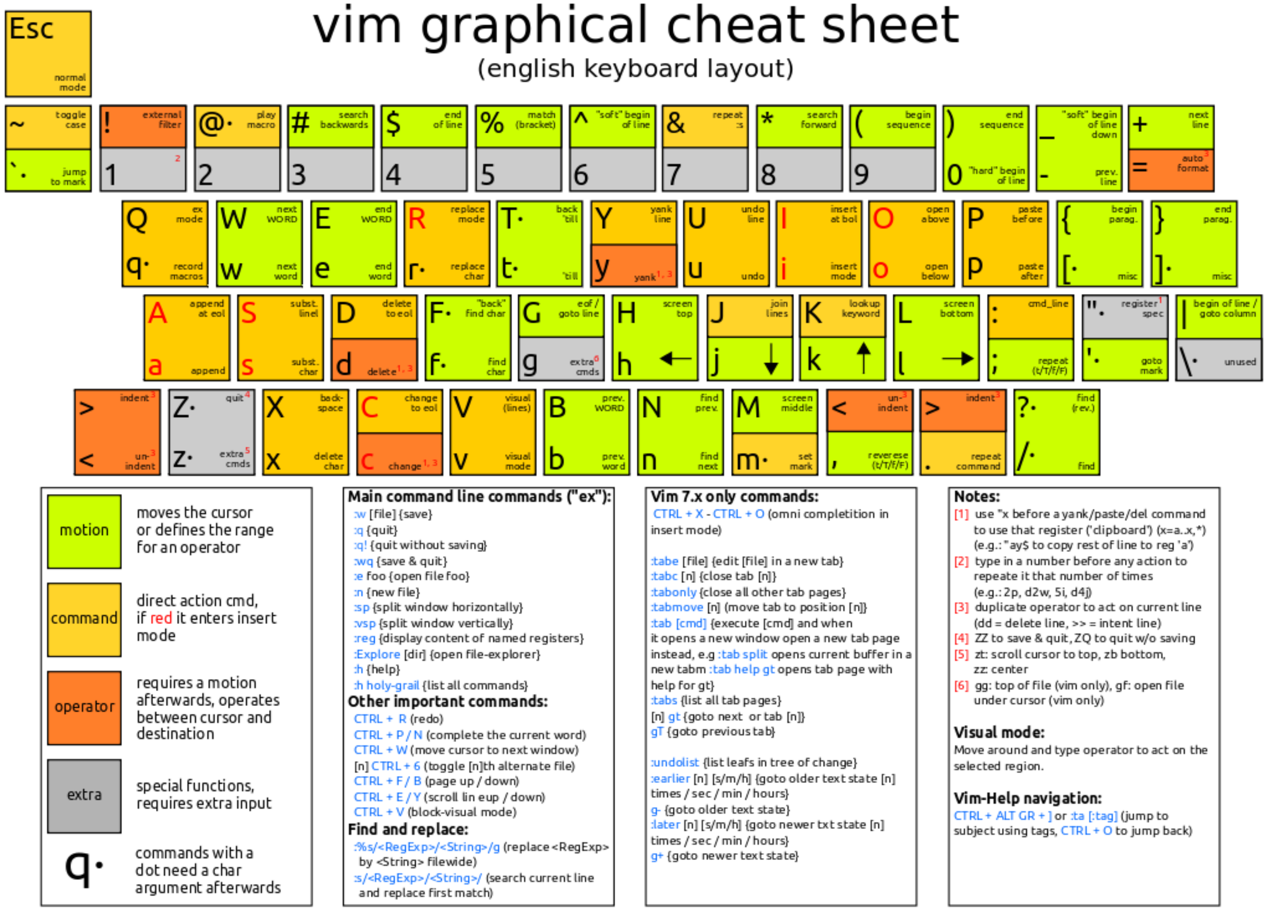
Menial and basic tasks. Obviously very low wages

Data Labeling in the era of AI. But still very boring and not inspiring.





Professional programmers don't use mouse. This was supposed to be more efficient for them but not so for people like me with difficult in hand mobility.



vim graphical cheat sheet

(english keyboard layout)

motion	moves the cursor or defines the range for an operator
command	direct action cmd, if red it enters insert mode
operator	requires a motion afterwards, operates between cursor and destination
extra	special functions, requires extra input
q·	commands with a dot need a char argument afterwards

Main command line commands ("ex"):

:w [file] {save}
:q {quit}
:q! {quit without saving}
:wq {save & quit}
:e foo {open file foo}
:n {new file}
:sp {split window horizontally}
:vsp {split window vertically}
:reg {display content of named registers}
:Explore [dir] {open file-explorer}
:h {help}
:h holy-grail {list all commands}

Other important commands:

CTRL + R (redo)
CTRL + P / N (complete the current word)
CTRL + W (move cursor to next window)
[n] CTRL + 6 (toggle [n]th alternate file)
CTRL + F / B (page up / down)
CTRL + E / Y (scroll lin eup / down)
CTRL + V (block-visual mode)

Find and replace:

:%s/<RegExp>/<String>/g (replace <RegExp> by <String> filewide)
:s/<RegExp>/<String>/ (search current line and replace first match)

Vim 7.x only commands:

CTRL + X - CTRL + O (omni completition in insert mode)

:tabe [file] {edit [file] in a new tab}
:tabc [n] {close tab [n]}
:tabonly {close all other tab pages}
:tabmove [n] (move tab to position [n]}
:tab [cmd] {execute [cmd] and when it opens a new window open a new tab page instead, e.g :tab split opens current buffer in a new tabm :tab help gt opens tab page with help for gt}
:tabs {list all tab pages}
[n] gt {goto next or tab [n]}
gT {goto previous tab}

:undolist {list leafs in tree of change}
:earlier [n] [s/m/h] {goto older text state [n] times / sec / min / hours}
g- {goto older text state}
:later [n] [s/m/h] {goto newer txt state [n] times / sec / min / hours}
g+ {goto newer text state}

Notes:

[1] use "x before a yank/paste/del command to use that register ('clipboard') (x=a..x,*) (e.g.: "ay$ to copy rest of line to reg 'a')
[2] type in a number before any action to repeate it that number of times (e.g.: 2p, d2w, 5i, d4j)
[3] duplicate operator to act on current line (dd = delete line, >> = intent line)
[4] ZZ to save & quit, ZQ to quit w/o saving
[5] zt: scroll cursor to top, zb bottom, zz: center
[6] gg: top of file (vim only), gf: open file under cursor (vim only)

Visual mode:

Move around and type operator to act on the selected region.

Vim-Help navigation:

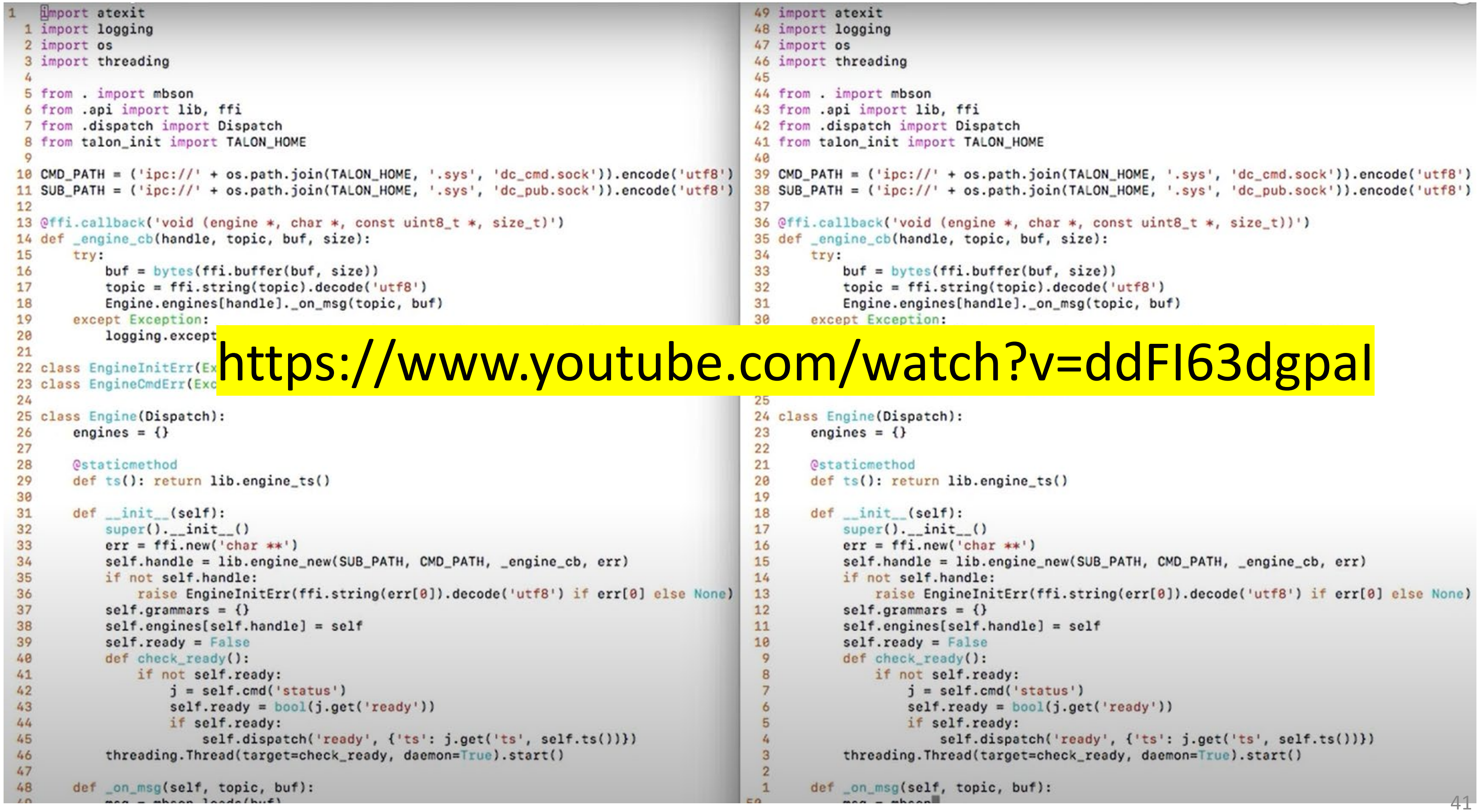
CTRL + ALT GR +] or :ta [:tag] (jump to subject using tags, CTRL + O to jump back)

Repetitive Strain Injury (RSI) is common among professional developers, and since 2013 following the presentation of Tavis Rudd, several solutions have been made to overcome this problem.

https://www.youtube.com/watch?v=8SkdfdXWYaI

https://www.youtube.com/watch?v=YRyYIIFKsdU

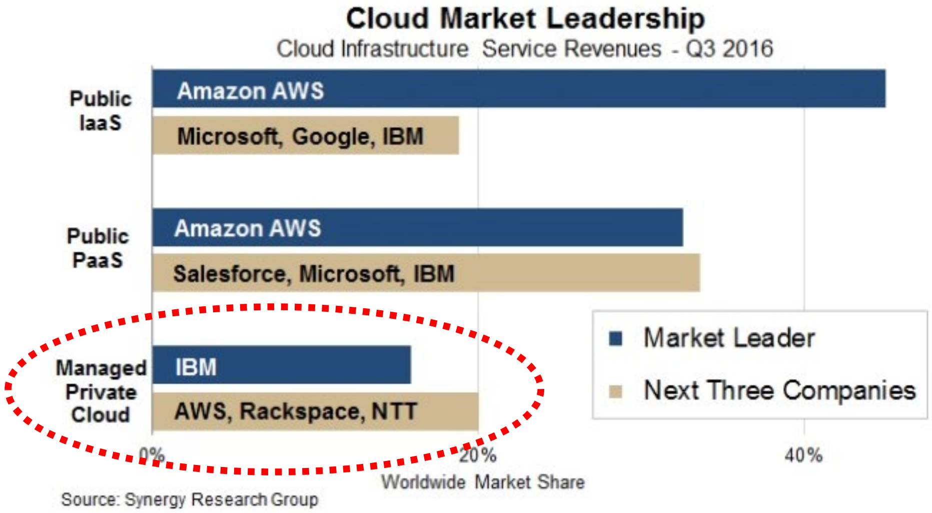
Comparison between hand typed and voice coding



```
import atexit
import logging
import os
import threading

from . import mbson
from .api import lib, ffi
from .dispatch import Dispatch
from talon_init import TALON_HOME

CMD_PATH = ('ipc://' + os.path.join(TALON_HOME, '.sys', 'dc_cmd.sock')).encode('utf8')
SUB_PATH = ('ipc://' + os.path.join(TALON_HOME, '.sys', 'dc_pub.sock')).encode('utf8')

@ffi.callback('void (engine *, char *, const uint8_t *, size_t)')
def _engine_cb(handle, topic, buf, size):
    try:
        buf = bytes(ffi.buffer(buf, size))
        topic = ffi.string(topic).decode('utf8')
        Engine.engines[handle]._on_msg(topic, buf)
    except Exception:
        logging.except

class EngineInitErr(Ex
class EngineCmdErr(Exc

class Engine(Dispatch):
    engines = {}

    @staticmethod
    def ts(): return lib.engine_ts()

    def __init__(self):
        super().__init__()
        err = ffi.new('char **')
        self.handle = lib.engine_new(SUB_PATH, CMD_PATH, _engine_cb, err)
        if not self.handle:
            raise EngineInitErr(ffi.string(err[0]).decode('utf8') if err[0] else None)
        self.grammars = {}
        self.engines[self.handle] = self
        self.ready = False
        def check_ready():
            if not self.ready:
                j = self.cmd('status')
                self.ready = bool(j.get('ready'))
                if self.ready:
                    self.dispatch('ready', {'ts': j.get('ts', self.ts())})
        threading.Thread(target=check_ready, daemon=True).start()

    def _on_msg(self, topic, buf):
        msg = mbson.loads(buf)
```

```
import atexit
import logging
import os
import threading

from . import mbson
from .api import lib, ffi
from .dispatch import Dispatch
from talon_init import TALON_HOME

CMD_PATH = ('ipc://' + os.path.join(TALON_HOME, '.sys', 'dc_cmd.sock')).encode('utf8')
SUB_PATH = ('ipc://' + os.path.join(TALON_HOME, '.sys', 'dc_pub.sock')).encode('utf8')

@ffi.callback('void (engine *, char *, const uint8_t *, size_t))')
def _engine_cb(handle, topic, buf, size):
    try:
        buf = bytes(ffi.buffer(buf, size))
        topic = ffi.string(topic).decode('utf8')
        Engine.engines[handle]._on_msg(topic, buf)
    except Exception:


class Engine(Dispatch):
    engines = {}

    @staticmethod
    def ts(): return lib.engine_ts()

    def __init__(self):
        super().__init__()
        err = ffi.new('char **')
        self.handle = lib.engine_new(SUB_PATH, CMD_PATH, _engine_cb, err)
        if not self.handle:
            raise EngineInitErr(ffi.string(err[0]).decode('utf8') if err[0] else None)
        self.grammars = {}
        self.engines[self.handle] = self
        self.ready = False
        def check_ready():
            if not self.ready:
                j = self.cmd('status')
                self.ready = bool(j.get('ready'))
                if self.ready:
                    self.dispatch('ready', {'ts': j.get('ts', self.ts())})
        threading.Thread(target=check_ready, daemon=True).start()

    def _on_msg(self, topic, buf):
        msg = mbson
```

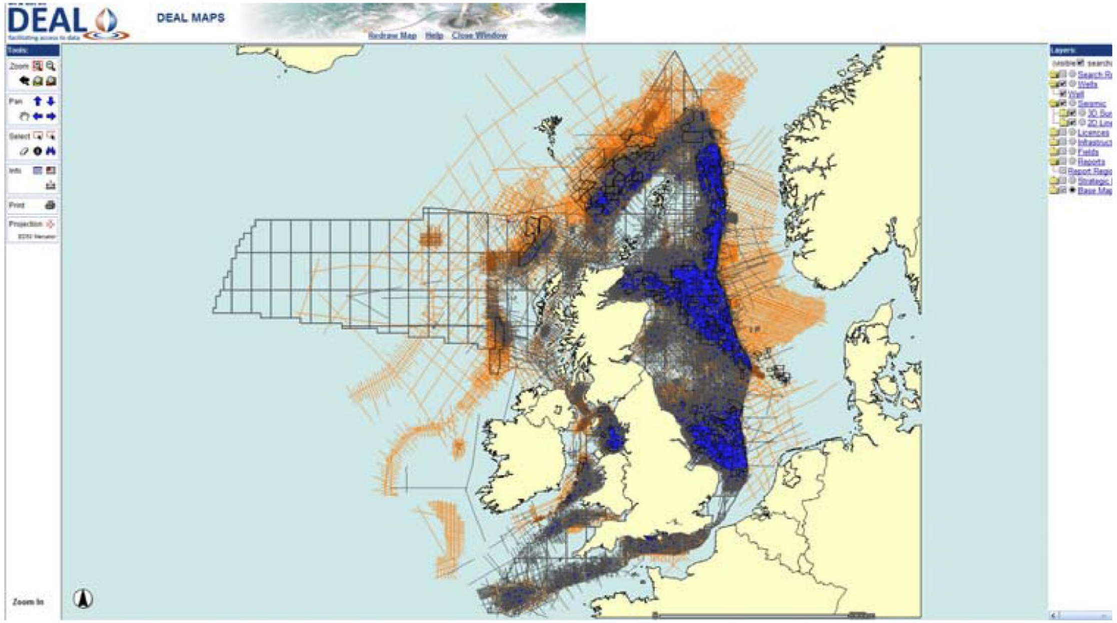
https://www.youtube.com/watch?v=ddFI63dgpaI



Cloud Market Leadership

Cloud Infrastructure Service Revenues - Q3 2016

Source: Synergy Research Group



Cloud computing and remote working have become a new norm as a result of coronavirus. People with a high level of disability find their home most comforting.

Here is an example of public data. Shown here are national database of UK as part of offshore petroleum survey. My goal is to set up an organization such that people with disability who has high-level of computer knowledge and skills would participate in this global issue of environment and energy.

At the moment, I am testing this idea with nine people with disability.

Thank You Very Much for Your Attention