Introduction to Research
(in Data Management)

Amr Magdy
Assistant Professor
Computer Science and Engineering
www.cs.ucr.edu/~amr/
What is research about?

› We solve unsolved problems
  › Unsolved not unsolvable

› Examples:
  › How to fly faster?
  › How to cure cancer?
  › How to enable a machine to see?
  › How to understand all data generated by human?

› Research is about finding solutions for unsolved problems in different life aspects
  › Health, engineering, behavioral sciences…etc
Is research that big?

- It is,
  - but many MANY teams contribute to that big problems

- Example: how to enable a machine to see?
  Subproblems:
  - How a human see?
  - How to make an artificial eye?
    - How to take photos? (Camera) - How to make it better? …..etc
  - How to make an artificial vision brain?
    - How to segment photos? - How to know objects? …..etc
  - How to connect eye to brain?
  - How to ……………
Is research that big?

- It is a big puzzle
Is research that big?

- It is a big puzzle
Is research that big?

- It is a big puzzle
Research problem elements

To address a research problem, answer six questions:

1. What is the problem?
2. Why it is important?
3. Why it is challenging?
4. What are the limitations of the related work?
5. What are the novel contributions?
6. What is the validation method?
Introduction to Data Management Research
The digital world

Data sources:

- Business: retailers, enterprise, etc
- Humans: games, social media, crowdsourcing, etc
- Science: astronomy, biology, satellites, etc
- ....

199x: Internet services provided to public
200x: Internet comes to mobile devices (revolution)
Now: ?
The digital world

Data sources:
Business: retailers, enterprise, etc
Humans: games, social media, crowdsourcing, etc
Science: astronomy, biology, satellites, etc

..., 199x: Internet services provided to public
200x: Internet comes to mobile devices (revolution)
Now: ?

Startups Using Big Data

Cybersecurity
Marketing
Real Estate
Energy
Agriculture
Finance
Healthcare
Supply Chain
Retail

CB INSIGHTS

CounterTack
CYLANCE
custora
[ BECKON ]
RISKIQ
RAPID7
6sense
Localytics
REONYM
homesnap
Agent Ace
COMPSTAK
FarmLogs
Farmeron
vts
smartzip
Granular
OnFARM
comprehend
FLATIRON
tonic
Healthcare
Ginger.io

HealthCatalyst
SolveBio
commerce
BLUE RIDGE
eyeQ
LogiNext
Mu Sigma
fusionops
stitchlabs

VoloMetrix
objective logistics
VISIER
GLINT1
HR/Recruiting

Visata
Credit Karma
OpenGamma
xignite
ZestFinance
OrochestrO
ORCHestRO

FirstInsight
Prayas Analytics

Data is the new Oil

The world’s most valuable resource is no longer oil, but data

The data economy demands a new approach to antitrust rules

BRAINSTORM TECH

FORTUNE

Why Data Is The New Oil

Jonathan Vanian
Jul 11, 2016

Tech

Forbes

Is Data The New Oil?
Oil Refinery vs. Data Refinery

“Data is just like crude. It’s valuable, but if unrefined it cannot really be used.”

Michael Palmer
“Data is just like crude. It’s valuable, but if unrefined it cannot really be used.”

Michael Palmer
Stages and Activities

- **Stages:**
  - Identifying a problem
  - Proposing a solution
  - Realizing and validating the solution
  - Documenting the research (writing papers)

- **Activities:**
  1. **Reading:** existing work, prior knowledge,…etc
  2. **Coding:** your solution, related work, utilities,…etc
  3. **Writing:** papers, summaries,…etc
  4. **Presenting:** your problem, solution, surveys, etc