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?
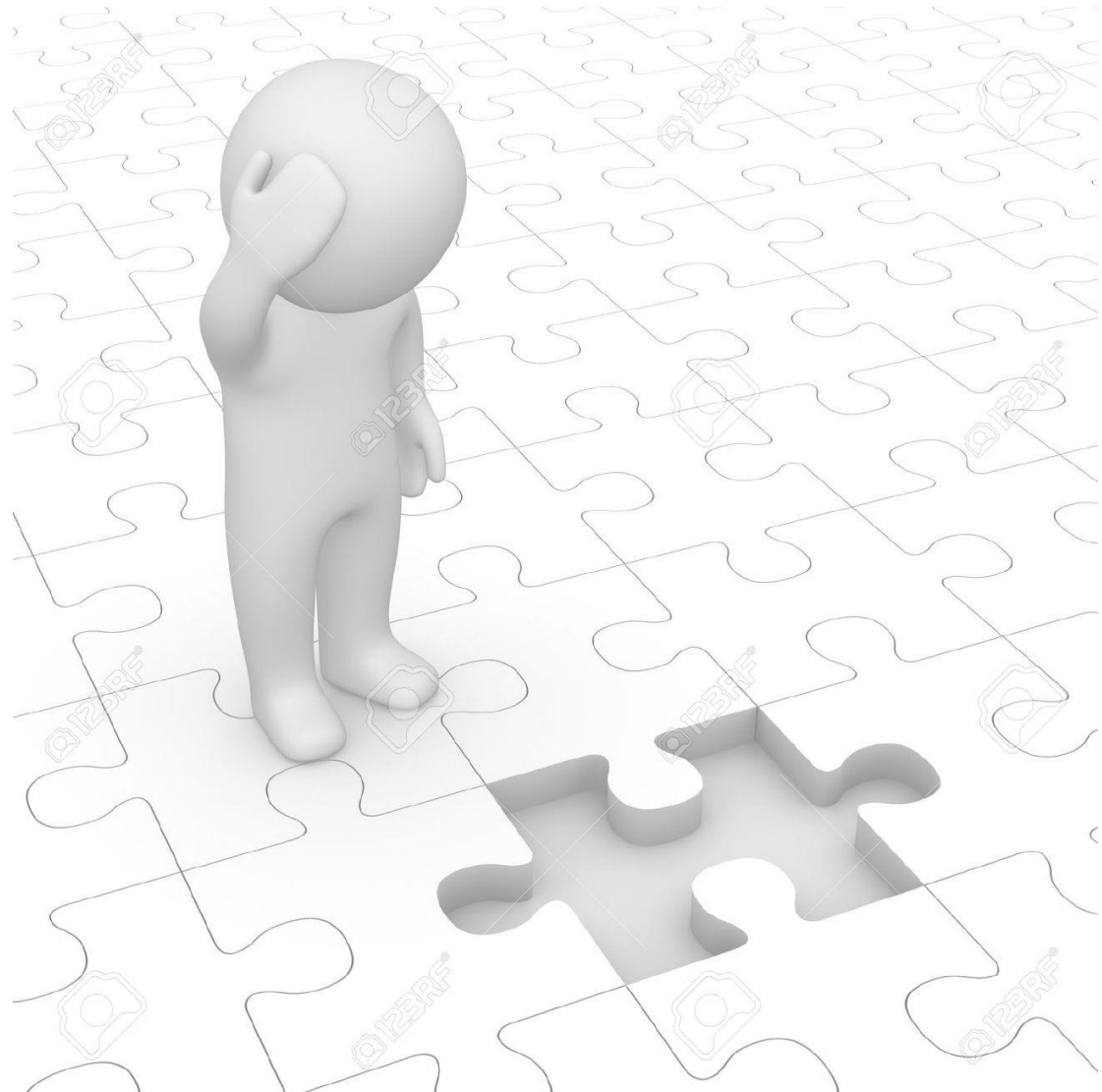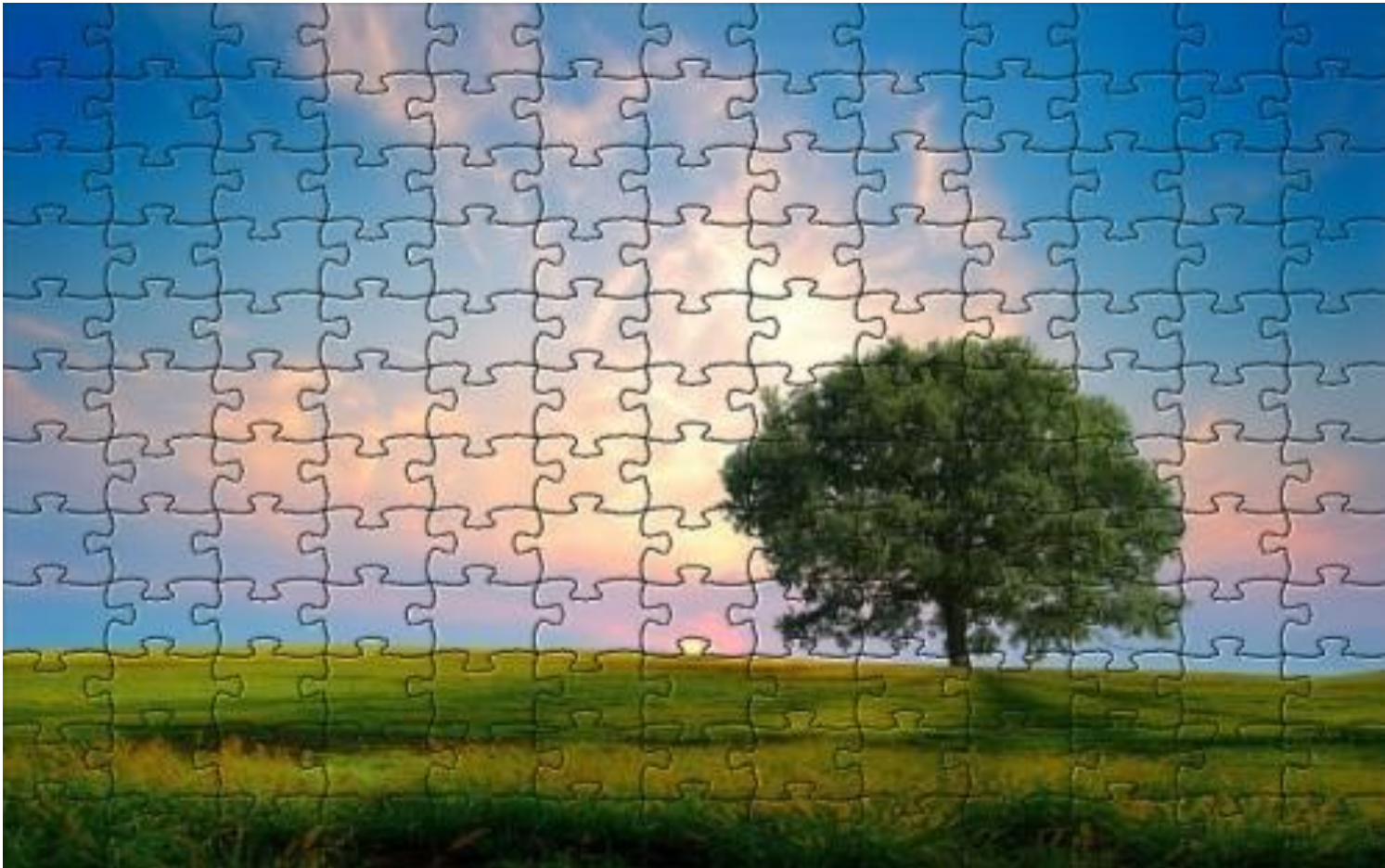
> It is a big puzzle

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?

Research problem elements

To address a research problem, answer six questions:

1. What is the problem?

2. Why is it important?

3. Why is it challenging?

4. What are the limitations of the related work?

5. What are the novel contributions?

6. What is the validation method?

Research problem elements

To address a research problem, answer six questions:

1. What is the problem?

Computational problem definitions:

      - What is the input?

      - What is the output?

      - What are the objectives/constraints?

      - Any other contextual elements for problem definition

Research problem elements

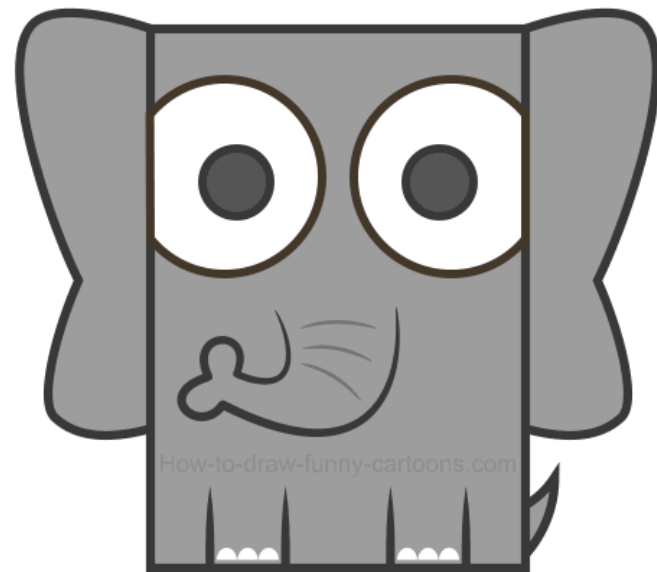To address a research problem, answer six questions:

1. What is the problem?

2. Why is it important?

3. Why is it challenging?

4. What are the limitations of the related work?

5. What are the novel contributions?

6. What is the validation method?

Research problem elements

To address a research problem, answer six questions:

2. Why is it important?

Getting a square elephant into a circle door is a problem, but not an important one

Research problem elements

To address a research problem, answer six questions:

1. What is the problem?

2. Why is it important?

3. Why is it challenging?

4. What are the limitations of the related work?

5. What are the novel contributions?

6. What is the validation method?

Research problem elements

To address a research problem, answer six questions:

2. Why is it challenging?

      calculating x*y is an important problem to solve, but it is not difficult, and it can be handled by hardware in O(1)

Research problem elements

To address a research problem, answer six questions:

1. What is the problem?

2. Why is it important?

3. Why is it challenging?

4. What are the limitations of the related work?

5. What are the novel contributions?

6. What is the validation method?

Research problem elements

To address a research problem, answer six questions:

4. What are the limitations of the related work?

Given: a point location, 100 millions spatial objects

Find: the nearest neighbor

Important and challenging! but solved before, why existing solutions do not work.

Research problem elements
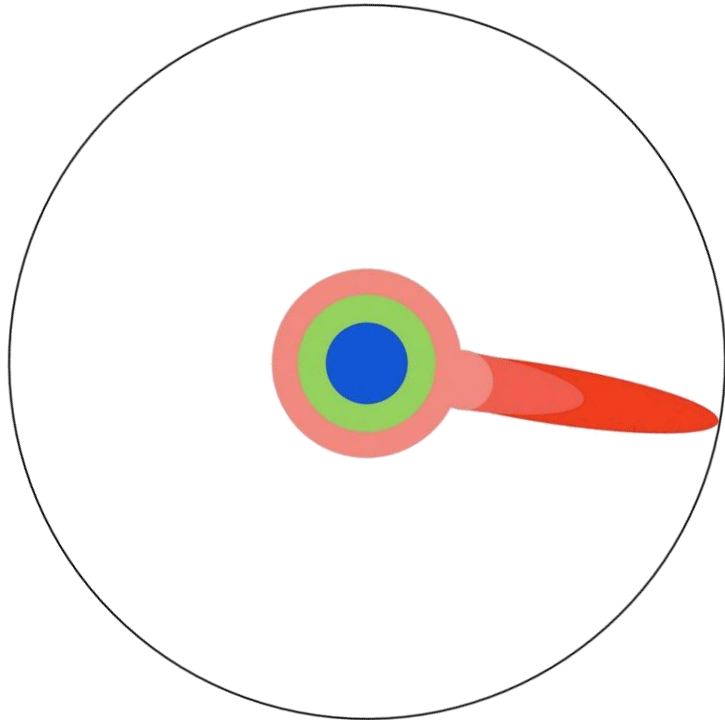
To address a research problem, answer six questions:

1. What is the problem?

2. Why is it important?

3. Why is it challenging?

4. What are the limitations of the related work?

5. What are the novel contributions?

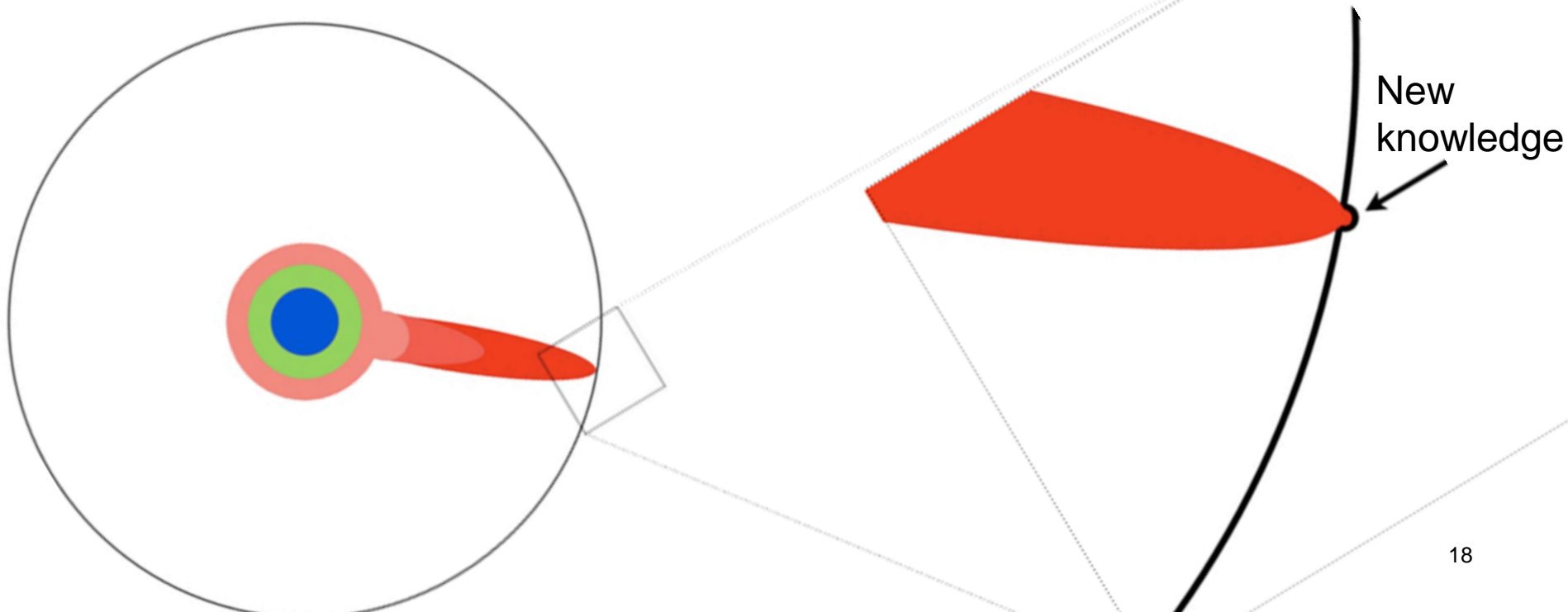6. What is the validation method?

Research problem elements

To address a research problem, answer six questions:

5. What are the novel contributions?

Research is about novel contributions to advance existing knowledge.

Research problem elements

To address a research problem, answer six questions:

5. What are the novel contributions?

Research is about novel contributions to advance existing knowledge.

New knowledge

Research problem elements

To address a research problem, answer six questions:

1. What is the problem?

2. Why is it important?

3. Why is it challenging?

4. What are the limitations of the related work?

5. What are the novel contributions?

6. What is the validation method?

Research problem elements

To address a research problem, answer six questions:

6. What is the validation method?
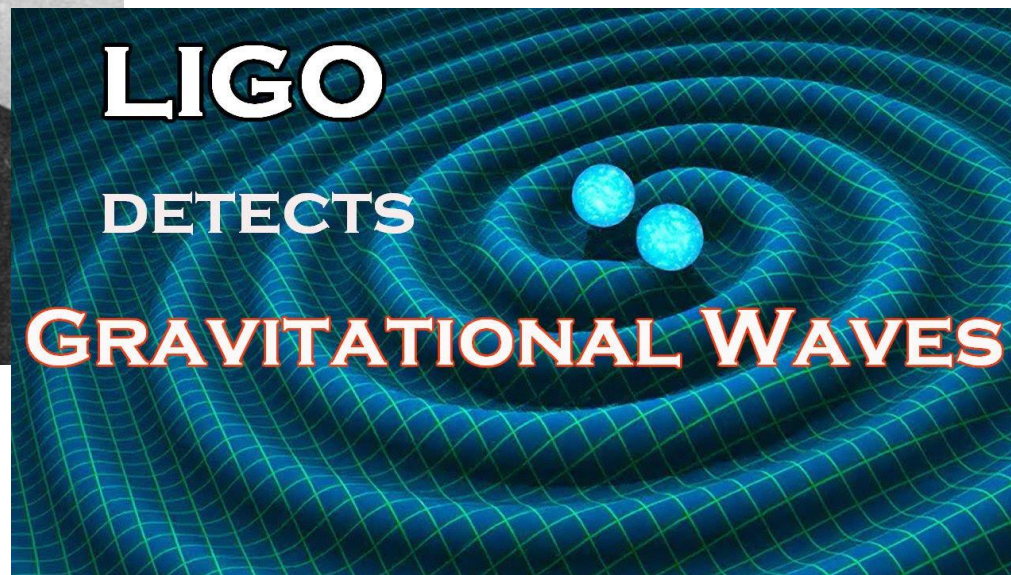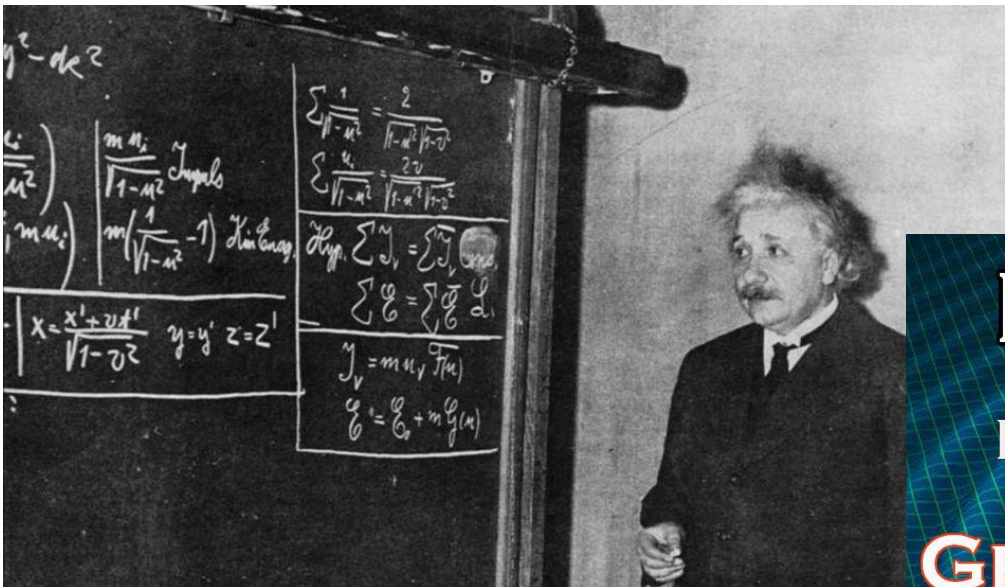
Introduction to Data Management Research

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: ?

Startups Using Big Data

Data is the new Oil

The Economist

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

The data economy demands a new approach to antitrust rules

FORTUNE

BRAINSTORM TECH

Why Data Is The New Oil

Jonathan Vanian
Jul 11, 2016

Forbes

Tech

APR 2, 2012 @ 11:09 AM
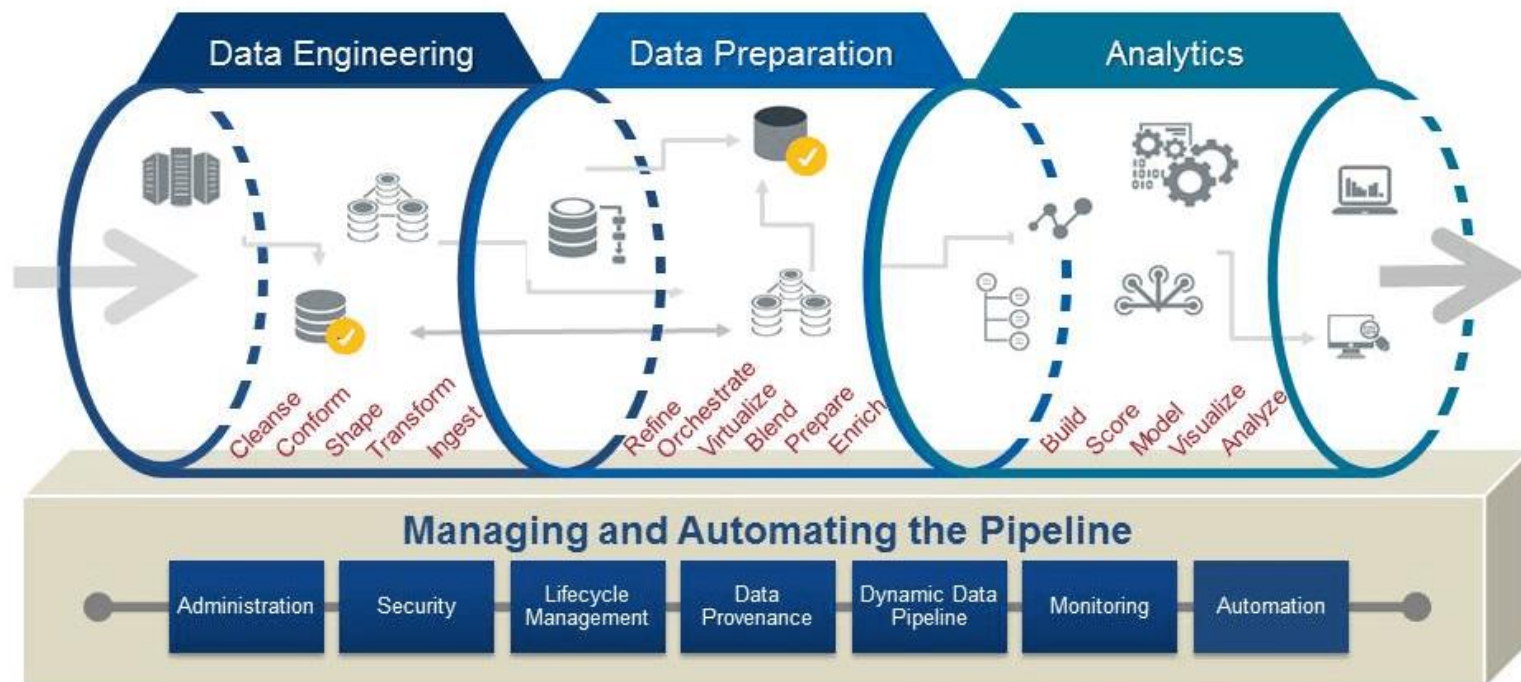
Is Data The New Oil?

Oil Refinery vs. Data Refinery

> "Data is just like crude. It's valuable, but if unrefined it cannot really be used."
>
> **Michael Palmer**

Oil Refinery vs. Data Refinery

> **"Data is just like crude. It's valuable, but if unrefined it cannot really be used."**
>
> **Michael Palmer**

Oil Refinery vs. Data Refinery

UCR

Forbes

Here's Why Data Is Not The New Oil

Bernard Marr Contributor ⓘ
Enterprise & Cloud

WIRED

BUSINESS   CULTURE   GEAR   IDEAS   SCIENCE   SECURITY   TRANSPORTATION

ANTONIO GARCÍA MARTÍNEZ   IDEAS   02.26.2019 07:00 AM

No, Data Is Not the New Oil

Proposals to "pay" users for the value of their data don't reflect how internet giants like Facebook and Google really operate.

Towards Data Science   DATA SCIENCE   MACHINE LEARNING   PROGRAMMING   VISUALIZATION   AI   JOURNALISM

Data is not the new oil

About the reality of working with data
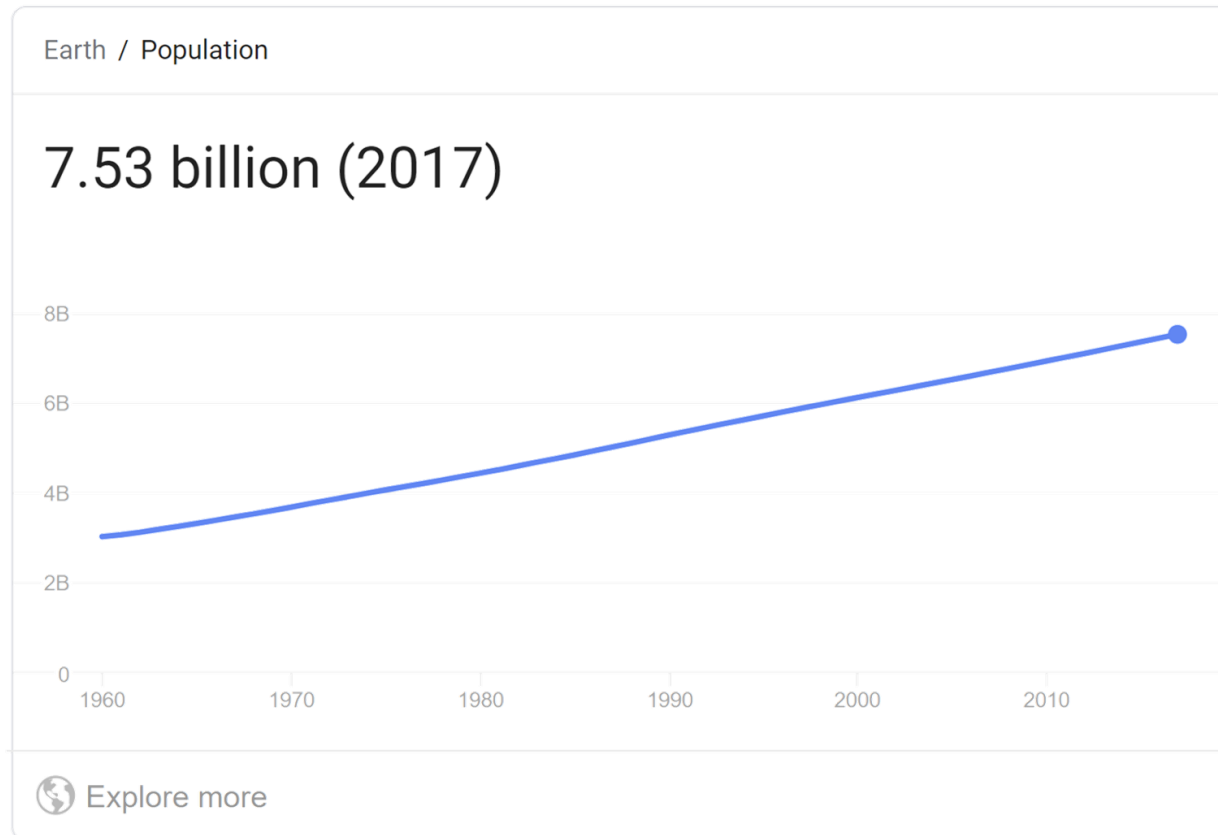
Samuel Flender   Follow

Oil Refinery vs. Data Refinery

> Why data might not be the new oil?
>> Oil refinery is much more deterministic (easier?)
>> Oil is finite, data is not

Oil Refinery vs. Data Refinery

> Why data might not be the new oil?
>> Oil refinery is much more deterministic (easier?)
>> Oil is finite, data is not

Earth / Population

7.53 billion (2017)

Explore more

Sources include: World Bank

Feedback

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