From Leibniz to Boole to Big Data

John Hooker
Carnegie Mellon University

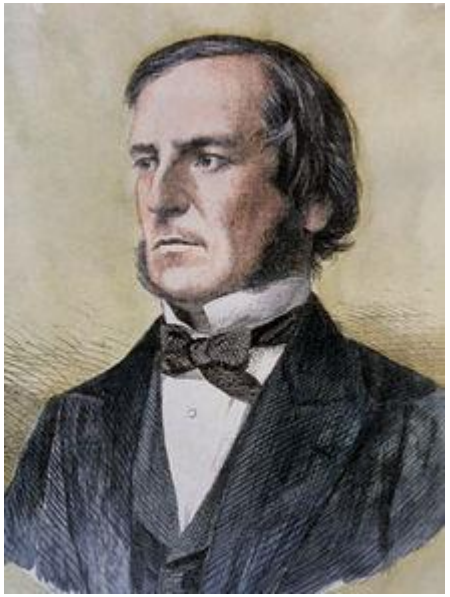
George Boole Bicentenary
University College Cork
August 2015

Far Ahead of His Time

George Boole is best known for Boolean logic.

But it has 200-year old roots in the ideas of **Leibniz**.

Innovative as it was, it was almost **traditional**, compared to Boole's **strikingly original** work in another type of logic…



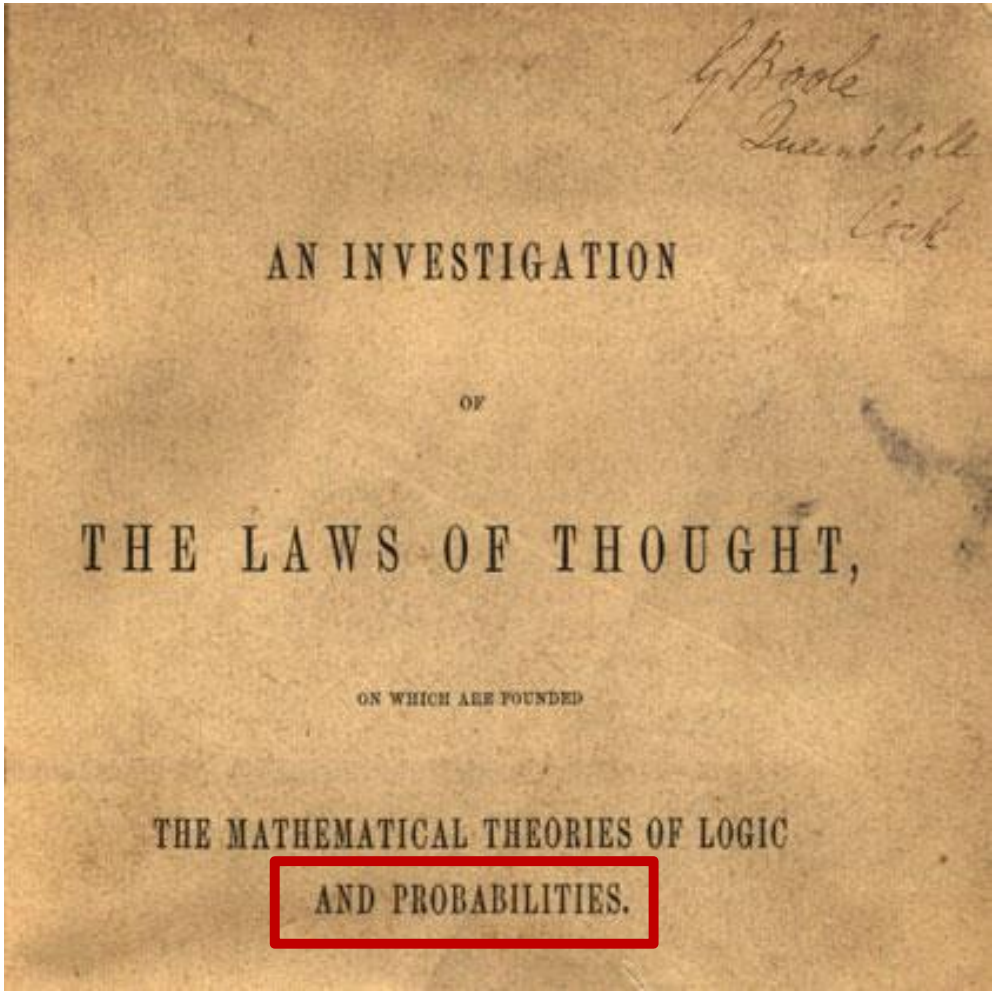
Far Ahead of His Time

…**probability logic**,

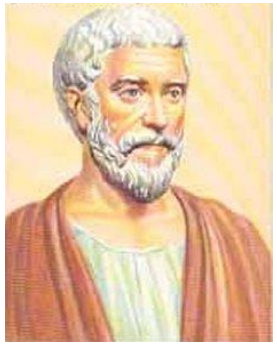
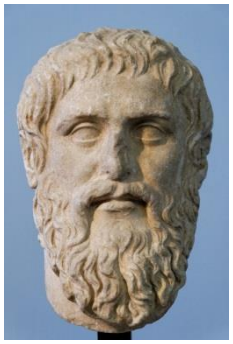
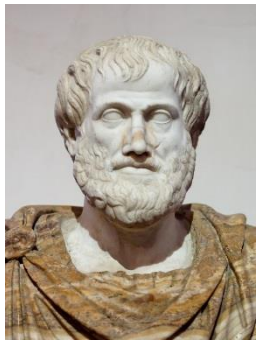
which was ignored or dismissed for a century.

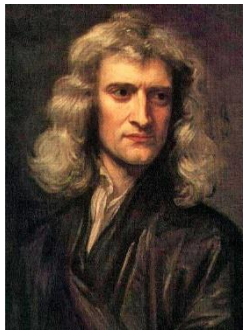
Boole was the first to address the central problem of our **information age**…

how to derive conclusions from "**big data**."



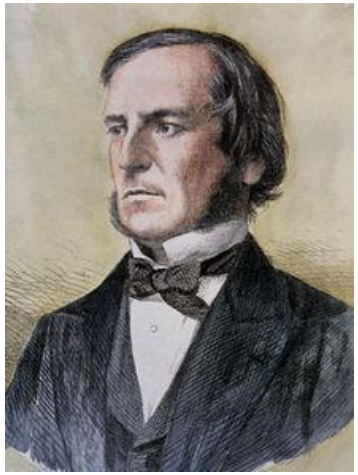
Far Ahead of His Time

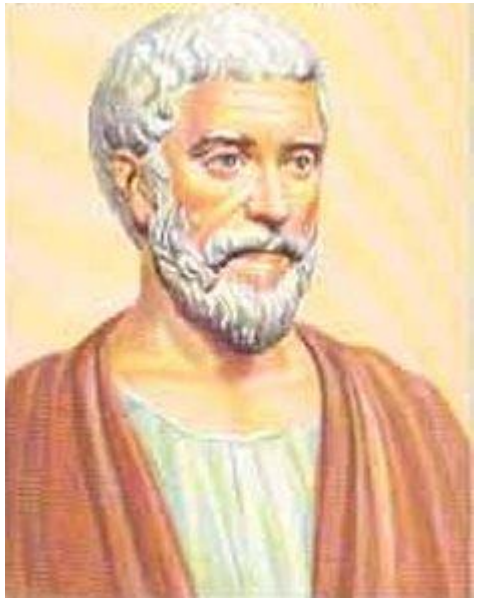


Boole represents a transition from 20 centuries of **rationalism** to our age of **data mining** and **crowd sourcing**.

To see this requires a brief excursion through Western cultural and intellectual history…

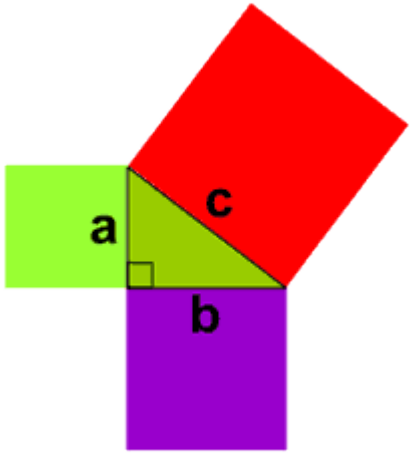


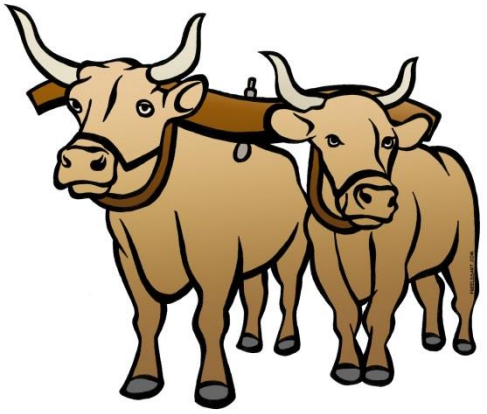
Back to the Beginning



Theorem of Pythagoras.
First proof in the Western tradition.

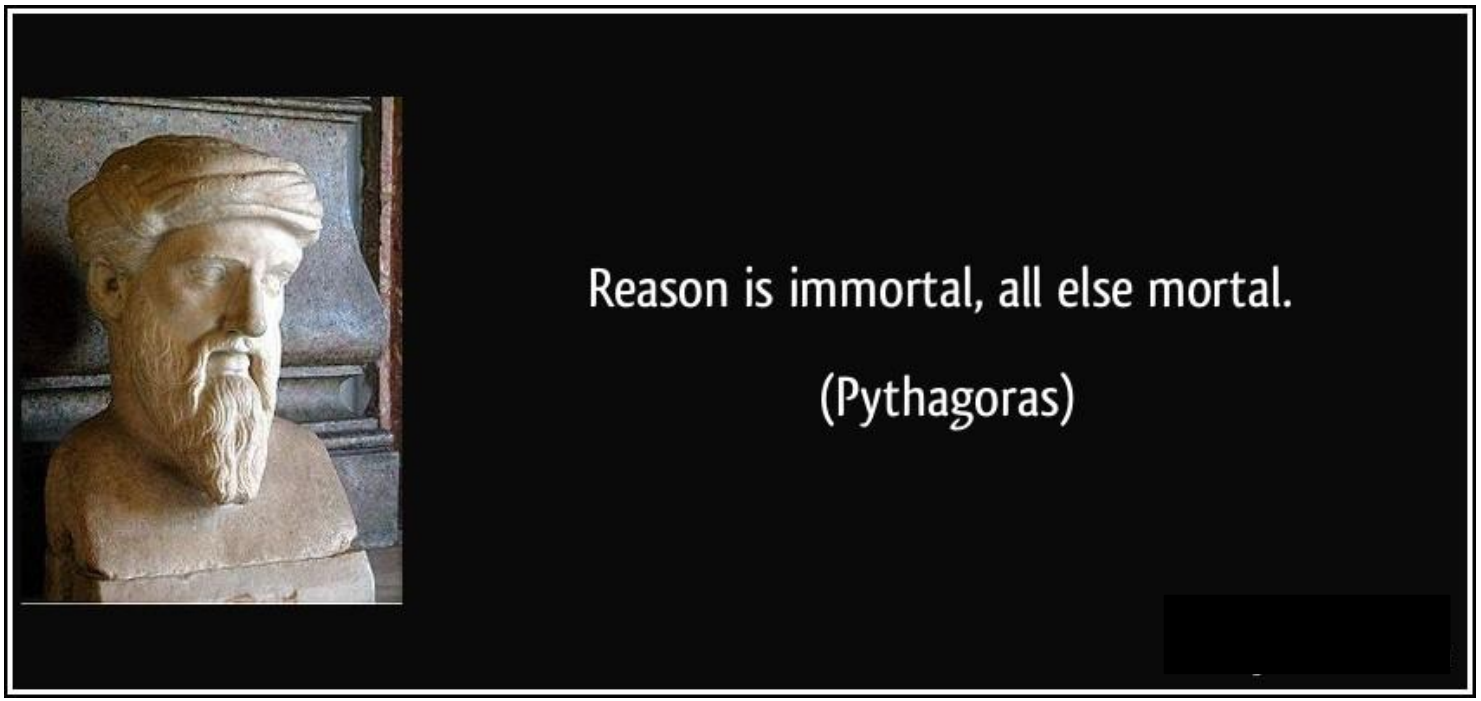
According to Legend, he sacrificed oxen to the gods on proving the theorem.





Why did he bother?

Back to the Beginning

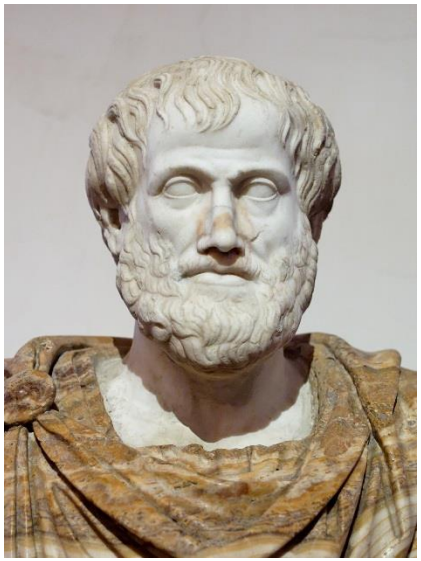


It proved the immortality of *psyche.*

In the *Meno,* Socrates argued further for transmigration of *psyche.*

Similar tradition in Hindu (Vedic) thought.

Back to the Beginning



Aristotle's logic remained the standard for 2000 years. Even Boole dealt with it in his first book.

Barbara syllogism:

All men are mortal
Socrates is a man
Socrates is mortal

256 types of syllogisms, including:

1st figure	2nd figure	3rd figure	4th figure
Barbara AAA-1	Cesare EAE-2	Darapti AAI-3	Bramantip AAI-4
Celarent EAE-1	Camestres AEE-2	Disamis IAI-3	Camenes AEE-4
Darii ALL-1	Festino EIO-2	Datisi ALL-3	Dimaris IAI-4
Ferio EIO-1	Baroco AOO-2	Felapton EAO-3	Fesapo EAO-4
		Bocardo OAO-3	Fresison EIO-4
		Ferison EIO-3	

Reliance on Reason

Why the emphasis on logic and reason?

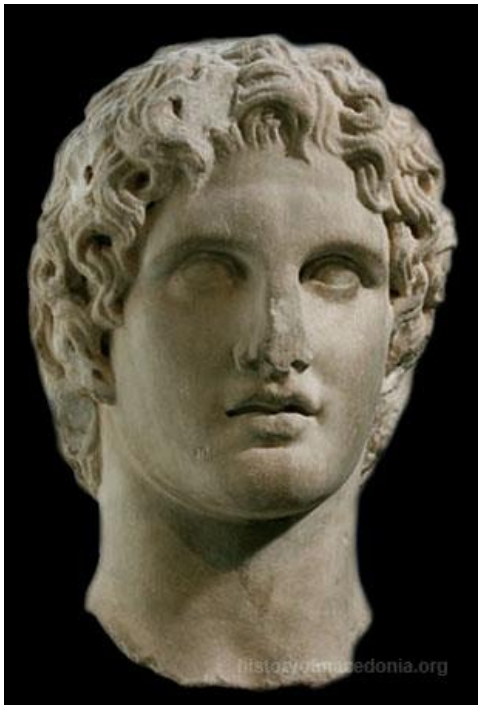
It is the primary **coping mechanism** in Western civilization.

We **relieve the stress** of uncertainty by finding order in the world.

Religion = reconnect

Rational worldview is compatible with **individualism**.

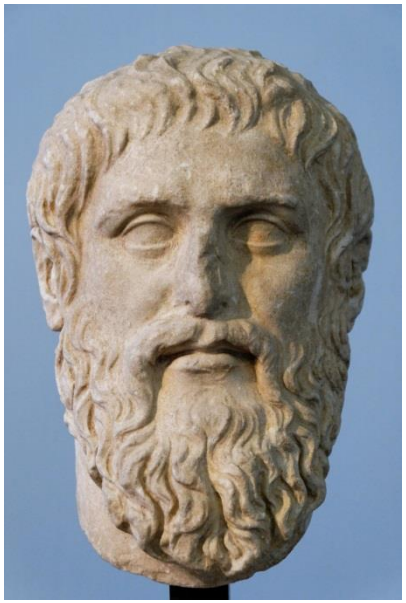
An intelligible world is **predictable**, **controllable** through technology.

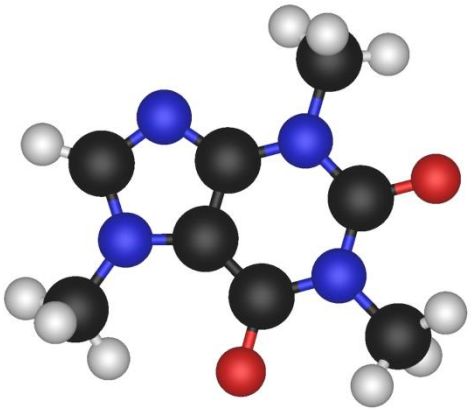


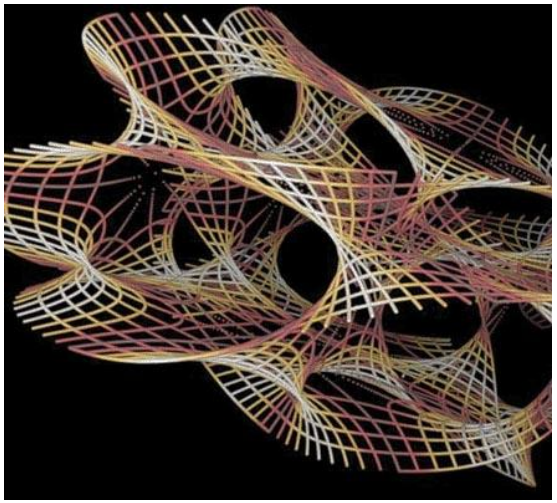
Alexander the Great

Reliance on Reason

Distrust of senses extends from Platonic forms to modern science







Mathematics and computer science today: Only proofs are publishable, no empirical observations (occasional exceptions).

Age of Reason

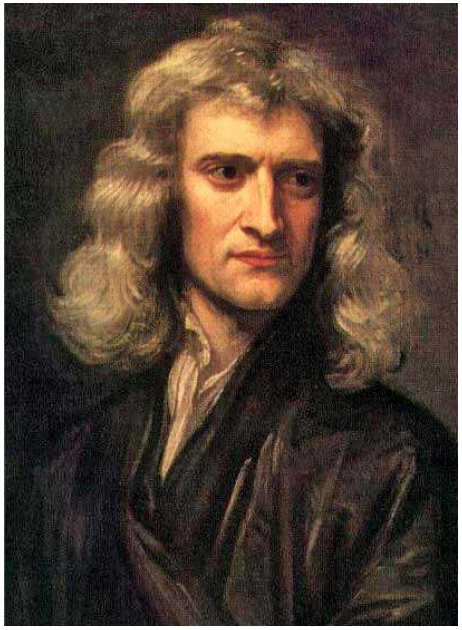
17th century: Zenith of rationalism
Reason can, in principle, deduce all truths.

René Descartes

Our senses deceive us

Baruch Spinoza

Theorem/proof format



Isaac Newton

Empirical observation is 2nd choice

Age of Reason

Gottfried Wilhelm
Freiherr von Leibniz

Chief rationalist

Conceived a logical calculus
(*calculus ratiocanator*)
that can deduce all truths…

….within a universal language
(*characteristica universalis*)

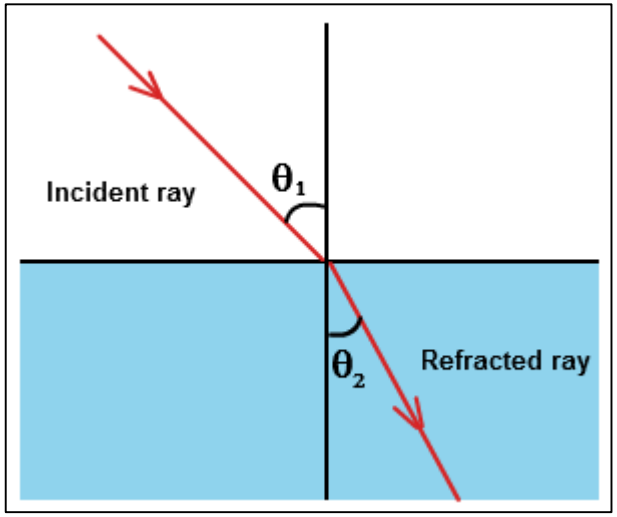
Age of Reason

Gottfried Wilhelm
Freiherr von Leibniz

Chief rationalist

Conceived a logical calculus
(*calculus ratiocanator*)
that can deduce all truths…

….within a universal language
(*characteristica universalis*)



Snell's Law

cf. Laplace

In the Leibnizian Tradition

Johann Heinreich Lambert, 18th c.

Used arithmetical notation for logic

From *Neues Organon:*


Inhalt des erſten Bandes.

Dianoiologie

oder

Lehre von den Geſetzen des Denkens.

"Laws of Thought"

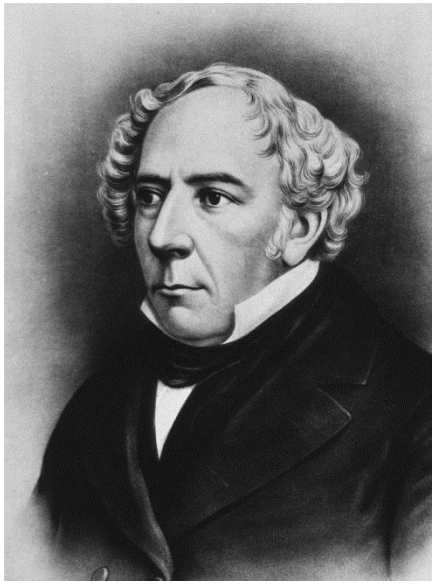
$$\alpha = \alpha\gamma + \alpha\delta = \alpha(\gamma + \delta)$$

This is "intensional" logic

In the Leibnizian Tradition

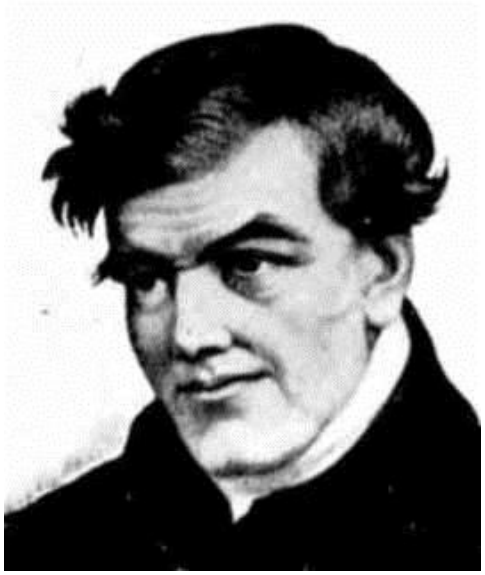
Augustus
De Morgan, 19th c.

Important influence
on Boole



William Hamilton

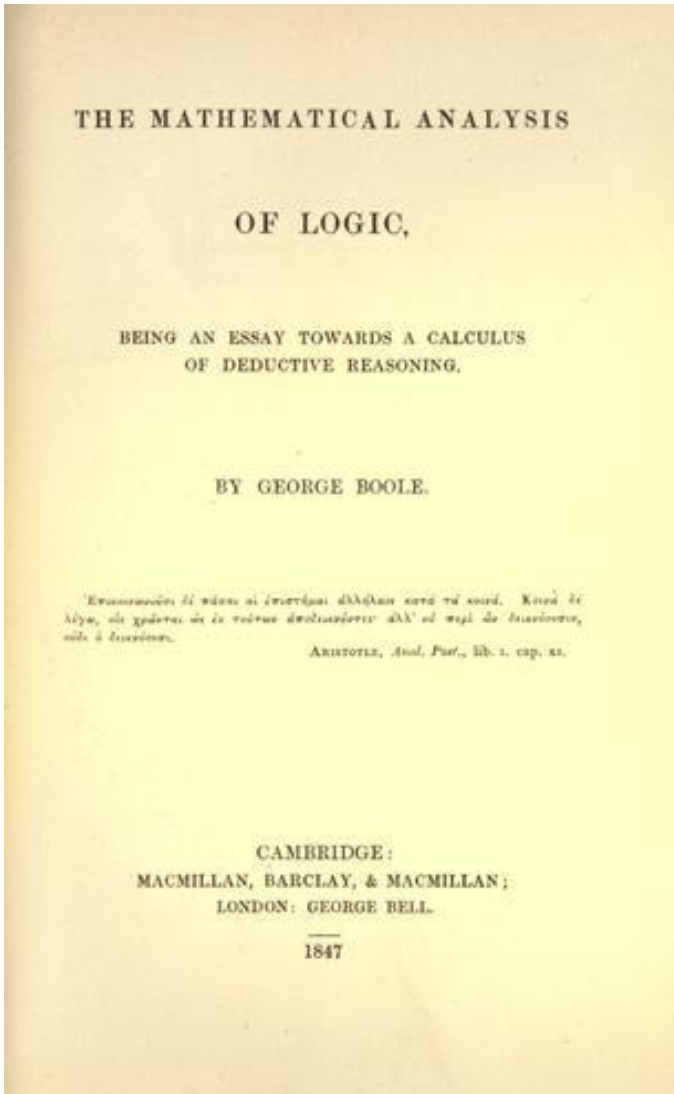
Dispute with De
Morgan inspired
Boole's first book



George Peacock,
19th c.

Emphasized
uninterpreted
symbols

Boole's First Book



THE MATHEMATICAL ANALYSIS

OF LOGIC,

BEING AN ESSAY TOWARDS A CALCULUS OF DEDUCTIVE REASONING.

BY GEORGE BOOLE.

[illegible]

ARISTOTLE, *Anal. Post.*, lib. I. cap. XI.

CAMBRIDGE:
MACMILLAN, BARCLAY, & MACMILLAN;
LONDON: GEORGE BELL.

1847

Recognized as beginning of mathematical logic,

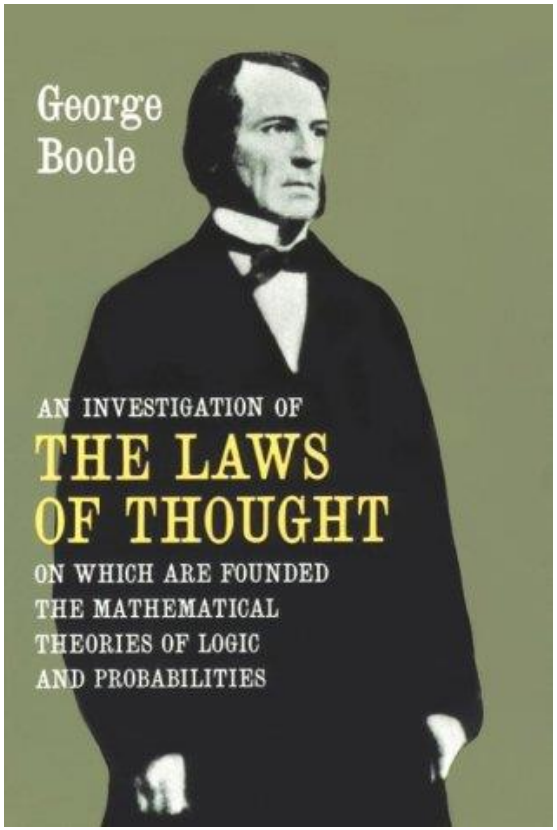
…but continues Leibnizian tradition

Uses arithmetical notation for logic.

$$x(u+v) = xu + xv,$$
$$xy = yx,$$
$$x^n = x,$$

From page 16

Boolean Algebra



Boole saw that symbols in his logic could be interpreted as **classes** (as in set theory) …

…or as propositions with truth values **0 and 1**.

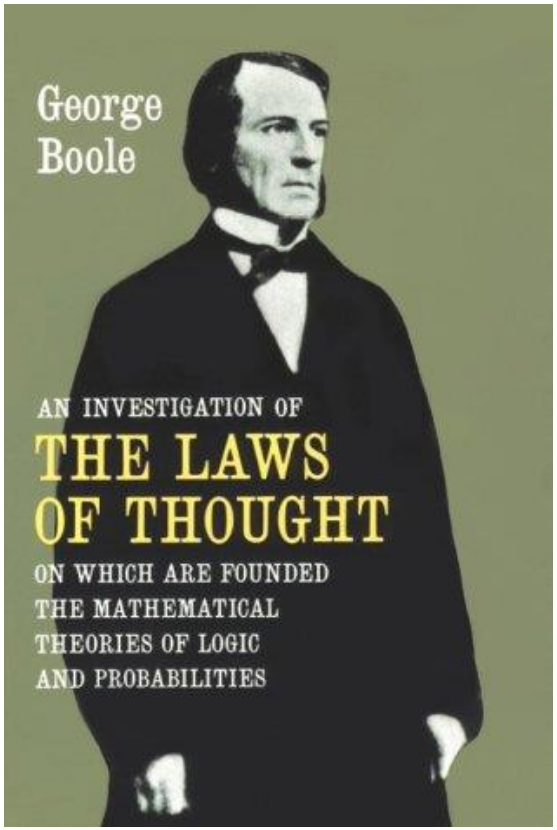
This gave rise to a **non-numerical algebra** and anticipated abstract algebra.

$$\begin{aligned} \{x(1-y)+y(1-x)\}^2 &= x(1-y)+y(1-x), \\ \{x+(1-x)\}^2 &= x+y(1-x). \end{aligned}$$

From page 40

This is “extensional” logic

Laws of Thought?



At the time, there was
no clear distinction between logic and human reasoning process.

This confusion ("psychologism")
wasn't cleared up until late 19th century
(Gottlob Frege, later William James).

We can see logic as a means of **justifying** conclusions we may arrive at by other means…

…not as a description of how we think.

Was Boole Aware of Leibnizian Logic?

Mary Everest Boole

Apparently not.

Boole's wife wrote:

"[Boole] felt as if Leibnitz had come and shaken hands with him across the centuries [after he heard about Leibniz]."

But it was part of the intellectual culture of the time.

Probability

The **rationalist project** had been **abandoned**, except in mathematics.

Science must be **empirical** and reason from **uncertain premises**.

We are interested in the **probability** of our conclusions.

Probability

This is the traditional, legal sense of probability.

Not the laws of chance, as in Bernoulli, Laplace et al.

Probability, *probity*, and *proof* have the same Latin root.

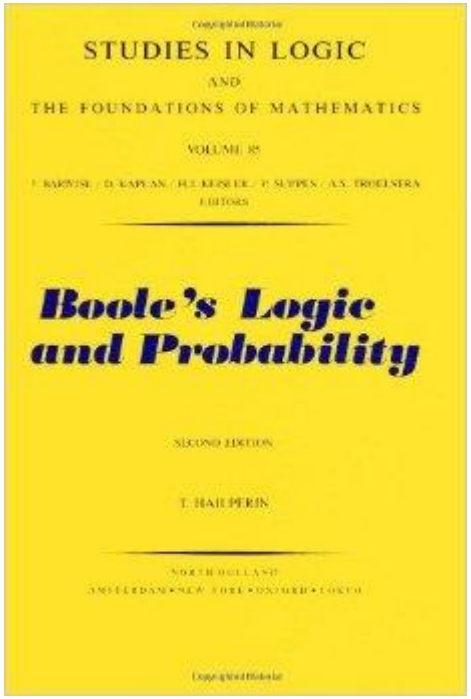
Pierre Simon Laplace

Probability

De Morgan and Boole saw the importance of reasoning under uncertainty.

Only Boole developed a **probability logic**, which was ignored or dismissed for a century.

In 1980s, Theodore Hailperin offered an interpretation of Boole's work using modern concept of **linear programming**.



Probability Logic

Statement	Probability
A	0.9
If A then B	0.8
If B then C	0.4

We can deduce C, but with what probability?

Boole's insights:

- We can only specify a **range** of probabilities for C.
- The range depends mathematically on the probabilities of **possible states of affairs** (possible worlds).

Probability Logic

Statement	Probability
A	0.9
If A then B	0.8
If B then C	0.4

Statement	Probability
A	0.9
not-A or B	0.8
not-B or C	0.4

First, interpret the if-then statements as "material conditionals"

Probability Logic

Statement	Probability
A	0.9
not-A or B	0.8
not-B or C	0.4

Identify the possible outcomes (possible worlds), each having an **unknown** probability.

There are 8 possible worlds:

A	B	C	Prob.
false	false	false	p_{000}
false	false	true	p_{001}
false	true	false	p_{010}
false	true	true	p_{011}
true	false	false	p_{100}
true	false	true	p_{101}
true	true	false	p_{110}
true	true	true	p_{111}

Probability Logic

Statement	Probability
A	0.9
not-A or B	0.8
not-B or C	0.4

$$p_{100} + p_{101} + p_{110} + p_{111} = 0.9$$

The worlds in which A is true must have probabilities that sum to 0.9.

There are 8 possible worlds:

A	B	C	Prob.
false	false	false	p_{000}
false	false	true	p_{001}
false	true	false	p_{010}
false	true	true	p_{011}
true	false	false	p_{100}
true	false	true	p_{101}
true	true	false	p_{110}
true	true	true	p_{111}

Probability Logic

Statement	Probability
A	0.9
not-A or B	0.8
not-B or C	0.4

$$p_{100} + p_{101} + p_{110} + p_{111} = 0.9$$

$$p_{000} + p_{001} + p_{010} + p_{011} + p_{110} + p_{111} = 0.8$$

There are 8 possible worlds:

A	B	C	Prob.
false	false	false	p_{000}
false	false	true	p_{001}
false	true	false	p_{010}
false	true	true	p_{011}
true	false	false	p_{100}
true	false	true	p_{101}
true	true	false	p_{110}
true	true	true	p_{111}

Probability Logic

Statement	Probability
A	0.9
not-A or B	0.8
not-B or C	0.4

$$p_{100} + p_{101} + p_{110} + p_{111} = 0.9$$

$$p_{000} + p_{001} + p_{010} + p_{011} + p_{110} + p_{111} = 0.8$$

$$p_{000} + p_{001} + p_{011} + p_{100} + p_{101} + p_{111} = 0.4$$

There are 8 possible worlds:

A	B	C	Prob.
false	false	false	p_{000}
false	false	true	p_{001}
false	true	false	p_{010}
false	true	true	p_{011}
true	false	false	p_{100}
true	false	true	p_{101}
true	true	false	p_{110}
true	true	true	p_{111}

Probability Logic

Statement	Probability
A	0.9
not-A or B	0.8
not-B or C	0.4

$$p_{100} + p_{101} + p_{110} + p_{111} = 0.9$$

$$p_{000} + p_{001} + p_{010} + p_{011} + p_{110} + p_{111} = 0.8$$

$$p_{000} + p_{001} + p_{011} + p_{100} + p_{101} + p_{111} = 0.4$$

$$p_{000} + \ldots + p_{111} = 1$$

There are 8 possible worlds:

A	B	C	Prob.
false	false	false	p_{000}
false	false	true	p_{001}
false	true	false	p_{010}
false	true	true	p_{011}
true	false	false	p_{100}
true	false	true	p_{101}
true	true	false	p_{110}
true	true	true	p_{111}

Probability Logic

Statement	Probability
A	0.9
not-A or B	0.8
not-B or C	0.4

$$p_{100} + p_{101} + p_{110} + p_{111} = 0.9$$

$$p_{000} + p_{001} + p_{010} + p_{011} + p_{110} + p_{111} = 0.8$$

$$p_{000} + p_{001} + p_{011} + p_{100} + p_{101} + p_{111} = 0.4$$

$$p_{000} + \ldots + p_{111} = 1$$

Minimize and maximize probability of C:

$$p_{001} + p_{011} + p_{101} + p_{111}$$

subject to these equations and $p_{ijk} \geq 0$

There are 8 possible worlds:

A	B	C	Prob.
false	false	false	p_{000}
false	false	true	p_{001}
false	true	false	p_{010}
false	true	true	p_{011}
true	false	false	p_{100}
true	false	true	p_{101}
true	true	false	p_{110}
true	true	true	p_{111}

Probability Logic

Statement	Probability
A	0.9
not-A or B	0.8
not-B or C	0.4

$$p_{100} + p_{101} + p_{110} + p_{111} = 0.9$$

$$p_{000} + p_{001} + p_{010} + p_{011} + p_{110} + p_{111} = 0.8$$

$$p_{000} + p_{001} + p_{011} + p_{100} + p_{101} + p_{111} = 0.4$$

$$p_{000} + \ldots + p_{111} = 1$$

Minimize and maximize probability of C:

$$p_{001} + p_{011} + p_{101} + p_{111}$$

subject to these equations and $p_{ijk} \geq 0$

This is a **linear programming** problem, first clearly stated by Leonid Kantorovich in the 1930s.

The result is a **range** of probabilities for C:

0.1 to 0.4

Probability Logic

Statement	Probability
A	0.9
not-A or B	0.8
not-B or C	0.4

$$p_{100} + p_{101} + p_{110} + p_{111} = 0.9$$

$$p_{000} + p_{001} + p_{010} + p_{011} + p_{110} + p_{111} = 0.8$$

$$p_{000} + p_{001} + p_{011} + p_{100} + p_{101} + p_{111} = 0.4$$

$$p_{000} + \ldots + p_{111} = 1$$

Minimize and maximize probability of C:

$$p_{001} + p_{011} + p_{101} + p_{111}$$

subject to these equations and $p_{ijk} \geq 0$

This is a **linear programming** problem, first clearly stated by Leonid Kantorovich in the 1930s.

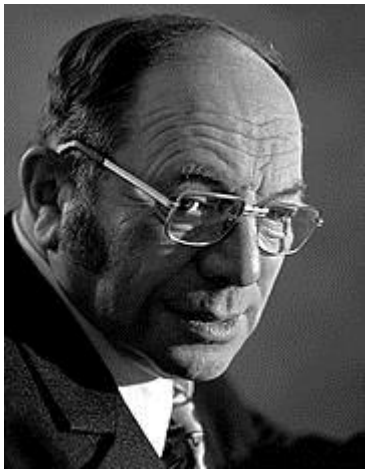
The result is a **range** of probabilities for C:

0.1 to 0.4

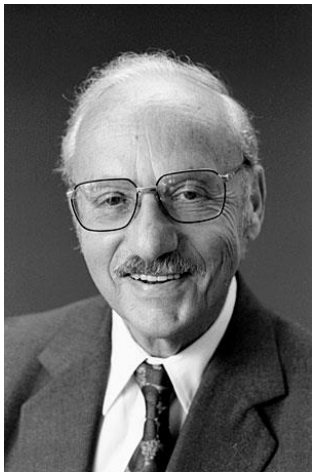
Probability Logic

Linear programming is one of the fundamental problems of applied mathematics today.

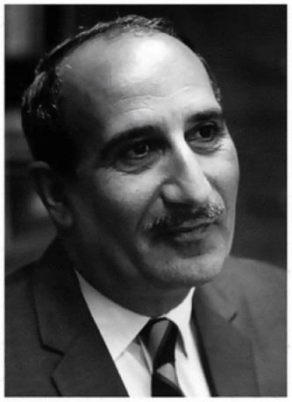
“Column generation” methods deal with the large (exponential) number of variables in Boole’s problem.



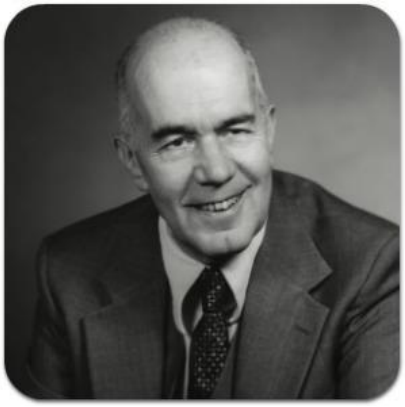
Leonid Kantorovich



George Dantzig



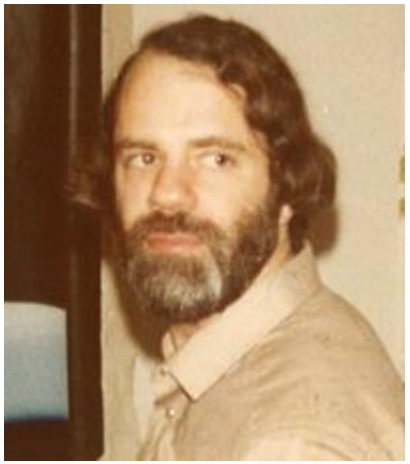
Andrew Charnes



William Cooper

Probability Logic

Possible-world semantics re-emerged in modal logic of 1960s



Saul Kripke

Probability logic re-invented in AI community of 1980s.

Nils Nilsson

Reasoning under Uncertainty

A central problem today.

Associate each inference with its **probability**, **relevance**, or **confidence**.

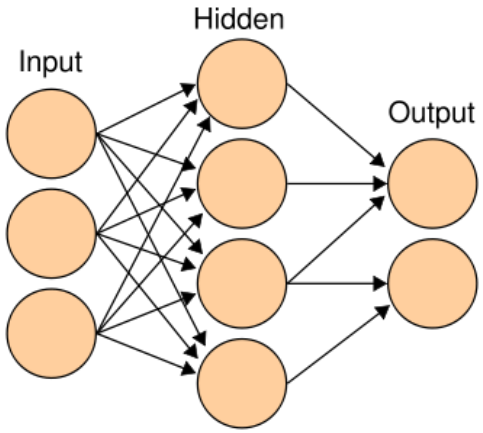
IBM's Watson, Jeopardy champion

Reasoning under Uncertainty

Two approaches to deriving information from big data:

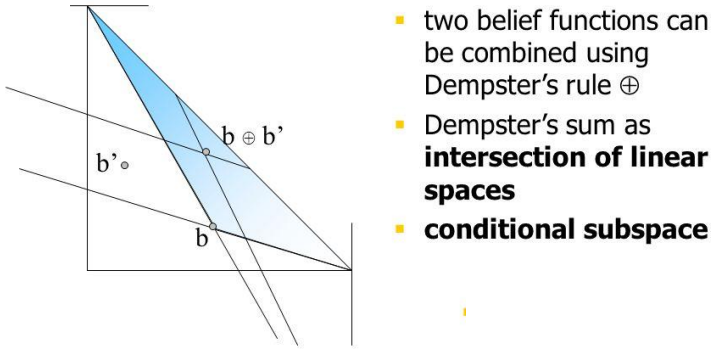
Statistical inference

Pattern recognition, neural networks, etc.



Logical inference

Probability logic, belief logic, epistemic logic, etc.



- Dempster's combination rule

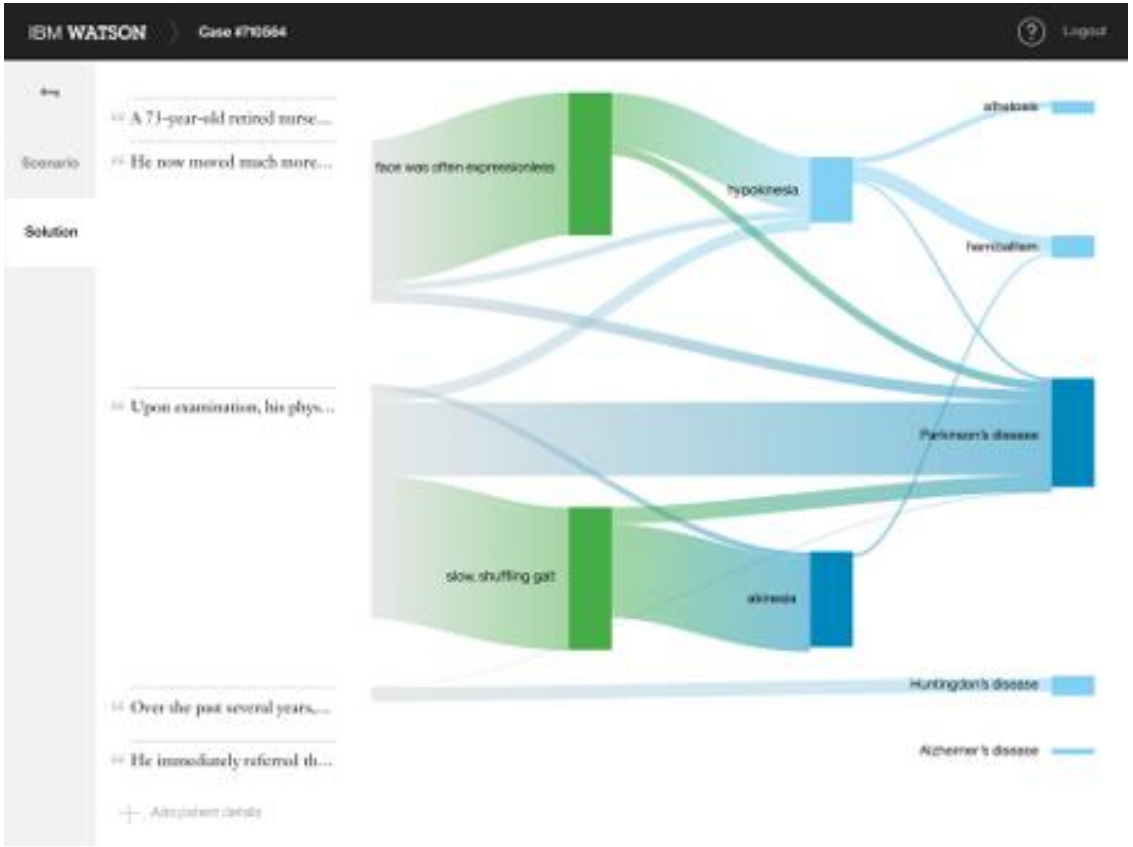
Mining the Medical Literature

Watson technology first applied to medicine (WatsonPaths).

Draws inferences from medical literature and clinical guidelines.

About 1 million articles listed per year in PubMed.

Probably 1.5-2 million overall.



Medical Crowdsourcing

Google
Flu Trends

Mobile phone
Ebola tracking
in east Africa

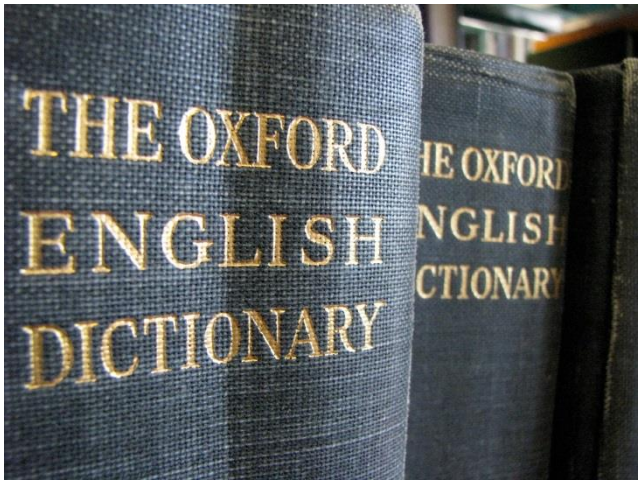
Mining
pharmacy
records to
predict
epidemics

Business Applications

Individual dossiers from social network connections & point-of sale transactions.

Used to place individualized ads and….

Linguistic Crowdsourcing



800+ volunteers sent in over 2 tons of quotation slips over 21 years.

Google Translate

For each language we need:

English translations of documents containing 200 millions words (e.g., EU Parliament records)

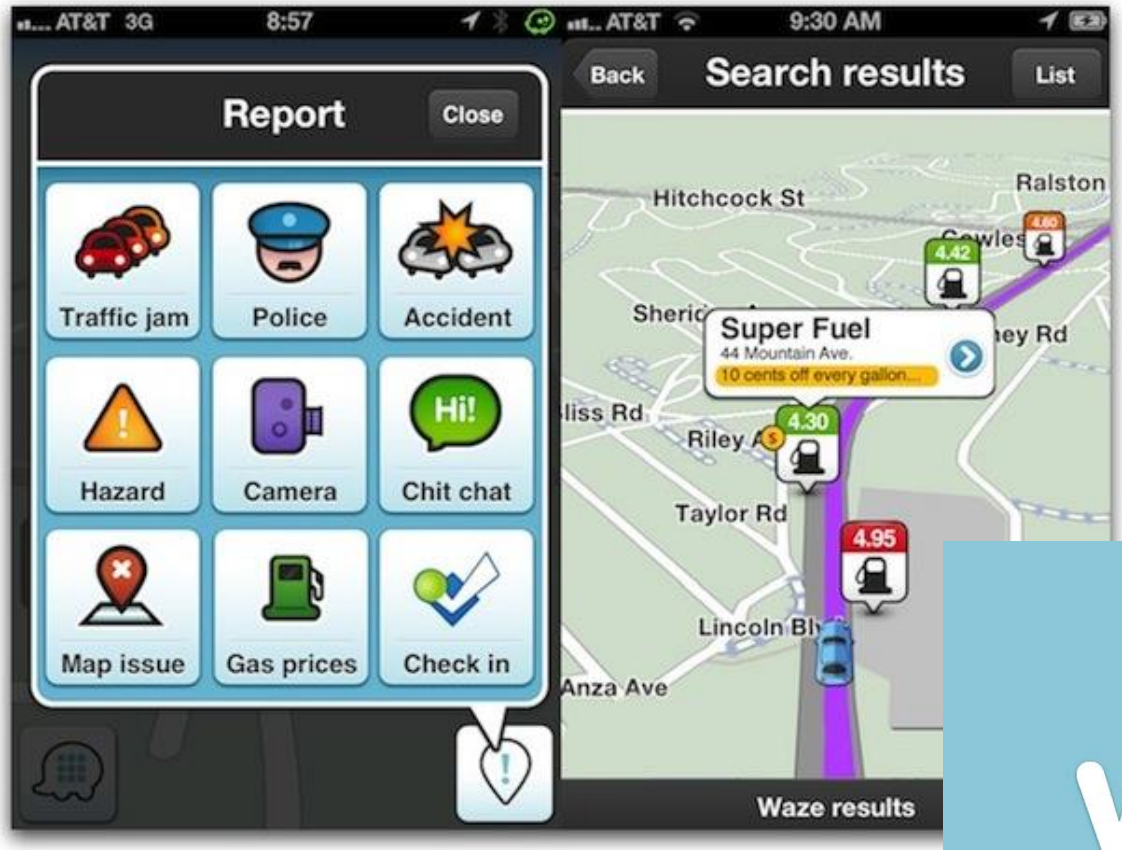
Documents containing 1 billion words in each language.

News Crowdsourcing

Monitors discussion of news items on Twitter & Instagram to check accuracy of reported story.

Traffic Crowdsourcing

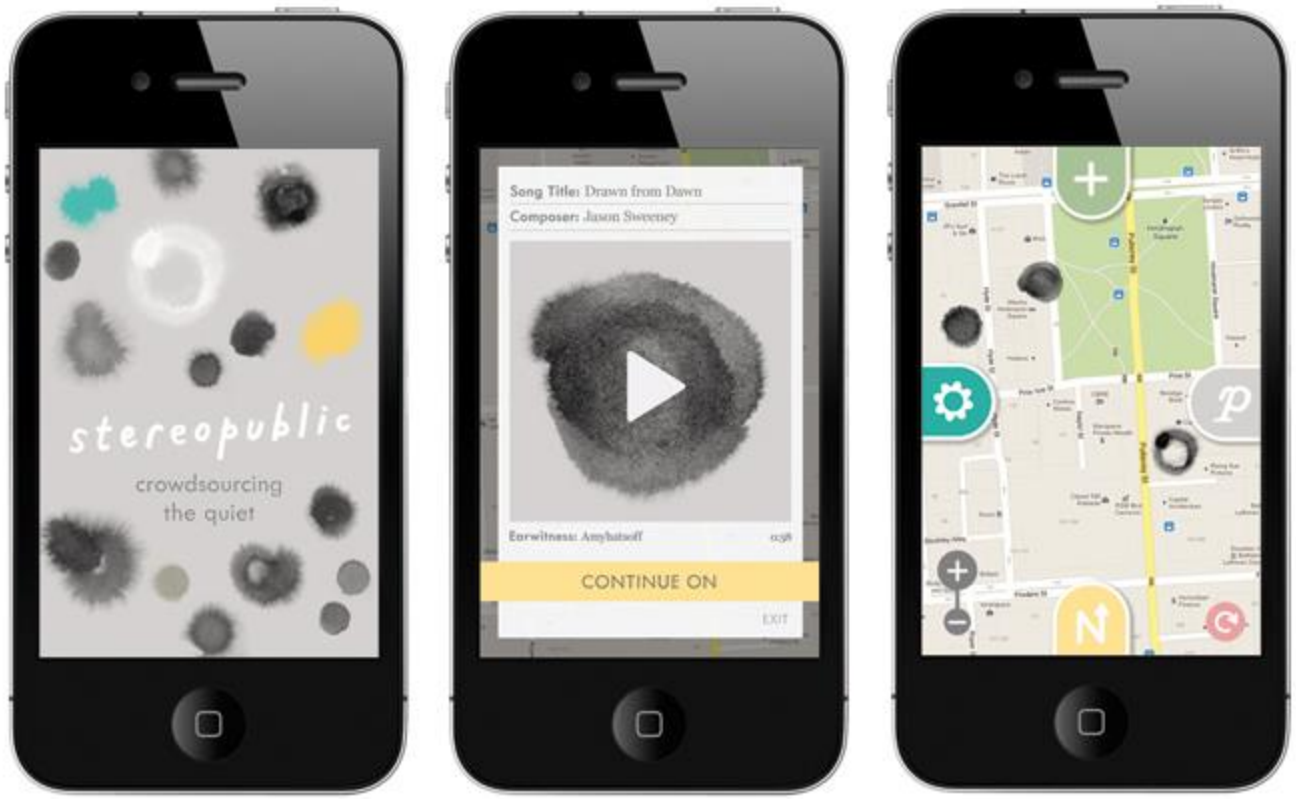
Monitors GPS coordinates of participating smart phone users.



waze
OUTSMARTING TRAFFIC, TOGETHER.

Peace & Quiet Crowdsourcing

Users pinpoint and monitor quiet spots for 30 seconds.

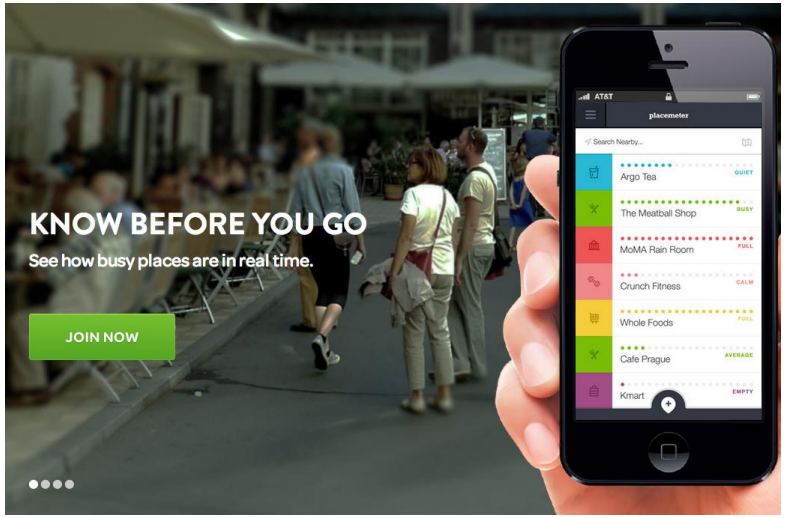


Toilet Crowdsourcing

Users review toilets by geolocation and use app to find the next one.

Crowd Crowdsourcing

Participants paid $50/month to mount iPhone on windows to observe pedestrian traffic..



placemeter

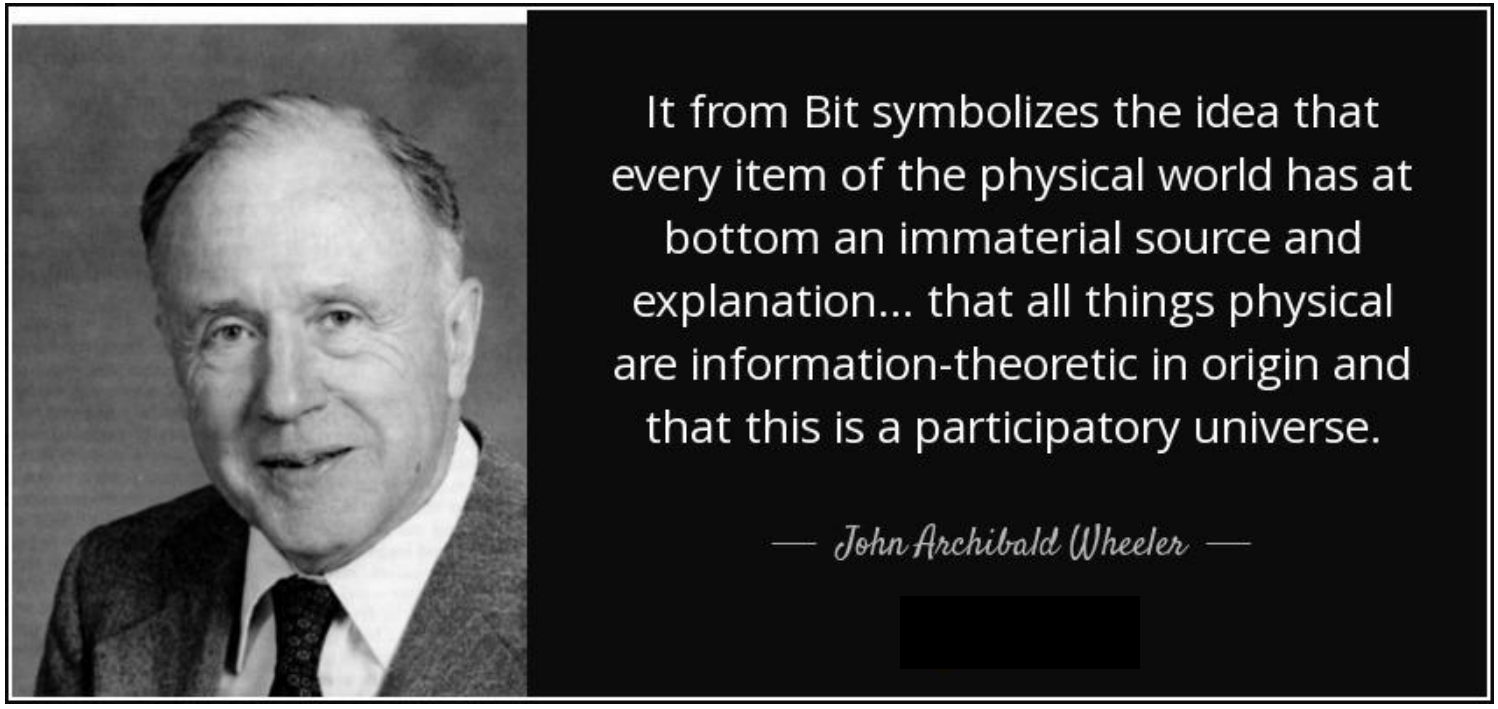
Why Our Faith in Big Data?

Data are now our world, our reality.

We must believe there is order, reason, and knowledge in that world.

The Universe as Data

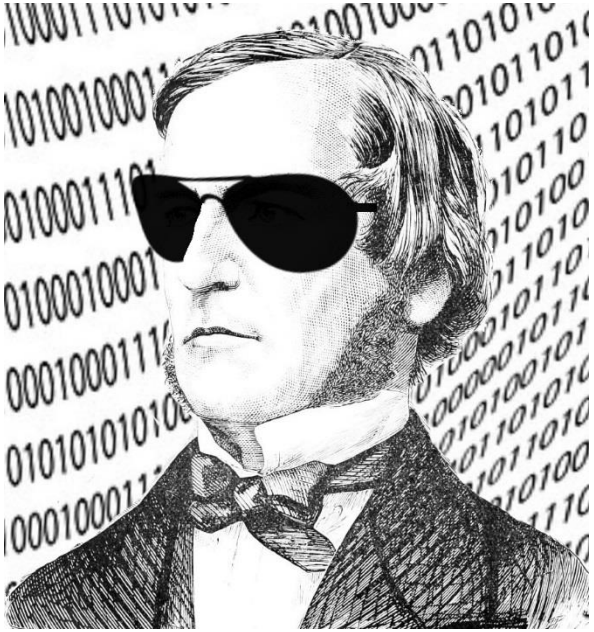
Physicist John Wheeler proposed (1990) that the universe is best understood as a **database** or **information source**.



"That which we call reality arises in the last analysis from the posing of yes-no questions and the registering of equipment-evoked responses."

The Universe as Data

Physicist John Wheeler proposed (1990) that the universe is best understood as a **database** or **information source**.



"That which we call reality arises in the last analysis from the posing of yes-no questions and the registering of equipment-evoked responses."

From Boole to (Really) Big Data

George Boole was the first to propose a logical system for deducing conclusions from a mass of uncertain data.