

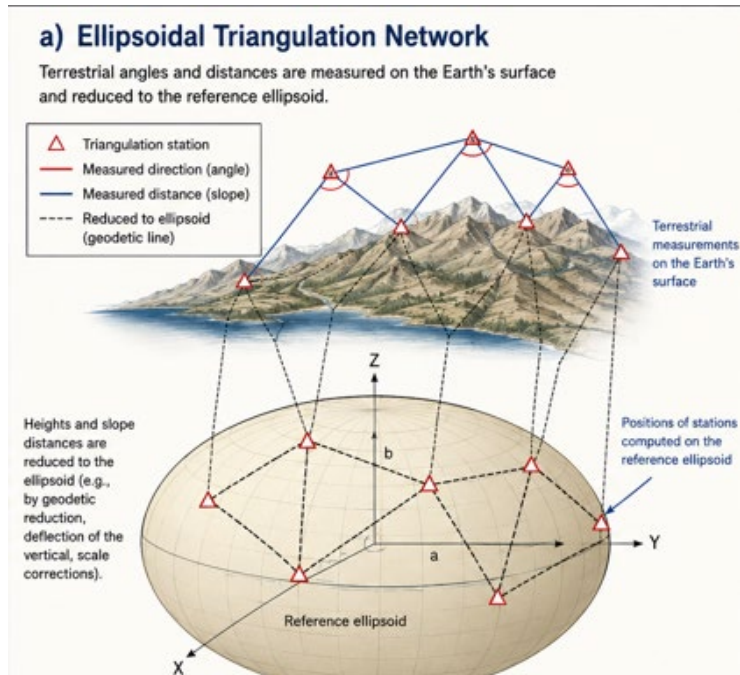
When Success Hides the Need to Understand

Mathematical Geodesy in an Age of Data and Computing Abundance

Peter J.G. Teunissen

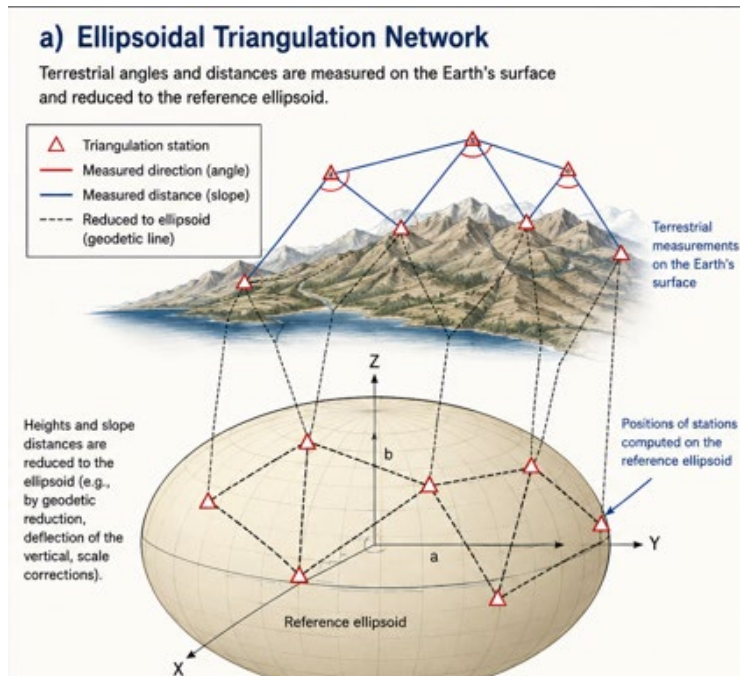
In the Past: Data and Computing Scarcity

- **Triangulation networks (2D+1D)**
 - Reduction + ellipsoidal networks



In the Past: Data and Computing Scarcity

- **Triangulation networks (2D+1D)**
 - Reduction + ellipsoidal networks

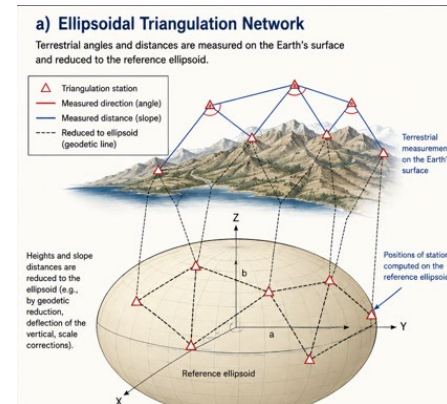


"Geodesists measure on a surface they cannot compute on, compute on a surface they cannot measure on, and therefore invented a third surface on which they can neither measure nor compute: the geoid."

In the Past: Data and Computing Scarcity

- Triangulation networks (2D+1D)
 - Reduction + ellipsoidal networks
- **Computing limitations:**
 - Elimination methods
 - Sparse matrix methods
 - Recursive algorithms

b) 1960s Punch Card Mainframe Computer

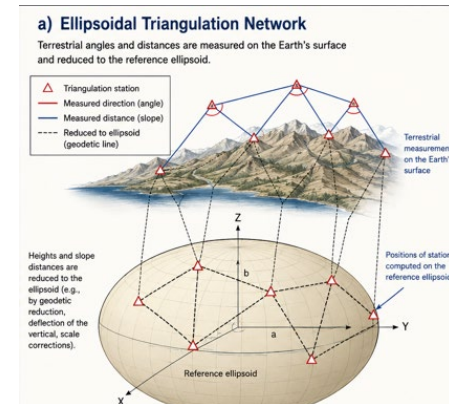


In the Past: Data and Computing Scarcity

- Triangulation networks (2D+1D)
 - Reduction + ellipsoidal networks
- **Computing limitations:**
 - Elimination methods
 - Sparse matrix methods
 - Recursive algorithms

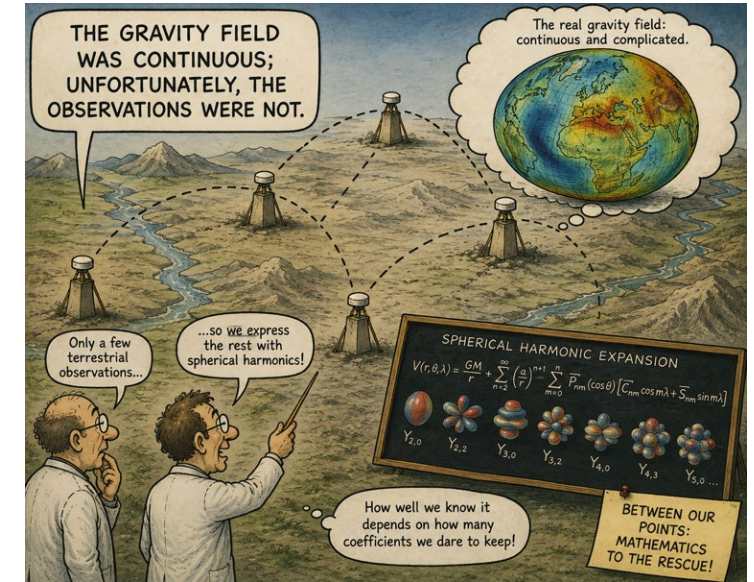
"In the 1970s, the easiest way to speed up a computation was to improve the mathematics."

b) 1960s Punch Card Mainframe Computer



In the Past: Data and Computing Scarcity

- Triangulation networks (2D+1D)
 - Reduction + ellipsoidal networks
- Computing limitations:
 - Elimination methods
 - Sparse matrix methods
 - Recursive algorithms
- **Earth gravity field (only sparsely observed):**
 - Spherical harmonics for global interpolation
 - Stokes' formula
 - Molodensky theory

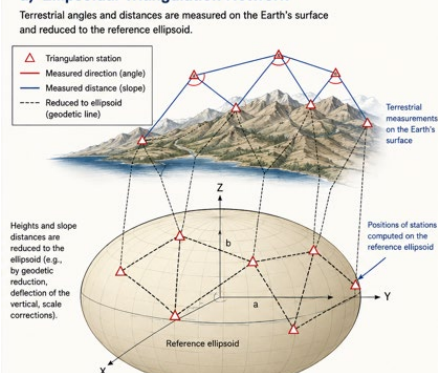


"Our gravity model had excellent global coverage
—
once you connected the dots."

b) 1960s Punch Card Mainframe Computer

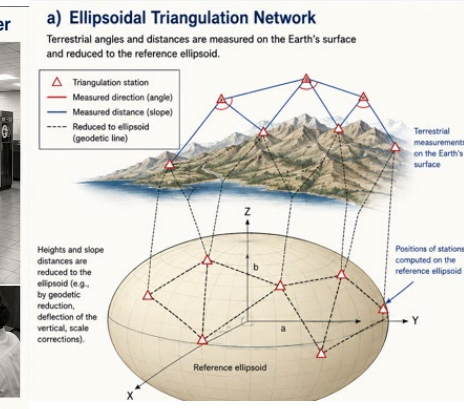
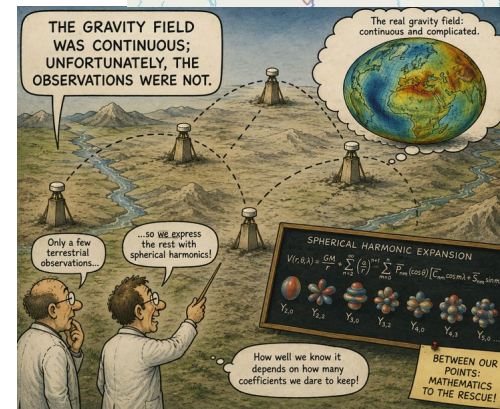


a) Ellipsoidal Triangulation Network



In the Past: Data and Computing Scarcity

- Triangulation networks (2D+1D)
 - Reduction + ellipsoidal networks
- Computing limitations:
 - Elimination methods
 - Sparse matrix methods
 - Recursive algorithms
- Earth gravity field (only sparsely observed):
 - Spherical harmonics for global interpolation
 - Stokes' formula
 - Molodensky theory
- **Early satellite geodesy**
(infrequent tracking/passes)
 - Orbit determination
 - Dynamic modelling
 - Kalman filtering



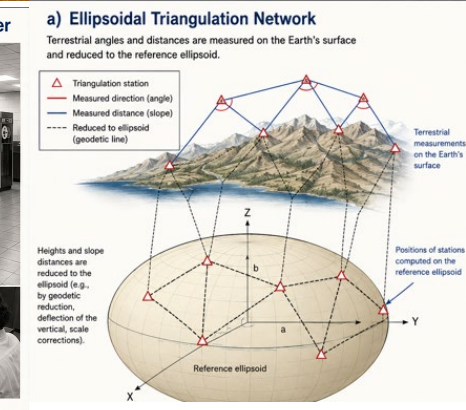
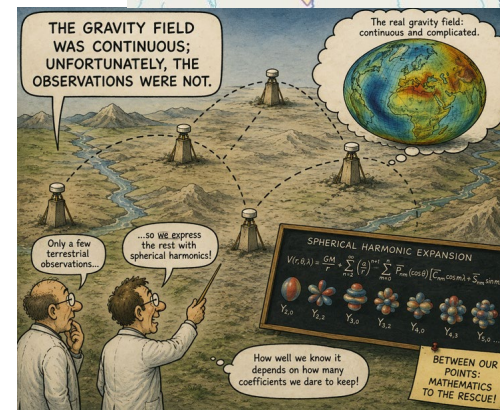
In the Past: Data and Computing Scarcity

- Triangulation networks (2D+1D)
 - Reduction + ellipsoidal networks
- Computing limitations:
 - Elimination methods
 - Sparse matrix methods
 - Recursive algorithms
- Earth gravity field (only sparsely observed):
 - Spherical harmonics for global interpolation
 - Stokes' formula
 - Molodensky theory

Early satellite geodesy (infrequent tracking/passes)

- Orbit determination
- Dynamic modelling
- Kalman filtering

"The satellite disappeared over the horizon
The equations kept following it"



In the Past: **Data and Computing Scarcity**

- Triangulation networks (2D+1D)
 - Reduction + ellipsoidal networks
- Computing limitations:
 - Elimination methods
 - Sparse matrix methods
 - Recursive algorithms
- Earth gravity field (only sparsely observed):
 - Spherical harmonics for global interpolation
 - Stokes' formula
 - Molodensky theory
- Early satellite geodesy (few tracking/infrequent passes)
 - Orbit determination
 - Dynamic modelling
 - Kalman filtering

Theory was not optional; Scarcity drove theory

- To extract every bit of information
- To enable feasible computation

Sherlock was right

Data! Data! Data!
I Can't Make Bricks Without Clay!



Sherlock Holmes in “The Adventure in the Copper Beeches” (Arthur Conan Doyle, 1892)

Today: World has Changed

Data and Computing Power in Abundance

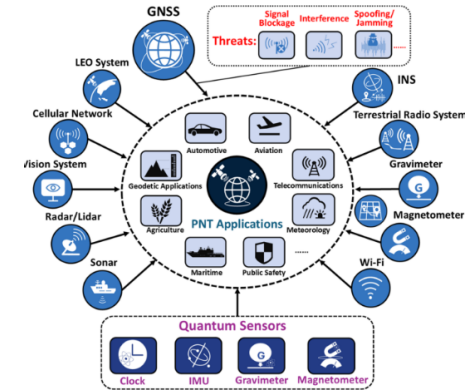


Then

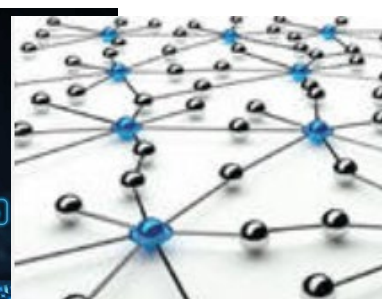
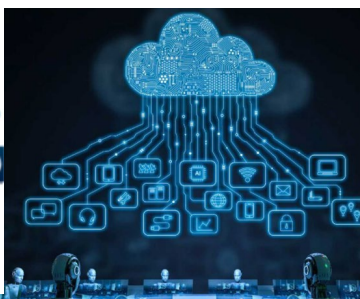
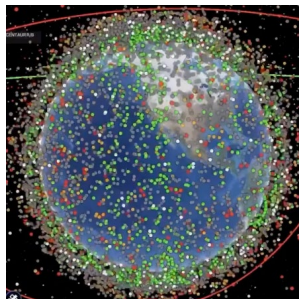
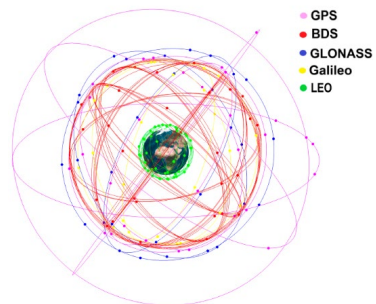
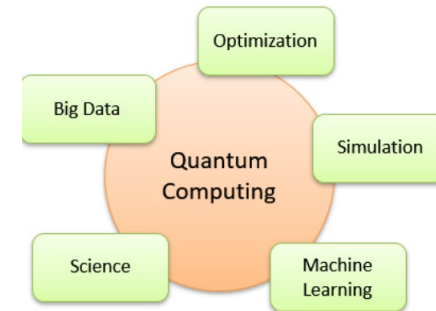
- few observations
- little computation
- simple software

Now

- continuous GNSS
- multi satellite missions
- cloud computing
- AI
- data centres
- quantum ...



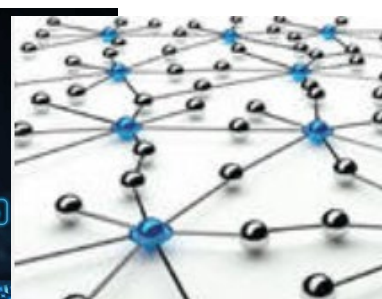
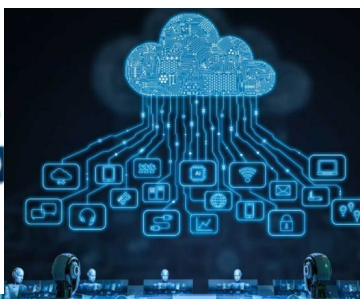
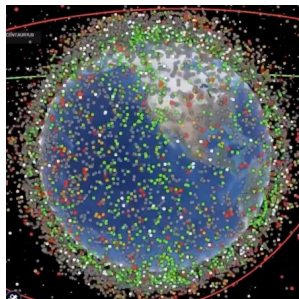
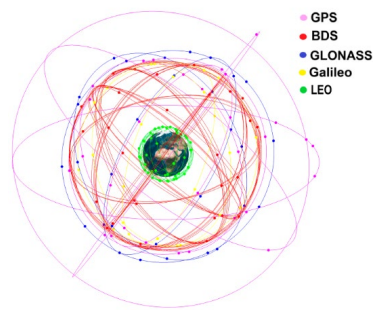
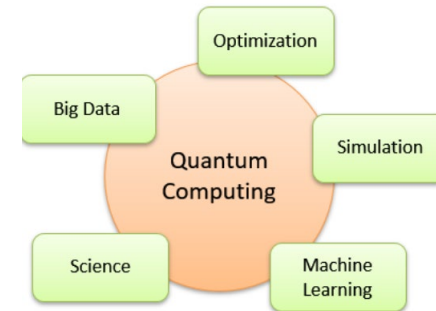
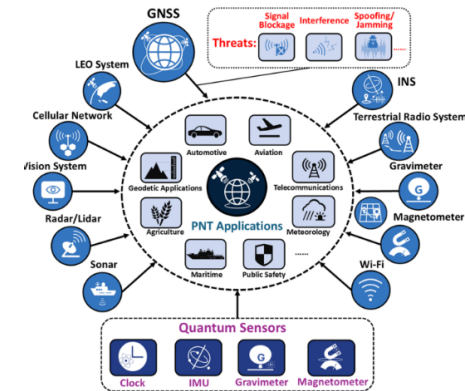
The bottleneck has moved.



Data and Computing Power in Abundance

Modern geodesy has reached an interesting stage: our measurements have become so precise, our datasets so large, and our software so powerful that we can obtain fantastic results

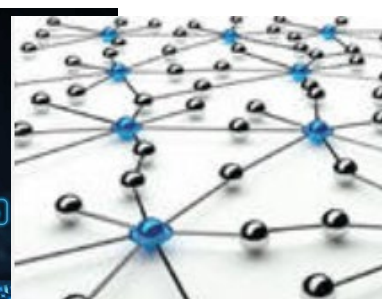
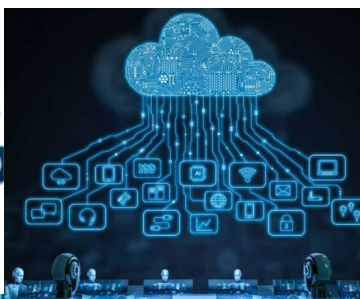
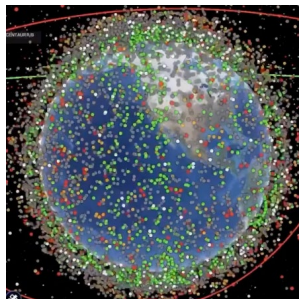
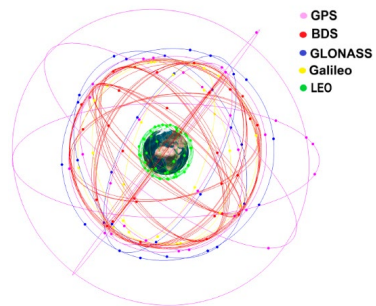
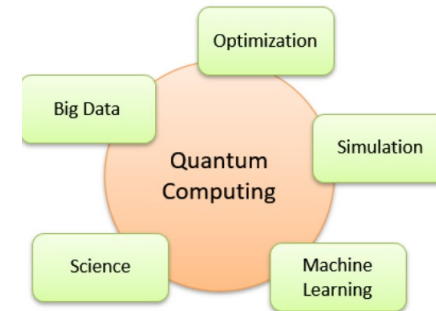
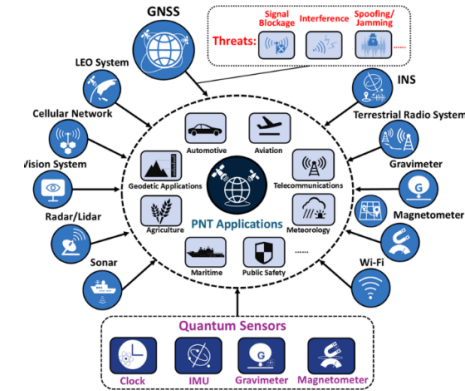
This is a triumph of modern geodesy



Data and Computing Power in Abundance

Modern geodesy has reached an interesting stage: our measurements have become so precise, our datasets so large, and our software so powerful that we can obtain fantastic results

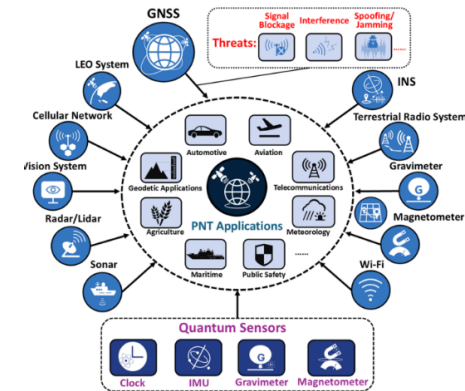
This is a triumph of modern geodesy & a very good reason not to abandon Mathematical Geodesy



Data and Computing Power in Abundance

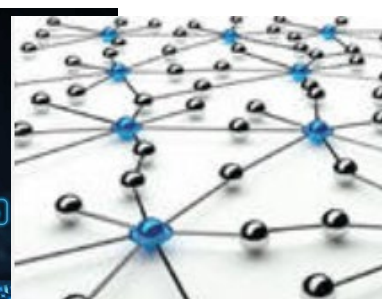
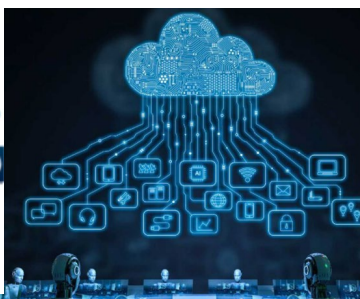
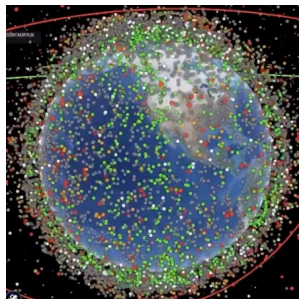
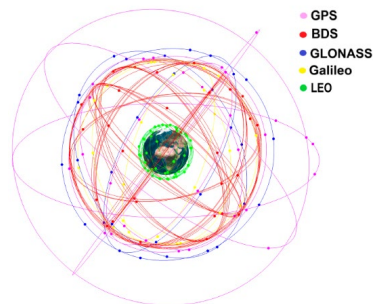
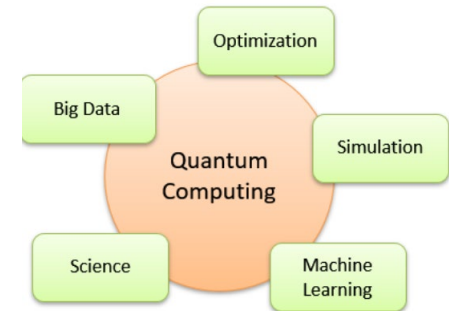
Modern geodesy has reached an interesting stage: our measurements have become so precise, our datasets so large, and our software so powerful that we can obtain fantastic results

This is a triumph of modern geodesy & a very good reason not to abandon Mathematical Geodesy



The geodesist of the past worried whether the theory was good enough for the observations.

The geodesist of today should worry whether the observations have become so good that they hide weaknesses in the theory



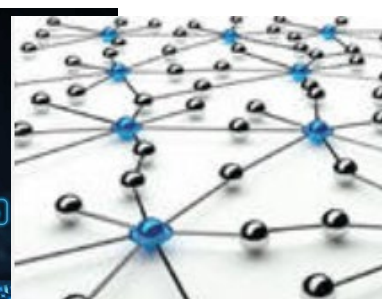
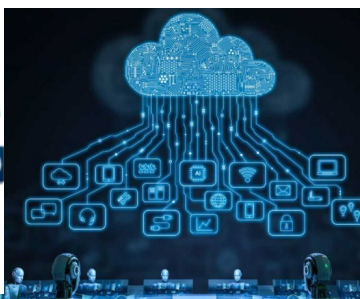
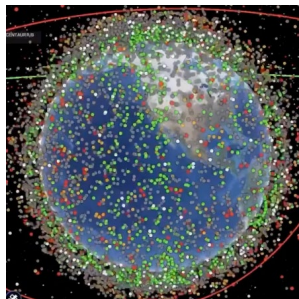
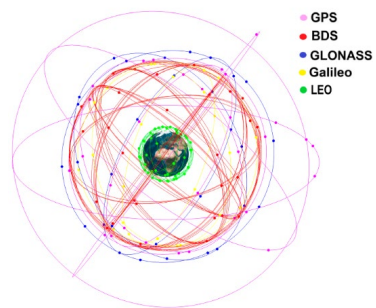
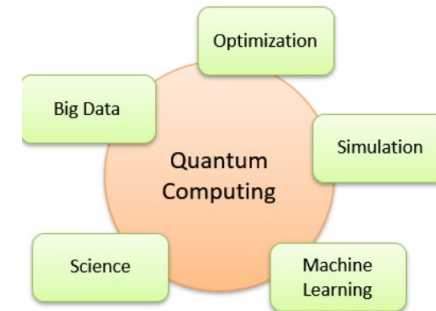
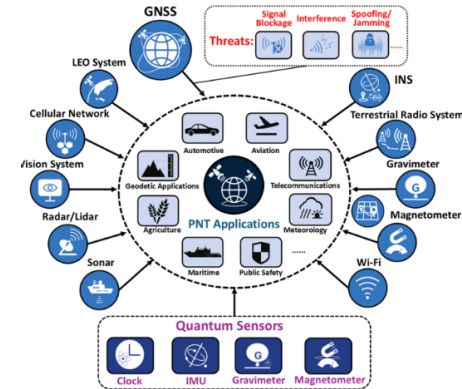
Data and Computing Power in Abundance

Modern geodesy has reached an interesting stage: our measurements have become so precise, our datasets so large, and our software so powerful that we can obtain fantastic results

This is a triumph of modern geodesy & a very good reason not to abandon Mathematical Geodesy

Success has become easier than understanding.

The geodesist of the past worried whether the theory was good enough for the observations.
The geodesist of today should worry whether the observations have become so good that they hide weaknesses in the theory



Data and Computing Power in Abundance

Modern geodesy has reached an interesting stage: our measurements have become so precise, our datasets so large, and our software so powerful that we can obtain fantastic results

This is a triumph of modern geodesy & a very good reason not to abandon Mathematical Geodesy

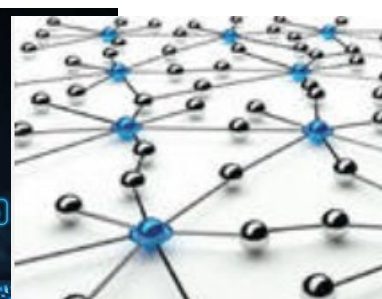
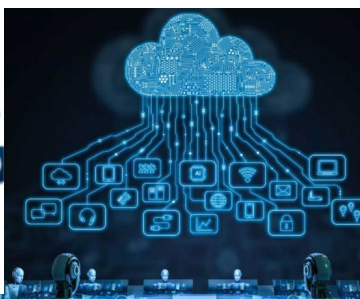
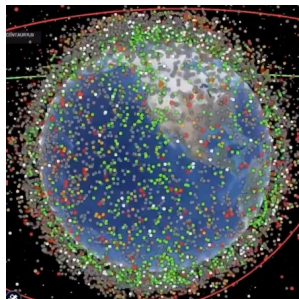
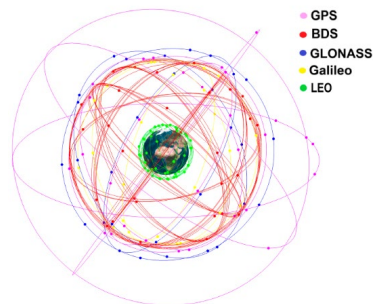
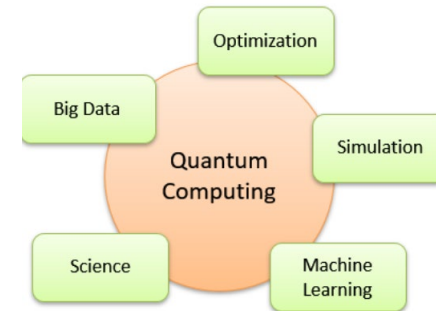
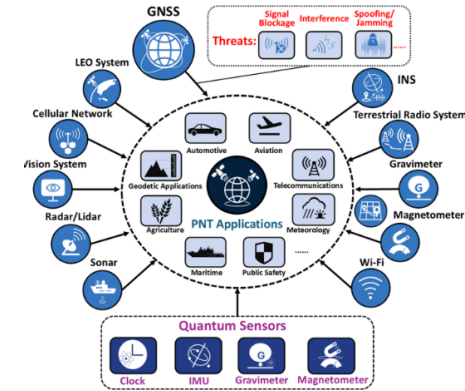
Success has become easier than understanding.



The geodesist of the past worried whether the theory was good enough for the observations.

The geodesist of today should worry whether the observations have become so good that they hide weaknesses in the theory

- **Big data** can compensate for many things. Lack of mathematical understanding is not one of them—it can merely conceal it for longer



Data and Computing Power in Abundance

Modern geodesy has reached an interesting stage: our measurements have become so precise, our datasets so large, and our software so powerful that we can obtain fantastic results

This is a triumph of modern geodesy & a very good reason not to abandon Mathematical Geodesy

Success has become easier than understanding.

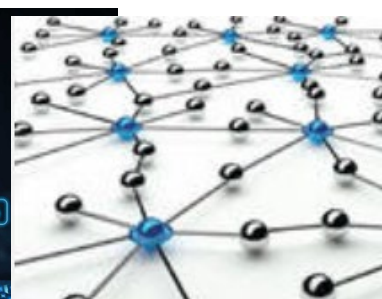
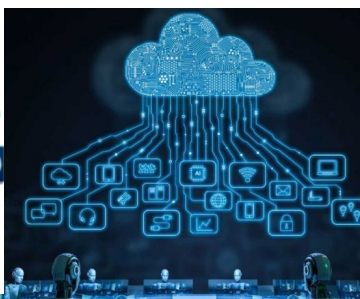
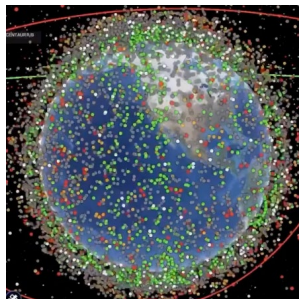
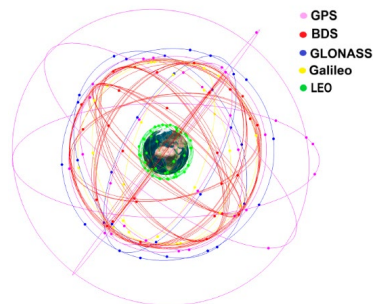
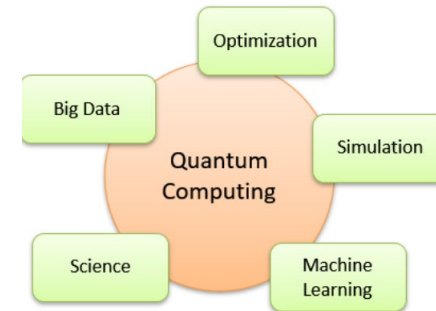
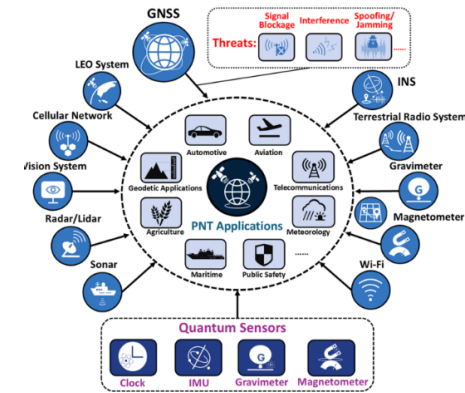


The geodesist of the past worried whether the theory was good enough for the observations.

The geodesist of today should worry whether the observations have become so good that they hide weaknesses in the theory

- **Big data** can compensate for many things. Lack of mathematical understanding is not one of them—it can merely conceal it for longer
- **Computers** answer the question we asked. They don't tell us whether it was the right question.

With enough **data** and enough **computing** power, even questionable mathematics can produce many convincing decimal places. Unfortunately, decimal places do not come with a certificate of understanding.



The Hidden Risks of Taking Big Data at Face Value

- Big data can make a wrong answer look increasingly convincing, because they reduce the uncertainty while leaving the misconception untouched.

Medicine:

Suppose a blood-pressure sensor has a calibration error of 5 mmHg.

Measure:

- once,
- one thousand times,
- one million times.

Average still contains the same 5 mmHg bias.

The only thing that improves is your confidence in the biased value.

The Hidden Risks of Taking Big Data at Face Value

- Big data can make a wrong answer look increasingly convincing, because they reduce the uncertainty while leaving the misconception untouched.

Medicine:

Suppose a blood-pressure sensor has a calibration error of 5 mmHg.

Measure:

- once,
- one thousand times,
- one million times.

Average still contains the same 5 mmHg bias.

The only thing that improves is your confidence in the biased value.

$$\text{MSE} = \text{VAR} + \text{BIAS}^2$$



The Hidden Risks of Taking Big Data at Face Value

- Big data can make a wrong answer look increasingly convincing, because they reduce the uncertainty while leaving the misconception untouched.
- Big data can postpone the moment at which we discover that we have misunderstood something.

Ptolemy's system of epicycles (used >1400years) :

As astronomical observations became more accurate, the Ptolemaic model did **not** immediately collapse.

Instead,

- more epicycles were added,
- parameters were refined,
- predictions improved.

The increasing amount of data actually **kept the model alive** for centuries.

Only much later did a simpler theoretical framework emerge

The Hidden Risks of Taking Big Data at Face Value

- Big data can make a wrong answer look increasingly convincing, because they reduce the uncertainty while leaving the misconception untouched.
- Big data can postpone the moment at which we discover that we have misunderstood something.

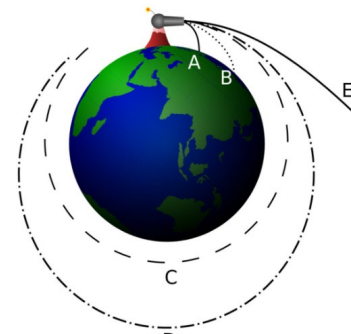
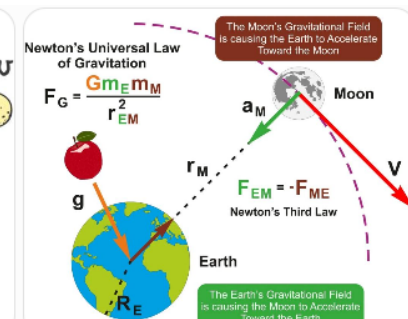
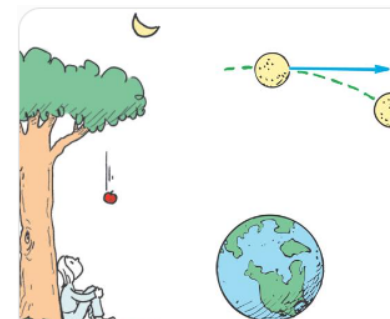
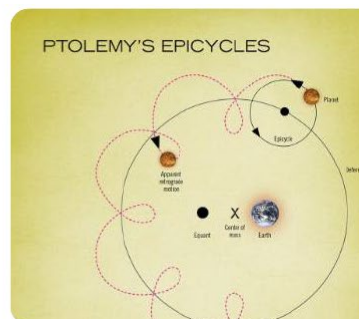
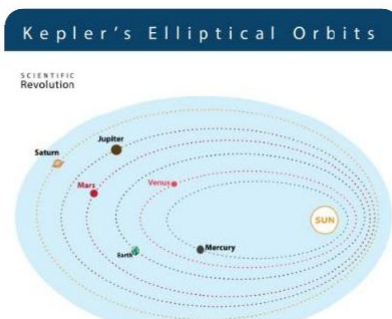
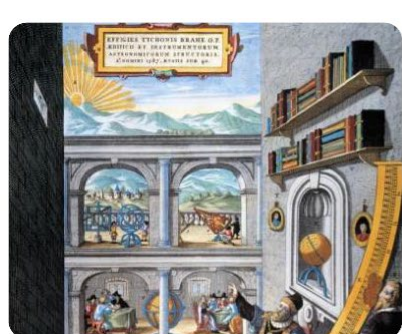
Ptolemy's system of epicycles (used >1400 years) :
As astronomical observations became more accurate, the Ptolemaic model did **not** immediately collapse.

Instead,

- more epicycles were added,
- parameters were refined,
- predictions improved.

The increasing amount of data actually **kept the model alive** for centuries.

Only much later did a simpler theoretical framework emerge



The Hidden Risks of Taking Big Data at Face Value

- Big data can make a wrong answer look increasingly convincing, because they reduce the uncertainty while leaving the misconception untouched.
- Big data can postpone the moment at which we discover that we have misunderstood something.
- Big data often test a method where it is already comfortable.

Imagine driving a car whose brakes fail only on steep downhill roads.

For years, every journey is uneventful.

Each successful trip increases your confidence.

Yet none of those journeys tested the assumption that the brakes always work.

The problem was present from the beginning.

Success merely concealed it.

The Hidden Risks of Taking Big Data at Face Value

- Big data can make a wrong answer look increasingly convincing, because they reduce the uncertainty while leaving the misconception untouched.
- Big data can postpone the moment at which we discover that we have misunderstood something.
- Big data often test a method where it is already comfortable.

Imagine driving a car whose brakes fail only on steep downhill roads.

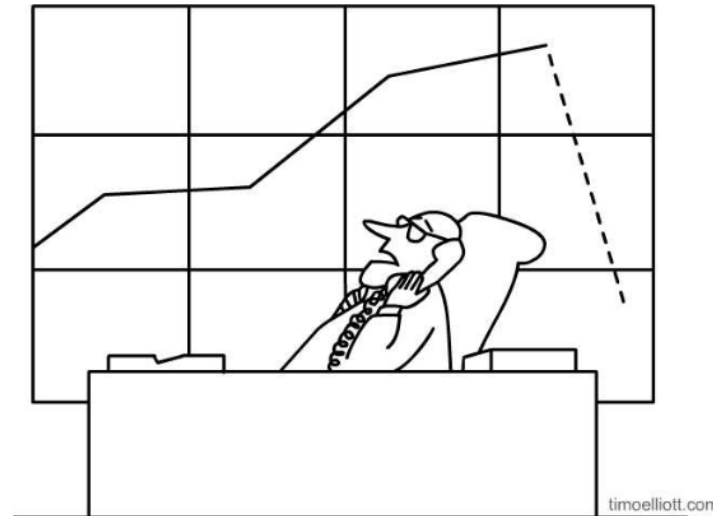
For years, every journey is uneventful.

Each successful trip increases your confidence.

Yet none of those journeys tested the assumption that the brakes always work.

The problem was present from the beginning.

Success merely concealed it.



"It turns out that failure WAS an option after all..."

The Hidden Risks of Taking Big Data at Face Value

- Big data can make a wrong answer look increasingly convincing, because they reduce the uncertainty while leaving the misconception untouched.
- Big data can postpone the moment at which we discover that we have misunderstood something.
- Big data often test a method where it is already comfortable.
- With big data and analytical flexibility, almost every hypothesis can find supporting evidence

Suppose someone develops a new neural network.

It performs

- worse on Dataset A,
- equal on Dataset B,
- better on Dataset C.

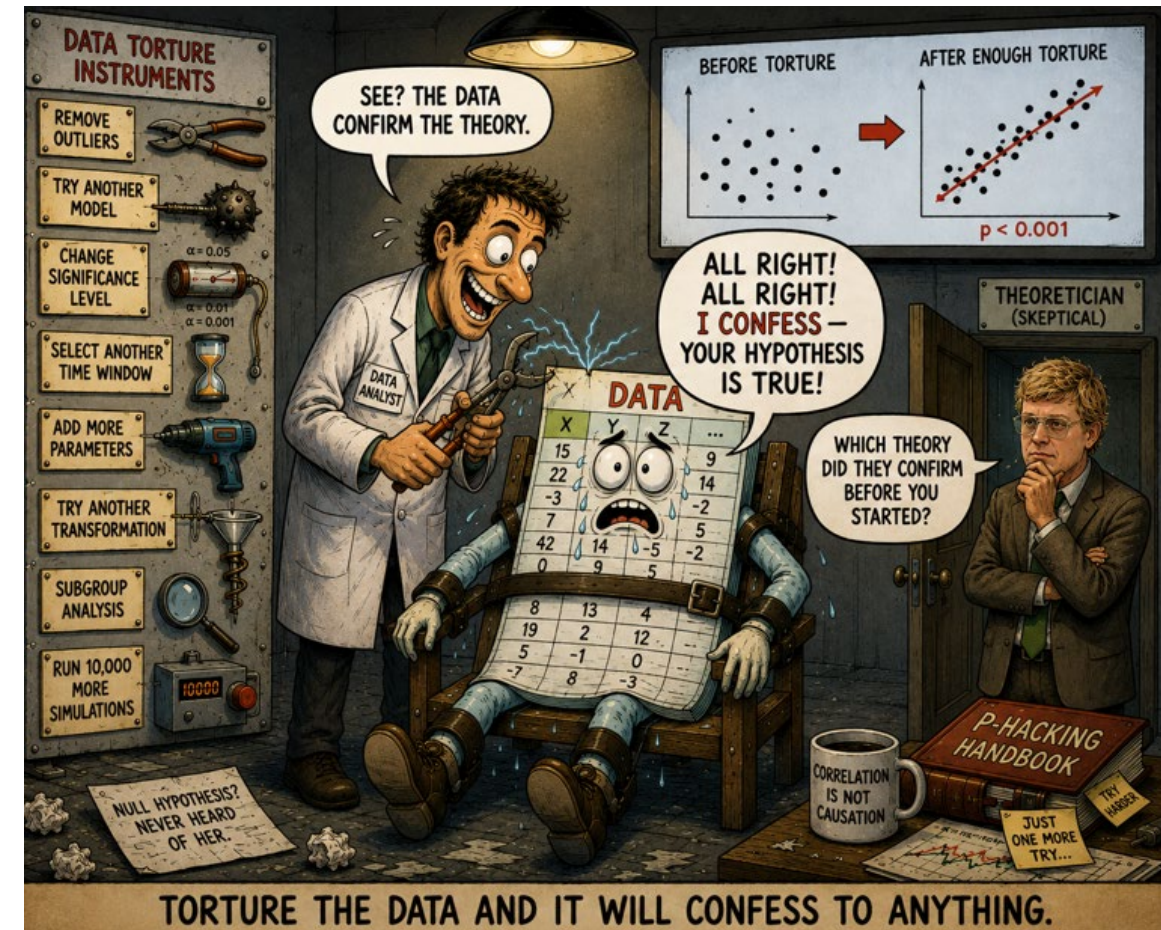
The paper emphasizes Dataset C.

Nothing dishonest occurred.

But with enough datasets, almost every algorithm eventually finds one on which it shines.

The Hidden Risks of Taking Big Data at Face Value

- Big data can make a wrong answer look increasingly convincing, because they reduce the uncertainty while leaving the misconception untouched.
- Big data can postpone the moment at which we discover that we have misunderstood something.
- Big data often test a method where it is already comfortable.
- With big data and analytical flexibility, almost every hypothesis can find supporting evidence



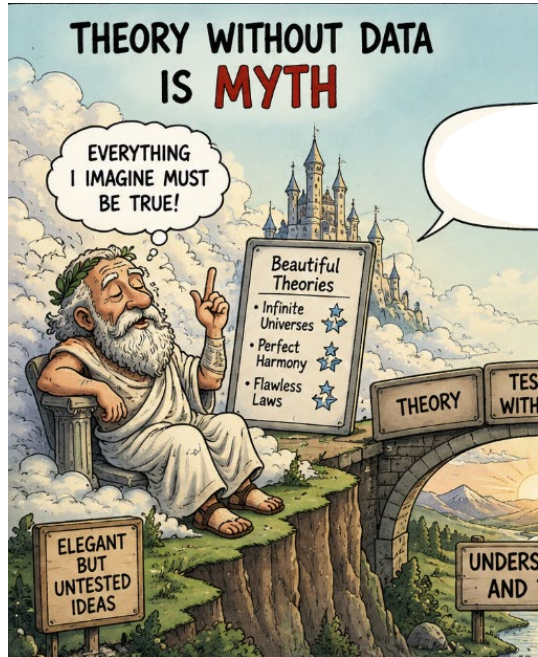
The Hidden Risks of Taking Big Data at Face Value

- Big data can make a wrong answer look increasingly convincing, because they reduce the uncertainty while leaving the misconception untouched.
- Big data can postpone the moment at which we discover that we have misunderstood something.
- Big data often test a method where it is already comfortable. Theory asks where it will become uncomfortable.
- With big data and analytical flexibility, almost every hypothesis can find supporting evidence

When data and computation are abundant, convincing results become easier to produce —but correctness no easier to establish.

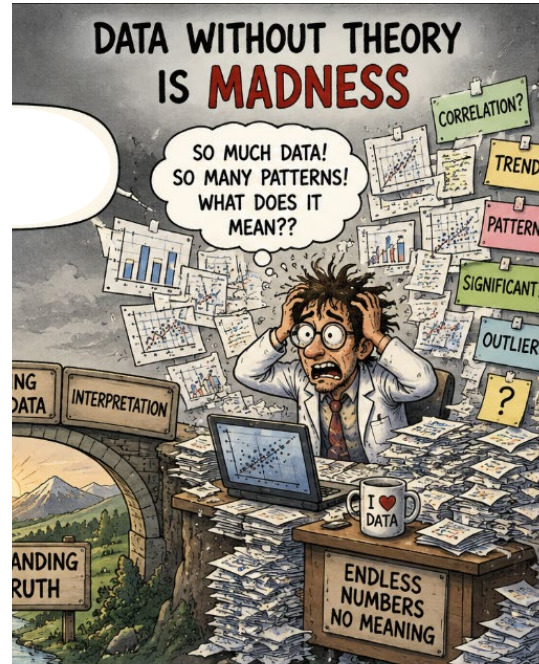
Mathematical understanding therefore becomes more, not less, essential.

Data & Theory



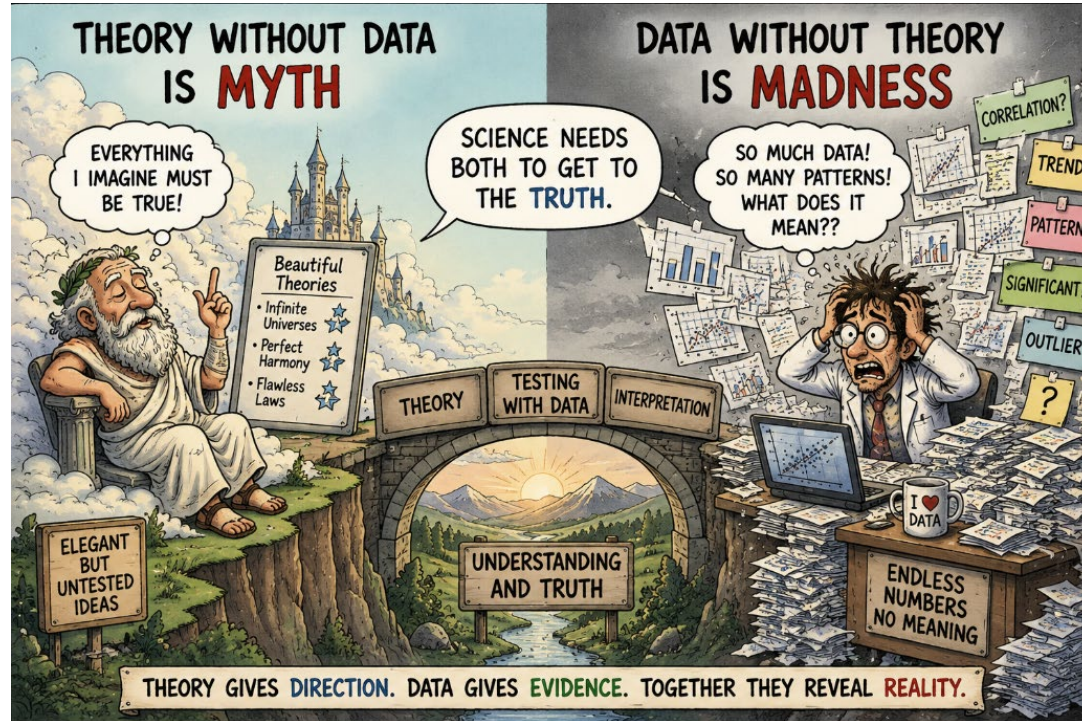
Theory without Data is Myth

Data & Theory



Data without Theory is Madness

Data & Theory



Theory without Data is Myth

Data without Theory is Madness

What Theory Really Contributes

Theory is not about equations; Theory is about understanding.

Data demonstrate success. Theory explains success.

What Theory Really Contributes

Theory is not about equations; Theory is about understanding.

Theory answers the questions that data alone cannot.

1. **Why does it work?** Data tell us what happened. Theory explains why it happened.

Example: Newton did not collect better planetary observations. He explained them.

2. **When does it work?** A method may work today. Theory tells us whether it should also work

- tomorrow,
- on another dataset, or
- under different conditions.

3. **How accurate is it?** Software gives a number. Theory gives its uncertainty, bias, and confidence.

Without theory, a number has no meaning.

4. **How reliable is it?** Two methods may produce the same answer. Theory tells us which one is more trustworthy.

5. **When will it fail?** Data usually reveal failure **after** it occurs. Theory predicts failure **before** it occurs.

6. **Can it be improved?** Theory identifies what limits performance. Data merely reveal performance.

What Theory Really Contributes

Theory is not about equations; Theory is about understanding.

Theory answers the questions that data alone cannot.

1. **Why does it work?** Data tell us what happened. Theory explains why it happened.

Example: Newton did not collect better planetary observations. He explained them.

2. **When does it work?** A method may work today. Theory tells us whether it should also work

- tomorrow,
- on another dataset, or
- under different conditions.

3. **How accurate is it?** Software gives a number. Theory gives its uncertainty, bias, and confidence.

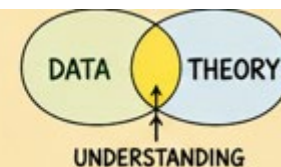
Without theory, a number has no meaning.

4. **How reliable is it?** Two methods may produce the same answer. Theory tells us which one is more trustworthy.

5. **When will it fail?** Data usually reveal failure **after** it occurs. Theory predicts failure **before** it occurs.

6. **Can it be improved?** Theory identifies what limits performance. Data merely reveal performance.

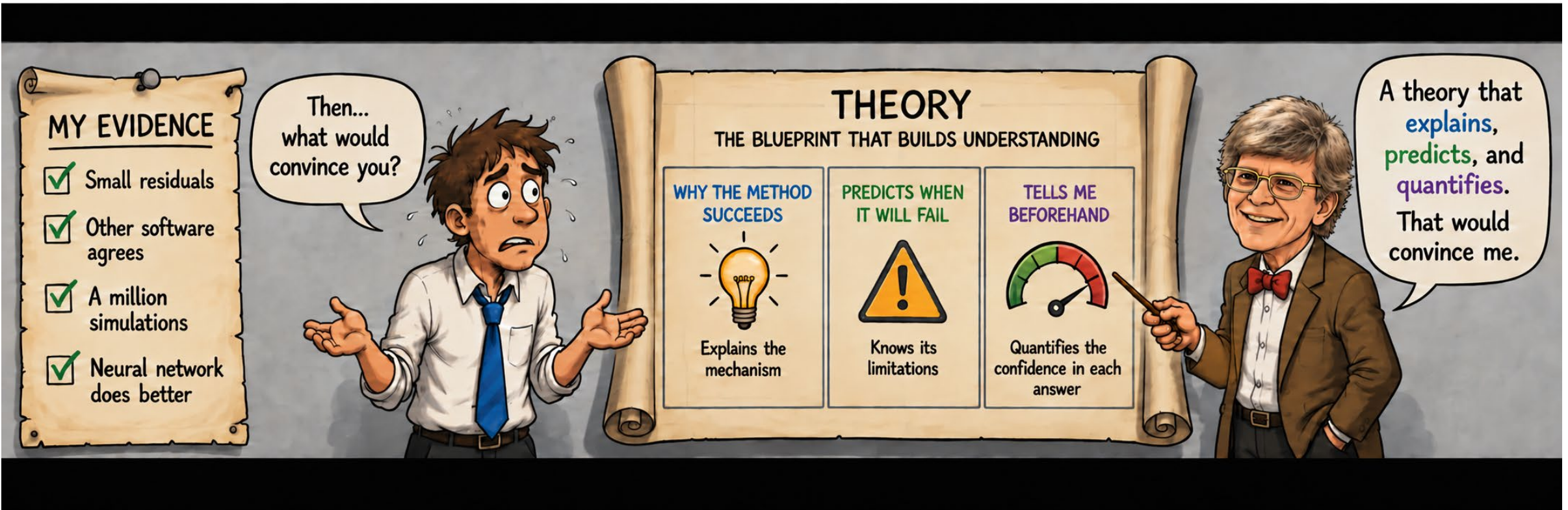
DATA TELL US **WHAT** HAPPENED.
THEORY TELLS US **WHY, WHEN, HOW, AND WHAT IF.**
TOGETHER THEY LEAD TO UNDERSTANDING.



What Theory Really Contributes

Theory is not about equations; Theory is about understanding.

Theory answers the questions that data alone cannot.



Data & Theory

So Holmes was right:

Data! Data! Data!
I can't make bricks without clay!



Data! Data! Data!
I can't make bricks without clay!



Today however, we have more clay than ever & we can manufacture more bricks than ever

But bricks alone do not make cathedrals.

They still need architecture!

That architecture is what mathematical theory provides.

Data! Data! Data!
I can't make bricks without clay!



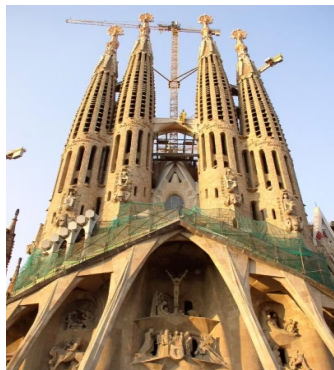
Today however, we have more clay than ever & we can manufacture more bricks than ever

But bricks alone do not make cathedrals.

They still need architecture!

That architecture is what mathematical theory provides.

Theory! Theory! Theory!
I Can't Make a Cathedral with Bricks Alone!



When Success Hides the Need to Understand

Mathematical Geodesy in an Age of Data and Computing Abundance

**Abundance changes the challenge:
from producing results to understanding why they
deserve to be trusted.**

Peter J.G. Teunissen