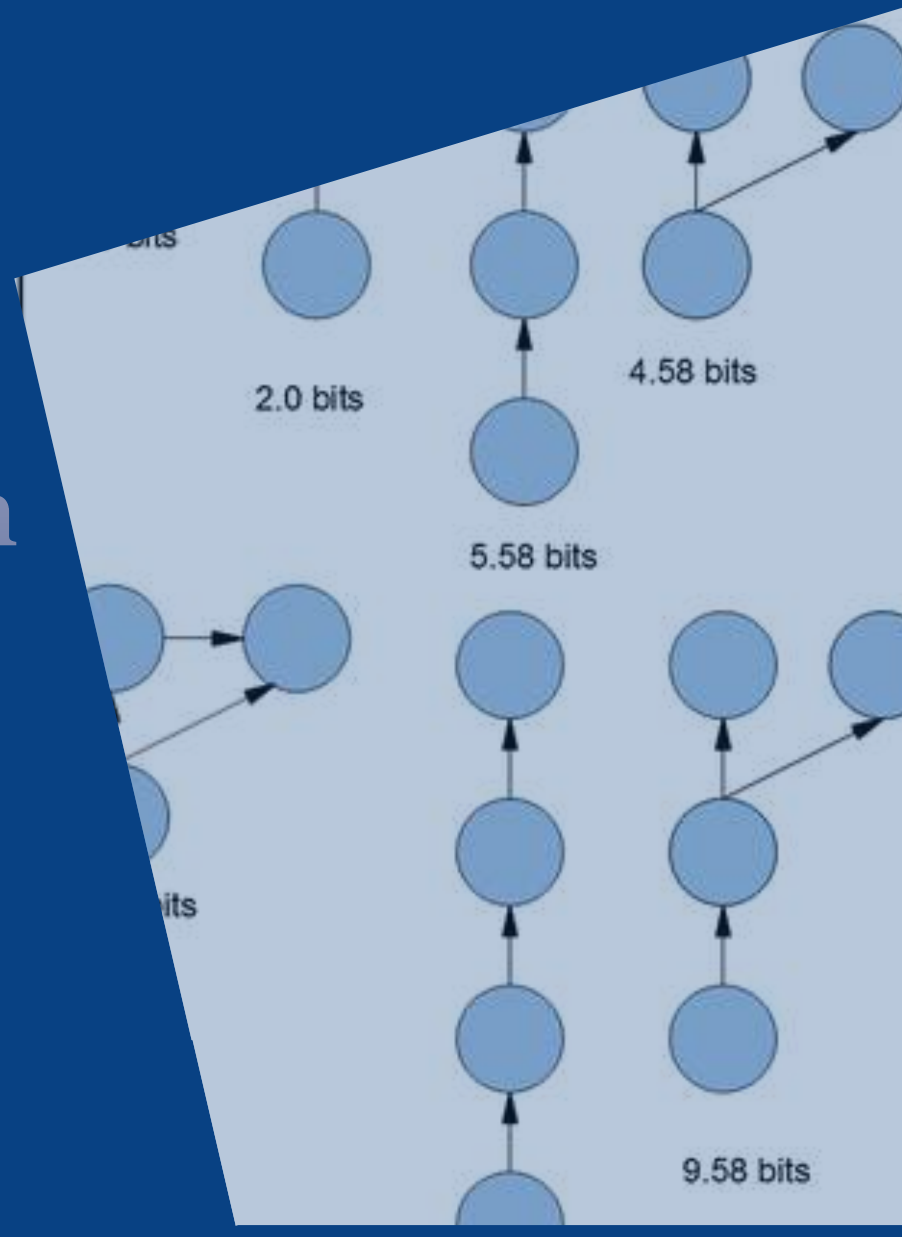
THE UNIVERSITY OF MELBOURNE

Comprehension and Compression

Scientific Understanding, Pattern Recognition, and Kolmogorov Complexity



John S. Wilkins

School of Historical and Philosophical Studies

Topics

1. Understanding in music, business, and science maybe

John S Wilkins john@wilkins.id.au

Topics

1. Understanding in music, business, and science maybe

2. Traditional accounts and recent work

John S Wilkins john@wilkins.id.au

Topics

1. Understanding in music, business, and science maybe
2. Traditional accounts and recent work
3. The mechanics of understanding?

John S Wilkins john@wilkins.id.au

Topics

1. Understanding in music, business, and science maybe
2. Traditional accounts and recent work
3. The mechanics of understanding?
4. Kolmogorov complexity and compression

John S Wilkins john@wilkins.id.au

Topics

1. Understanding in music, business, and science maybe
2. Traditional accounts and recent work
3. The mechanics of understanding?
4. Kolmogorov complexity and compression
5. Subjects

John S Wilkins john@wilkins.id.au

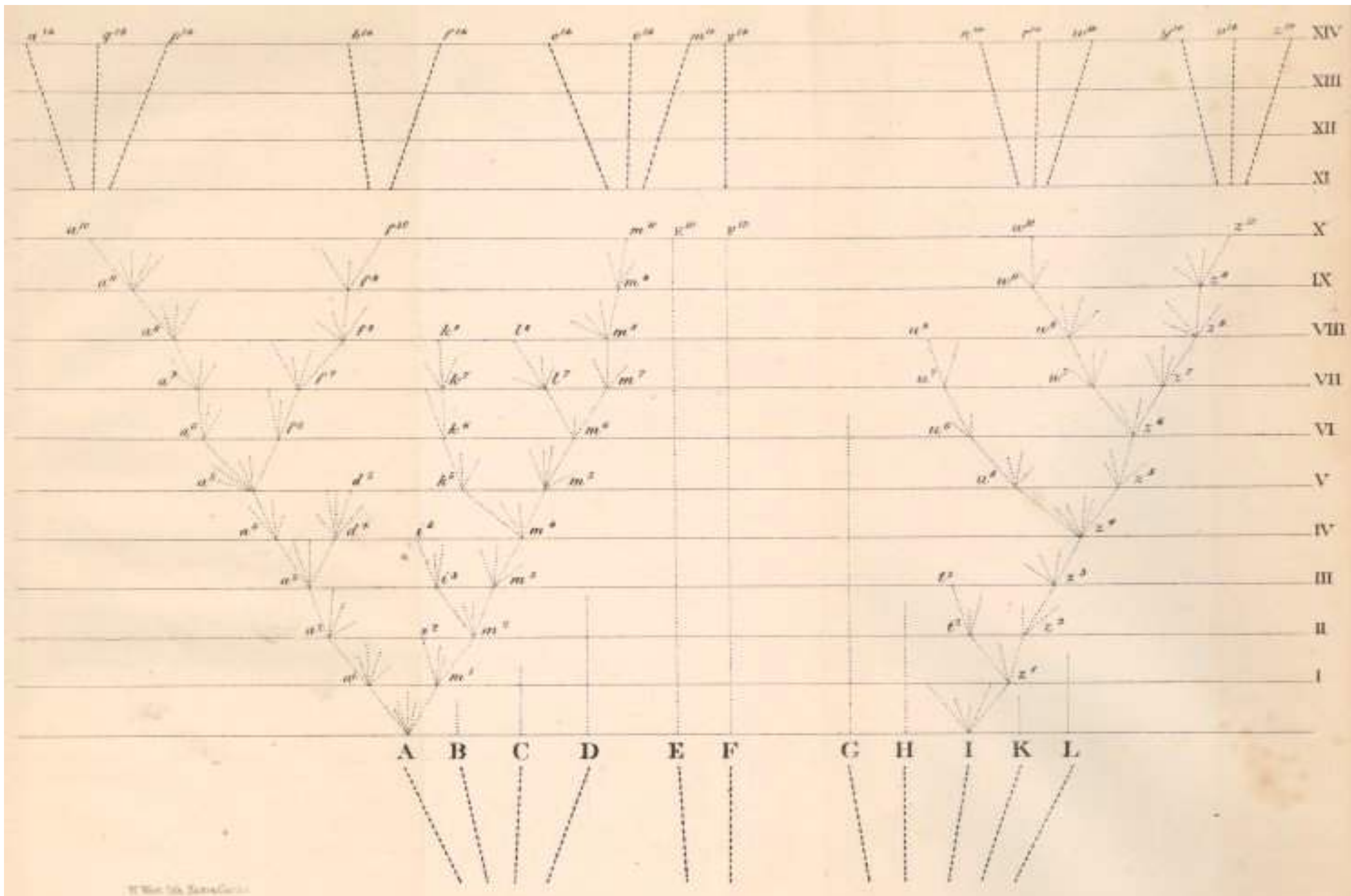
Topics

1. Understanding in music, business, and science maybe
2. Traditional accounts and recent work
3. The mechanics of understanding?
4. Kolmogorov complexity and compression
5. Subjects
6. Handwaving

John S Wilkins john@wilkins.id.au

Understanding perplexing subjects

"The accompanying diagram will aid us in understanding this rather perplexing subject." [Darwin, *Origin*, chapter 4]



John S Wilkins john@wilkins.id.au

Biology and the big data problem

AAGTCAAGCTGCTCTGTGGGCTGTGATCTGCCTCAAACCCACAGCCTGGGTAGCAGG
AGGACCTTGATGCTCCTGGCACAGATGAGGAGAATCTCTCTTTTCTCCTGCTTGAAG
GACAGACATGACTTTGGATTTCCCCAGGAGGAGTTTGGCAACCAGTTCCAAAAGGCT
GAAACCATCCCTGTCCTCCATGAGATGATCCAGCAGATCTTCAATCTCTTCAGCACA
AAGGACTCATCTGCTGCTTGGGATGAGACCCTCCTAGACAAATTCTACACTGAACTC
TACCAGCAGCTGAATGACCTGGAAGCCTGTGTGATACAGGGGGTGGGGGTGACAGAG
ACTCCCCTGATGAAGGAGGACTCCATTCTGGCTGTGAGGAAATACTTCCAAAGAATC
ACTCTCTATCTGAAAGAGAAGAAATACAGCCCTTGTGCCTGGGAGGTTGTCAGAGCA
GAAATCATGAGATCTTTTTCTTTGTCAACAAACTTGCAAGAAAGTTTAAGAAGTAAG
GAATGA, TGTGATCTGCCTCAAACCCACAGCCTGGGTAGCAGGAGGACCTTGATGC
TCCTGGCACAGATGAGGAGAATCTCTCTTTTCTCCTGCTTGAAGGACAGACATGACT
TTGGATTTCCCCAGGAGGAGTTTGGCAACCAGTTCCAAAAGGCTGAAACCATCCCTG
TCCTCCATGAGATGATCCAGCAGATCTTCAATCTCTTCAGCACAAAGGACTCATCTG
CTGCTTGGGATGAGACCCTCCTAGACAAATTCTACACTGAACTCTACCAGCAGCTGA
ATGACCTGGAAGCCTGTGTGATACAGGGGGTGGGGGTGACAGAGACTCCCCTGATGA
AGGAGGACTCCATTCTGGCTGTGAGGAAATACTTCCAAAGAATCACTCTCTATCTGA
AAGAGAAGAAATACAGCCCTTGTGCCTGGGAGGTTGTCAGAGCAGAAATCATGAGAT
CTTTTTCTTTGTCAACAAACTTGCAAGAAAGTTTAAGAAGTAAGGAATGA and

John S Wilkins john@wilkins.id.au

Biology and the big data problem

AAGTCAAGCTGCTCTGTGGGCTGTGATCTGCCTCAAACCCACAGCCTGGGTAGCAGG
AGGACCTTGATGCTCCTGGCACAGATGAGGAGAATCTCTCTTTTCTCCTGCTTGAAG
GACAGACATGACTTTGGATTTCCCCAGGAGGAGTTTGGCAACCAGTTCCAAAAGGCT
GAAACCATCCCTGTCCTCCATGAGATGATCCAGCAGATCTTCAATCTCTTCAGCACA
AAGGACTCATCTGCTGCTTGGGATGAGACCCTCCTAGACAAATTCTACACTGAACTC
TACCAGCAGCTGAATGACCTGGAAGCCTGTGTGATACAGGGGGTGGGGGTGACAGAG
ACTCCCCTGATGAAGGAGGACTCCATTCTGGCTGTGAGGAAATACTTCCAAAGAATC
ACTCTCTATCTGAAAGAGAAGAAATACAGCCCTTGTGCCTGGGAGGTTGTCAGAGCA
GAAATCATGAGATCTTTTTCTTTGTCAACAAACTTGCAAGAAAGTTTAAGAAGTAAG
GAATGA, TGTGATCTGCCTCAAACCCACAGCCTGGGTAGCAGGAGGACCTTGATGC
TCCTGGCACAGATGAGGAGAATCTCTCTTTTCTCCTGCTTGAAGGACAGACATGACT
TTGGATTTCCCCAGGAGGAGTTTGGCAACCAGTTCCAAAAGGCTGAAACCATCCCTG
TCCTCCATGAGATGATCCAGCAGATCTTCAATCTCTTCAGCACAAAGGACTCATCTG
CTGCTTGGGATGAGACCCTCCTAGACAAATTCTACACTGAACTCTACCAGCAGCTGA
ATGACCTGGAAGCCTGTGTGATACAGGGGGTGGGGGTGACAGAGACTCCCCTGATGA
AGGAGGACTCCATTCTGGCTGTGAGGAAATACTTCCAAAGAATCACTCTCTATCTGA
AAGAGAAGAAATACAGCCCTTGTGCCTGGGAGGTTGTCAGAGCAGAAATCATGAGAT
CTTTTTCTTTGTCAACAAACTTGCAAGAAAGTTTAAGAAGTAAGGAATGA and

“It’s human DNA!”

John S Wilkins john@wilkins.id.au

From the data up

"But besides this practical concern, there is a second basic motivation for the scientific quest, namely, man's insatiable intellectual curiosity, his deep concern to *know* the world he lives in, and to *explain*, and thus to *understand*, the unending flow of phenomena it presents to him."

[Carl Hempel 1962]

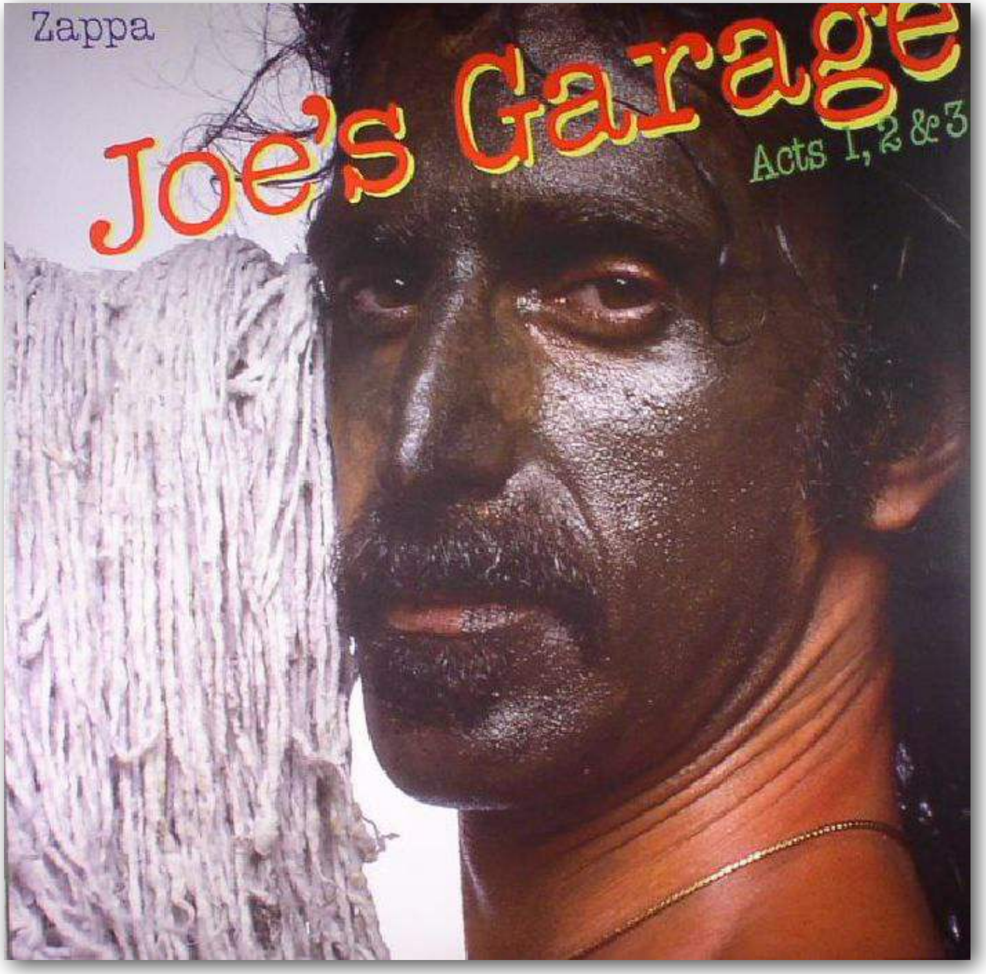
From the data up

"But besides this practical concern, there is a second basic motivation for the scientific quest, namely, man's insatiable intellectual curiosity, his deep concern to *know* the world he lives in, and to *explain*, and thus to *understand*, the unending flow of phenomena it presents to him."

[Carl Hempel 1962]

"Information is not knowledge.
Knowledge is not wisdom.
Wisdom is not truth."

[Zappa 1979]



From the data up

“But besides this practical concern, there is a second basic motivation for the scientific quest, namely, man's insatiable intellectual curiosity, his deep concern to *know* the world he lives in, and to *explain*, and thus to *understand*, the unending flow of phenomena it presents to him.”

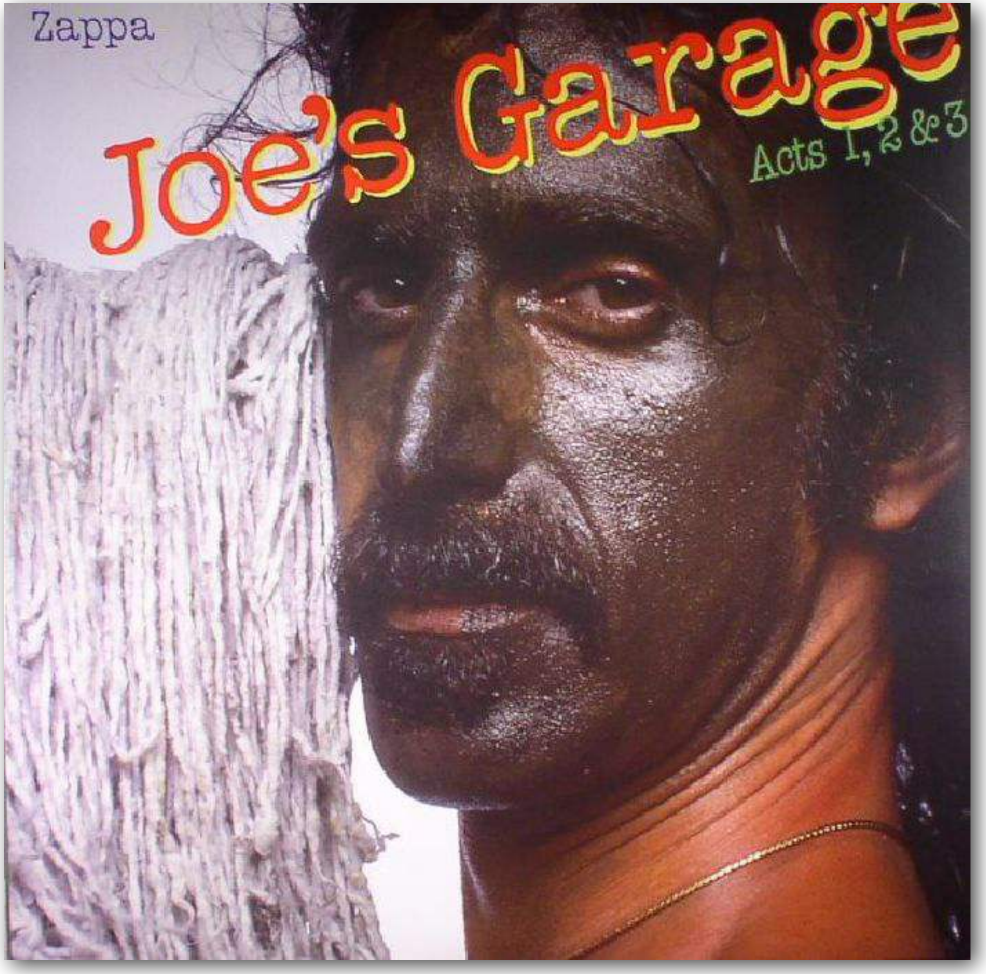
[Carl Hempel 1962]

“Information is not knowledge.
Knowledge is not wisdom.
Wisdom is not truth.”

[Zappa 1979]

“Data is not information, information is not knowledge, knowledge is not wisdom, wisdom is not truth.”

[Robert Royar 1994]



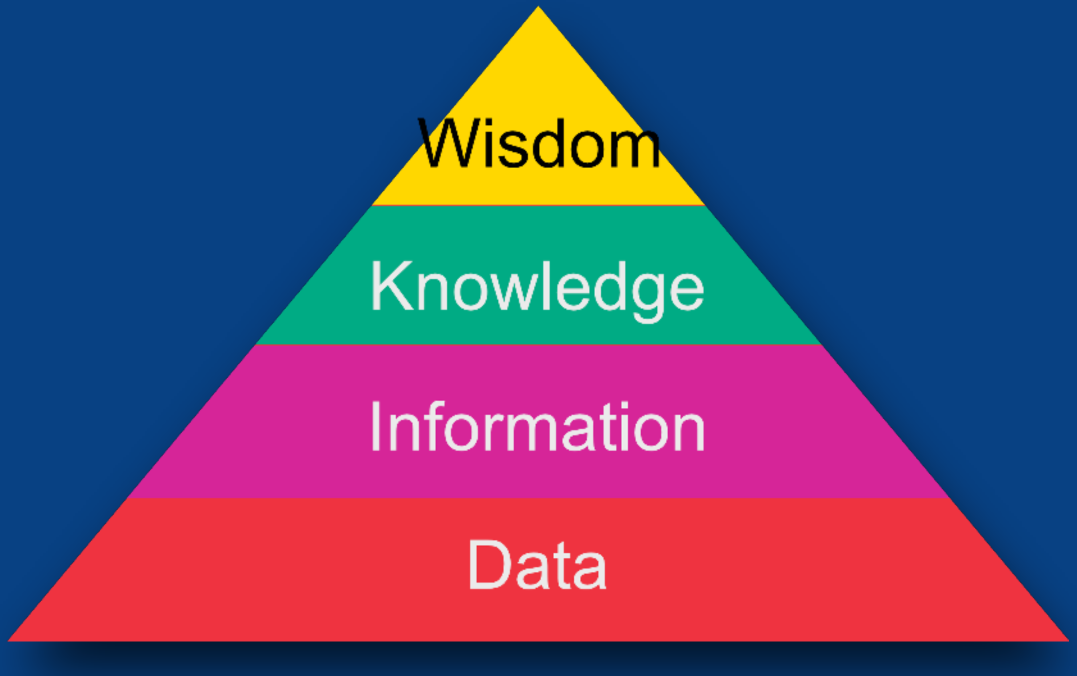
Hempel, Carl G. "Explanation in Science and History," in *Frontiers of Science and Philosophy*, ed. R.C. Colodny, 1962, pp. 9- 19. Pittsburgh: The University of Pittsburgh Press.

Royar, Robert. “New Horizons, Clouded Vistas.” *Computers and Composition* 11, no. 2 (January 1, 1994): 93–105.

Zappa, Frank. “Packard Goose”. 1979. *Joe’s Garage: Acts I, II & III*. FZ Records.

THE UNIVERSITY OF
MELBOURNE

The missing link in the DIKW pyramid

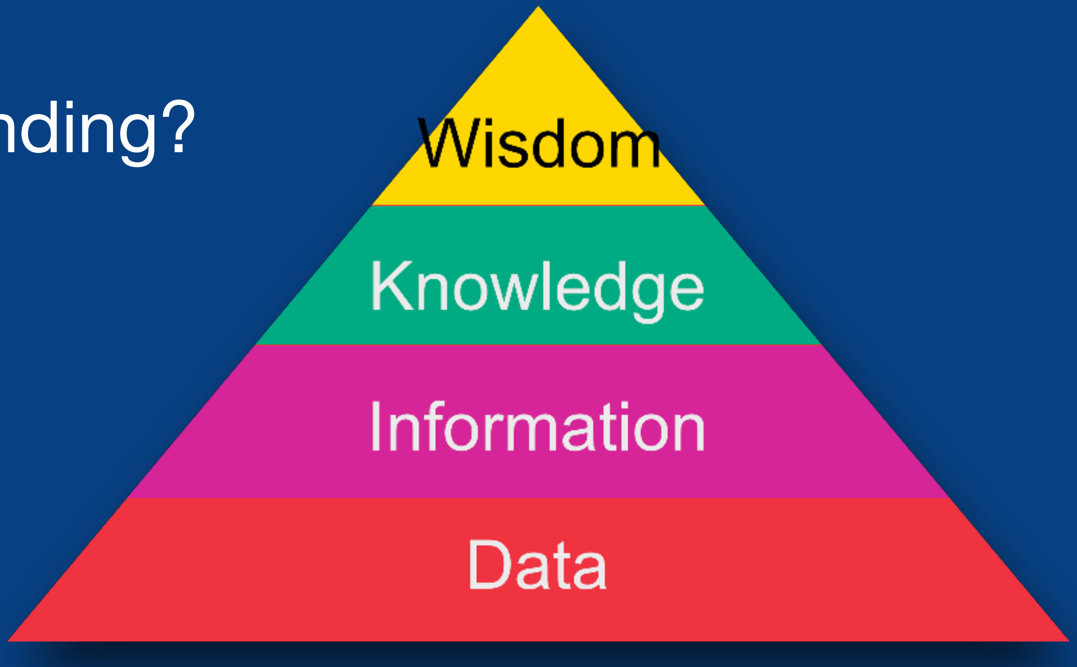


John S Wilkins john@wilkins.id.au

THE UNIVERSITY OF MELBOURNE

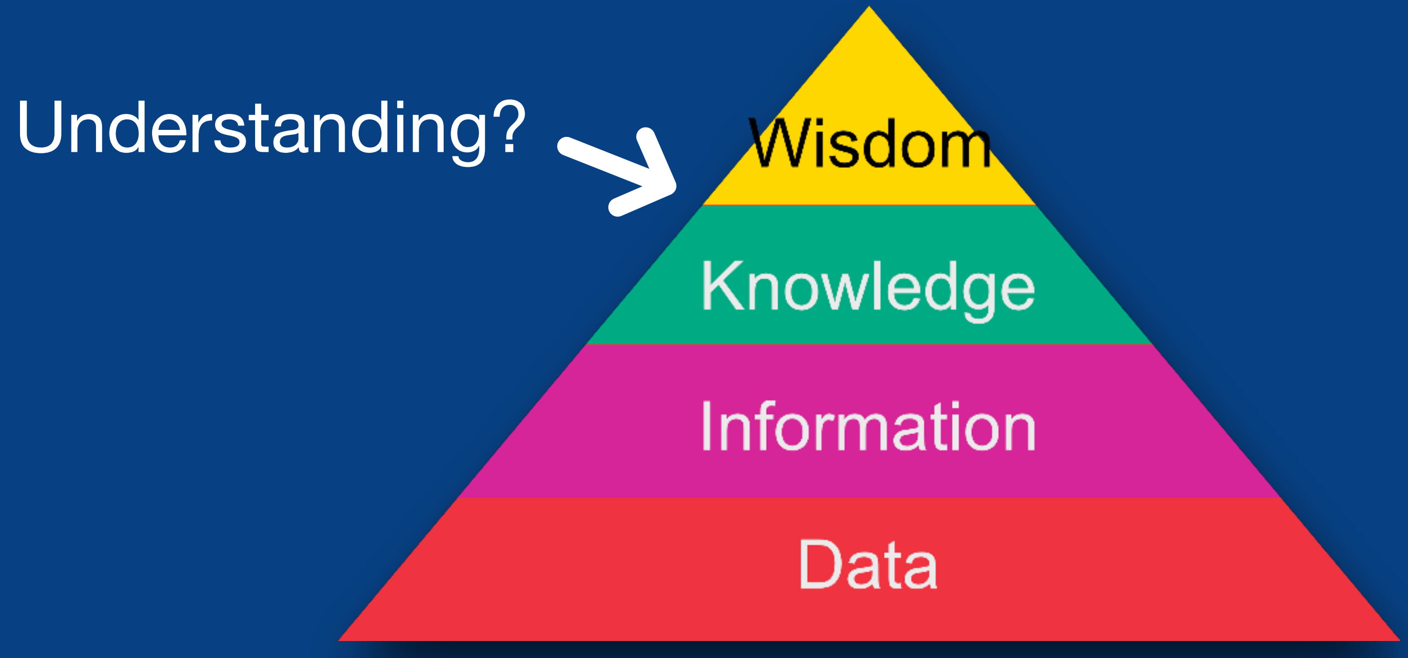
The missing link in the DIKW pyramid

Understanding?

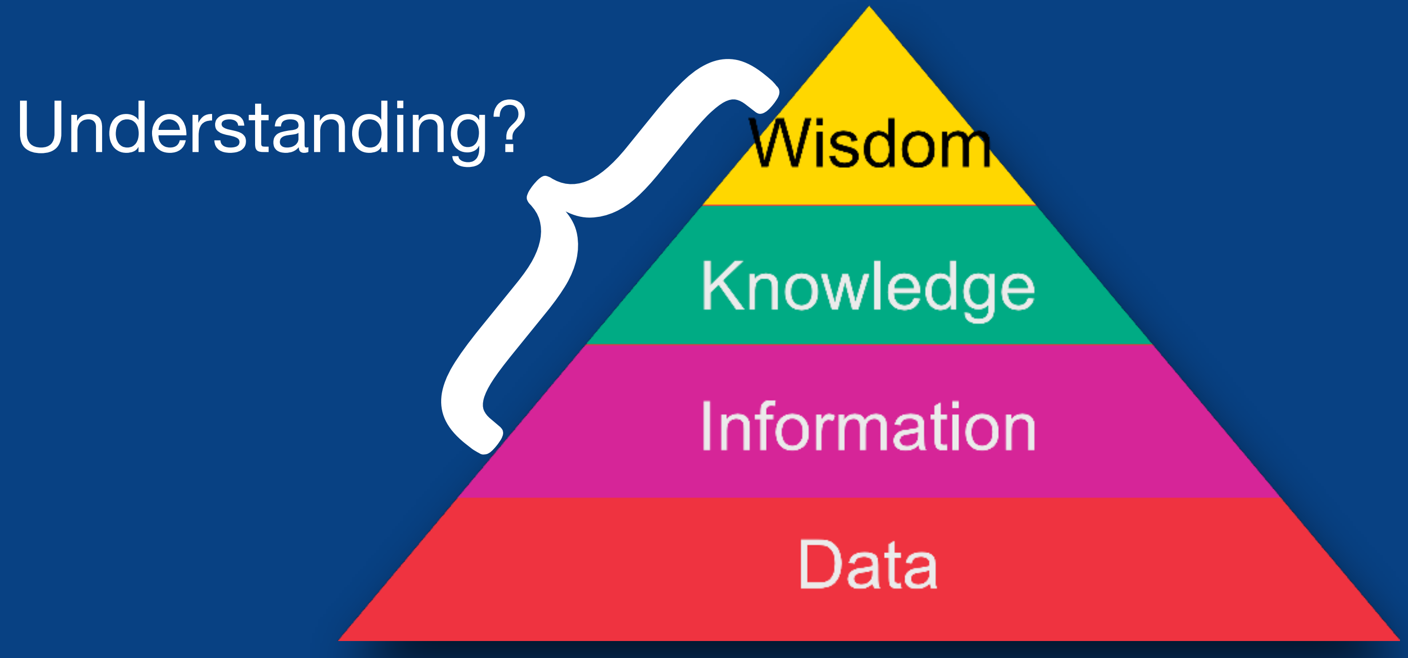


John S Wilkins john@wilkins.id.au

The missing link in the DIKW pyramid



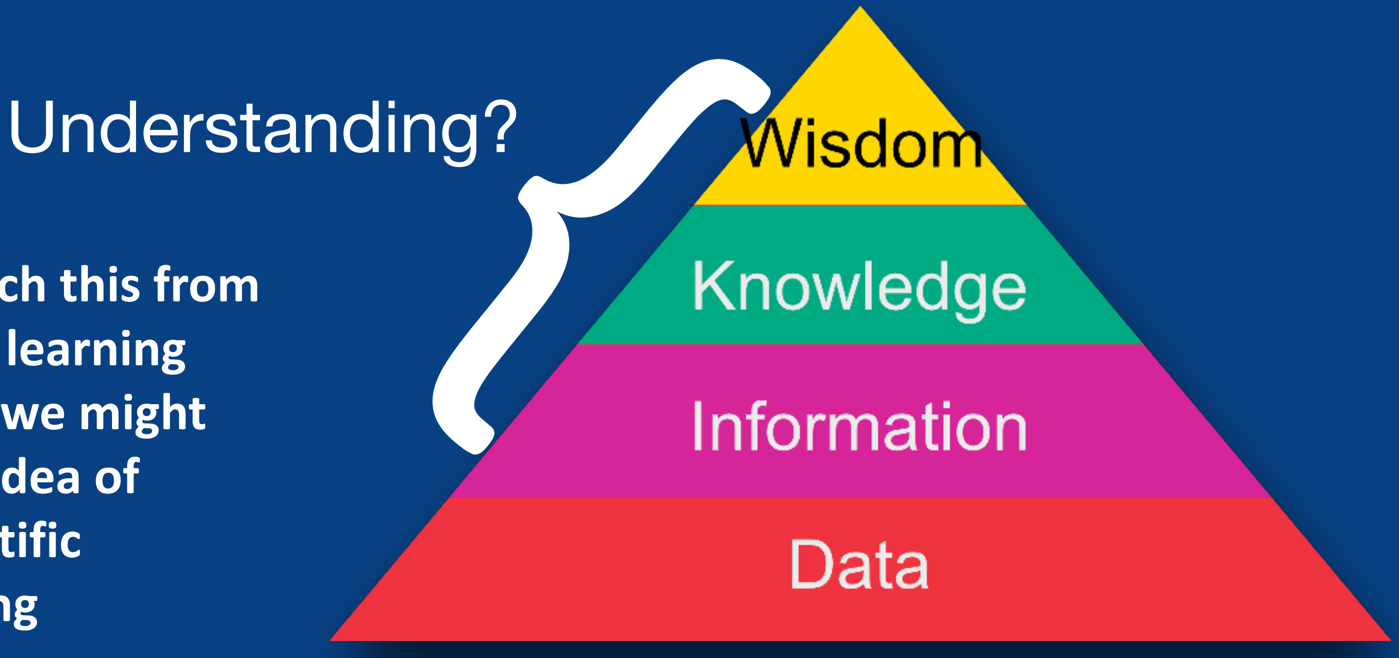
The missing link in the DIKW pyramid



John S Wilkins john@wilkins.id.au

THE UNIVERSITY OF MELBOURNE

The missing link in the DIKW pyramid



If we approach this from the machine learning perspective, we might get a better idea of human scientific understanding

Traditional accounts of understanding

Since Aristotle, understanding has been seen as based on causal explanation (cf. de Regt 2004 for summaries)

John S Wilkins john@wilkins.id.au

Traditional accounts of understanding

Since Aristotle, understanding has been seen as based on causal explanation (cf. de Regt 2004 for summaries)

Hempel and the logical empiricists rejected subjective elements such as feelings of confidence or enlightenment

John S Wilkins john@wilkins.id.au

Traditional accounts of understanding

Since Aristotle, understanding has been seen as based on causal explanation (cf. de Regt 2004 for summaries)

Hempel and the logical empiricists rejected subjective elements such as feelings of confidence or enlightenment

- “Empathic insight and subjective understanding provide no warrant of objective validity, no basis for the systematic prediction or explanation of phenomena…” [Hempel 1965, 163]

John S Wilkins john@wilkins.id.au

Traditional accounts of understanding

Since Aristotle, understanding has been seen as based on causal explanation (cf. de Regt 2004 for summaries)

Hempel and the logical empiricists rejected subjective elements such as feelings of confidence or enlightenment

- "Empathic insight and subjective understanding provide no warrant of objective validity, no basis for the systematic prediction or explanation of phenomena…" [Hempel 1965, 163]

The **objectivist** approach requires ultimately some pragmatist notion:

John S Wilkins john@wilkins.id.au

Traditional accounts of understanding

Since Aristotle, understanding has been seen as based on causal explanation (cf. de Regt 2004 for summaries)

Hempel and the logical empiricists rejected subjective elements such as feelings of confidence or enlightenment

- "Empathic insight and subjective understanding provide no warrant of objective validity, no basis for the systematic prediction or explanation of phenomena…" [Hempel 1965, 163]

The **objectivist** approach requires ultimately some pragmatist notion:

- Of prediction (Trout 2002, Wittgenstein 1953 §152)

John S Wilkins john@wilkins.id.au

Traditional accounts of understanding

Since Aristotle, understanding has been seen as based on causal explanation (cf. de Regt 2004 for summaries)

Hempel and the logical empiricists rejected subjective elements such as feelings of confidence or enlightenment

- "Empathic insight and subjective understanding provide no warrant of objective validity, no basis for the systematic prediction or explanation of phenomena…" [Hempel 1965, 163]

The **objectivist** approach requires ultimately some pragmatist notion:

- Of prediction (Trout 2002, Wittgenstein 1953 §152)
- Of explanation (van Fraassen 1980)

John S Wilkins john@wilkins.id.au

Traditional accounts of understanding

Since Aristotle, understanding has been seen as based on causal explanation (cf. de Regt 2004 for summaries)

Hempel and the logical empiricists rejected subjective elements such as feelings of confidence or enlightenment

- "Empathic insight and subjective understanding provide no warrant of objective validity, no basis for the systematic prediction or explanation of phenomena…" [Hempel 1965, 163]

The **objectivist** approach requires ultimately some pragmatist notion:

- Of prediction (Trout 2002, Wittgenstein 1953 §152)
- Of explanation (van Fraassen 1980)

The utility of causal accounts ultimately constitutes understanding.

John S Wilkins john@wilkins.id.au

Traditional accounts of understanding

Since Aristotle, understanding has been seen as based on causal explanation (cf. de Regt 2004 for summaries)

Hempel and the logical empiricists rejected subjective elements such as feelings of confidence or enlightenment

- "Empathic insight and subjective understanding provide no warrant of objective validity, no basis for the systematic prediction or explanation of phenomena…" [Hempel 1965, 163]

The **objectivist** approach requires ultimately some pragmatist notion:

- Of prediction (Trout 2002, Wittgenstein 1953 §152)
- Of explanation (van Fraassen 1980)

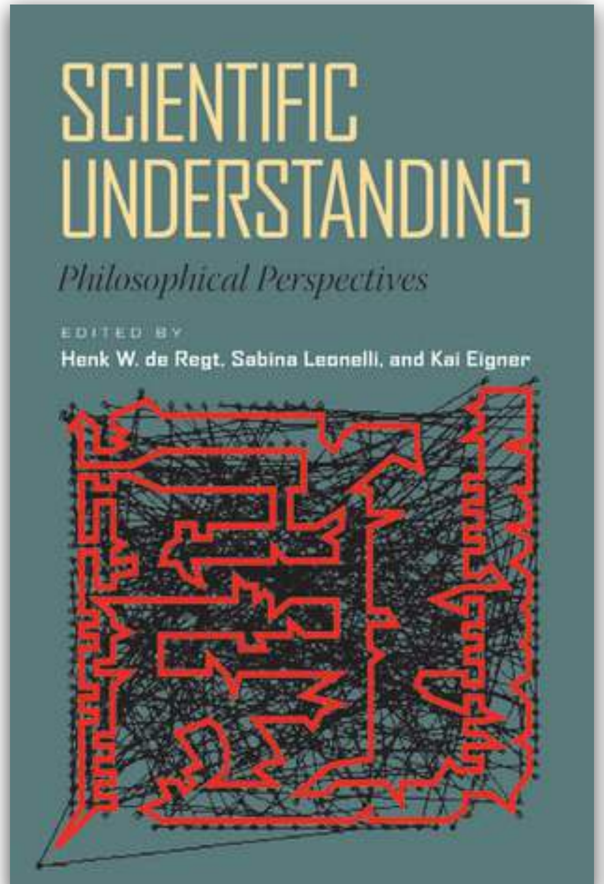
The utility of causal accounts ultimately constitutes understanding.

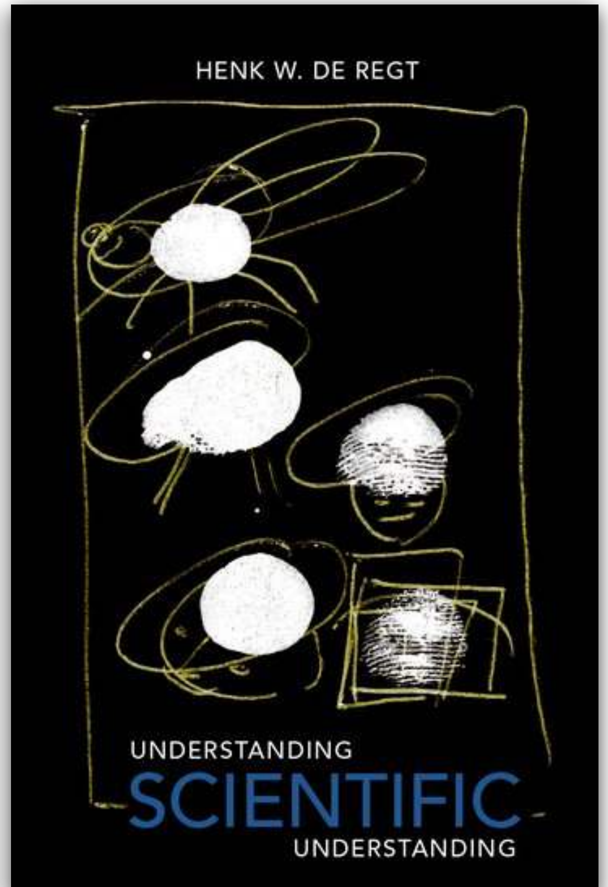
Subjectivist or phenomenological accounts of understanding are merely **psychologistic** on this approach.

Van Fraassen, Bas C. *The Scientific Image*. Oxford: Clarendon Press, 1980.
Hempel, Carl G. *Aspects of Scientific Explanation, and Other Essays in the Philosophy of Science*. New York: The Free Press, 1965.
Regt, Henk W. de. "Discussion Note: Making Sense of Understanding." *Philosophy of Science* 71, no. 1 (January 1, 2004): 98–109.
Trout, J. D. "Scientific Explanation and the Sense of Understanding." *Philosophy of Science* 69, no. 2 (June 1, 2002): 212–33.

John S Wilkins john@wilkins.id.au

Recent work on understanding

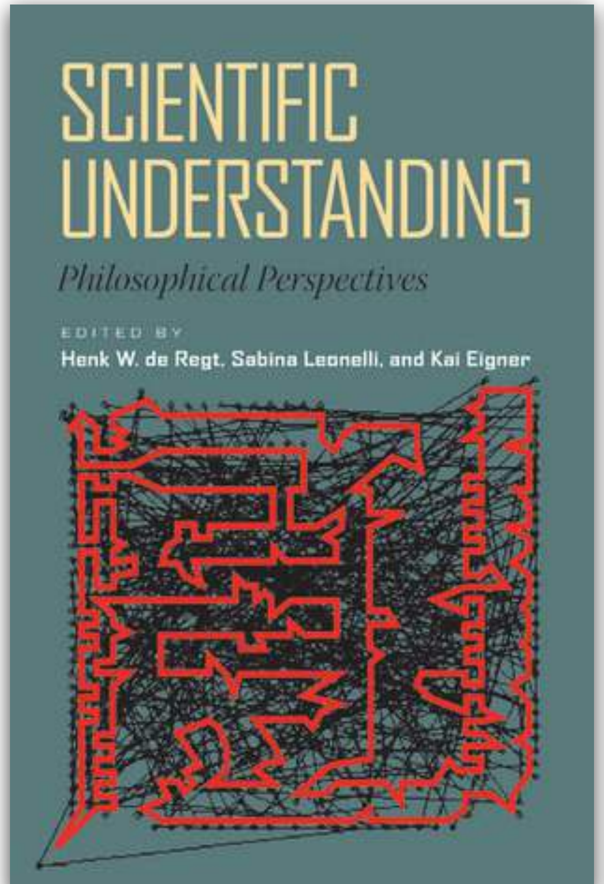


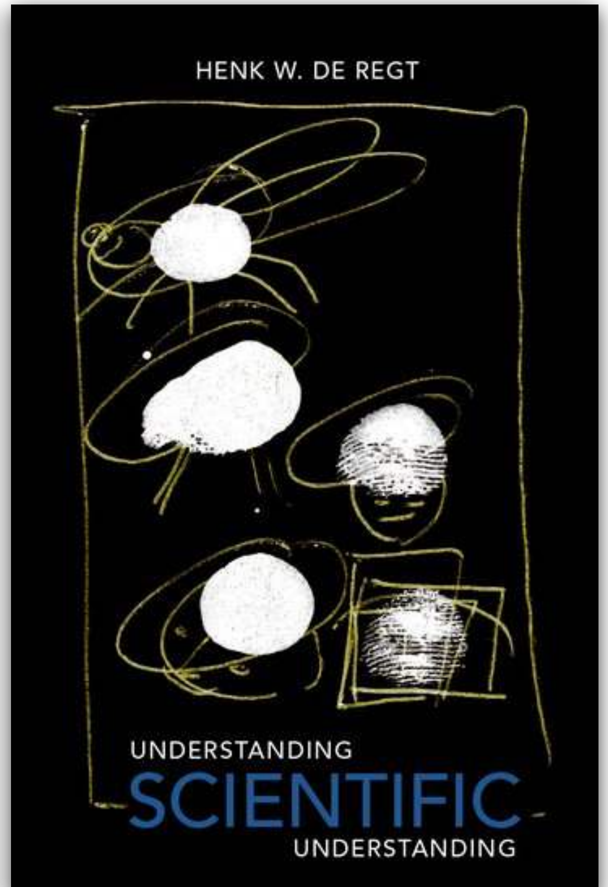


Henk de Regt and his collaborators have developed an account of understanding as:

John S Wilkins john@wilkins.id.au

Recent work on understanding

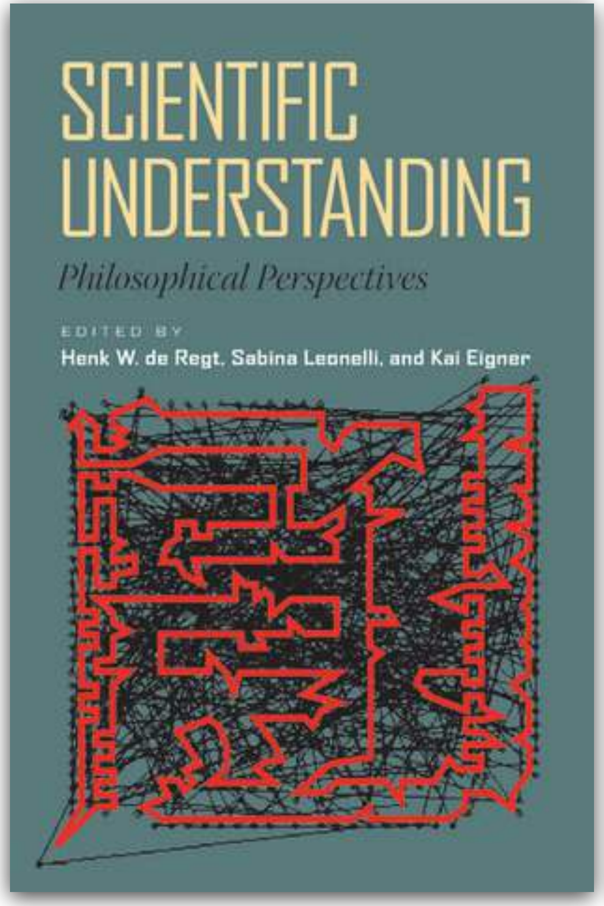


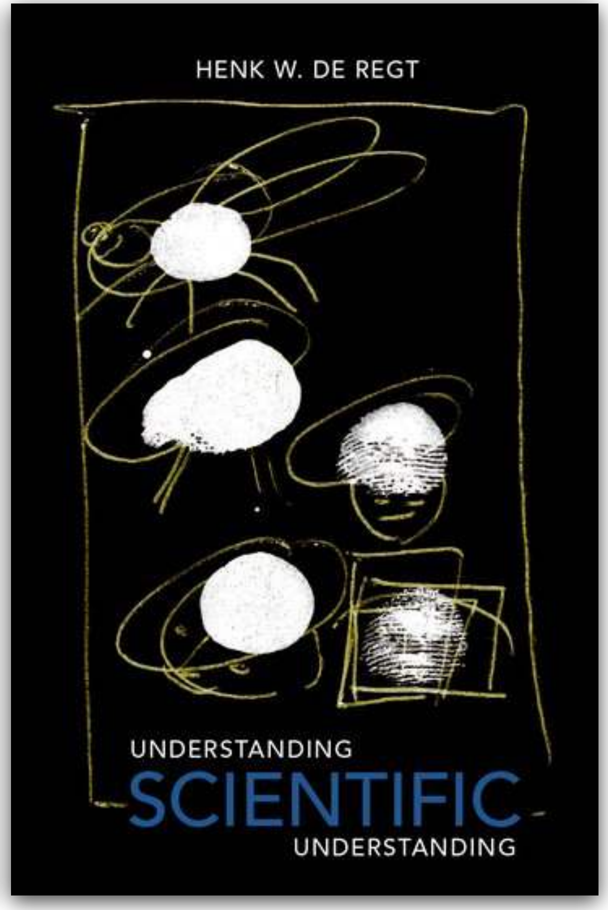


Henk de Regt and his collaborators have developed an account of understanding as:

- Intelligibility

Recent work on understanding

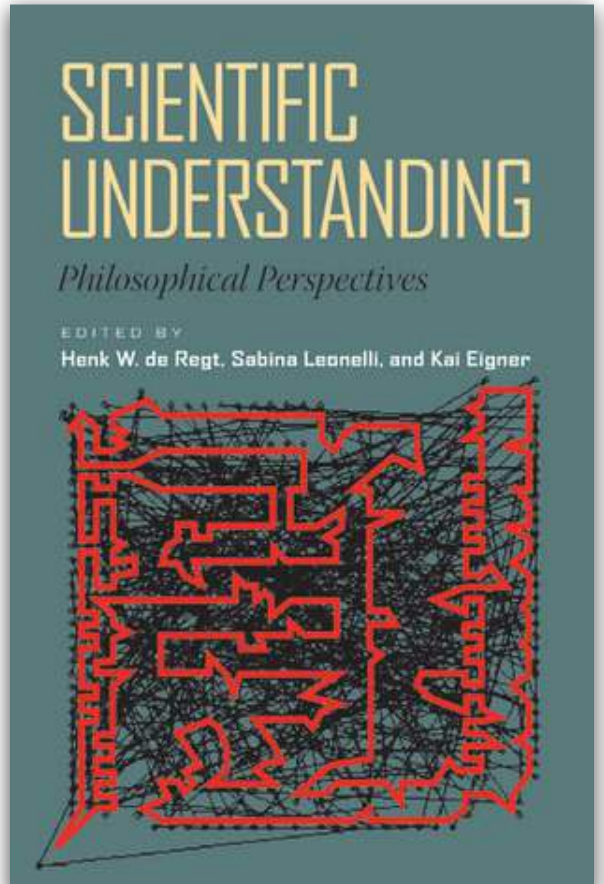


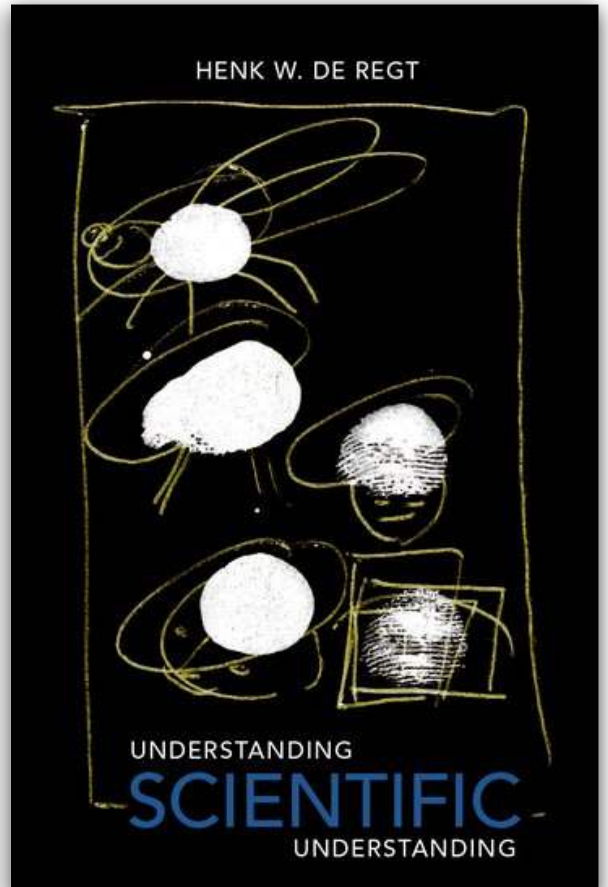


Henk de Regt and his collaborators have developed an account of understanding as:

- Intelligibility
- Empirical adequacy

Recent work on understanding

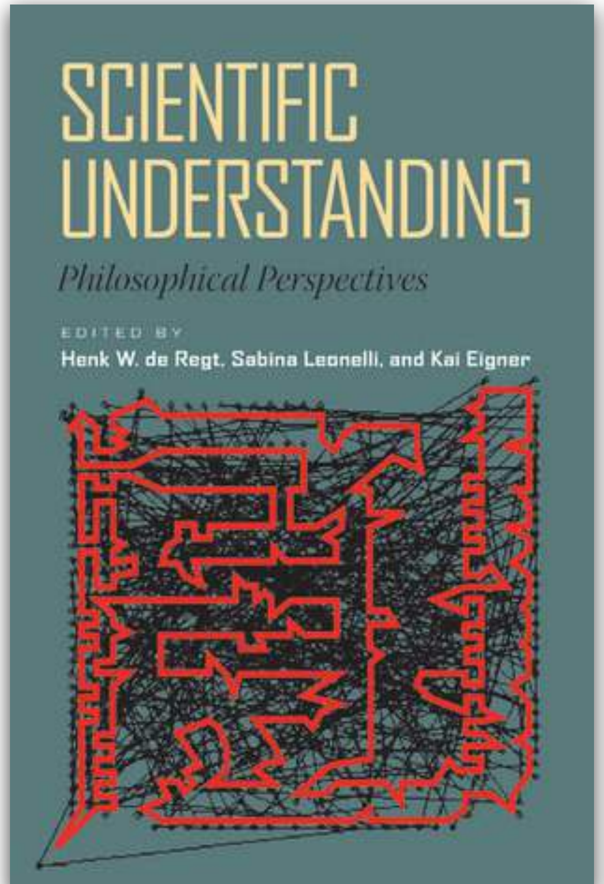


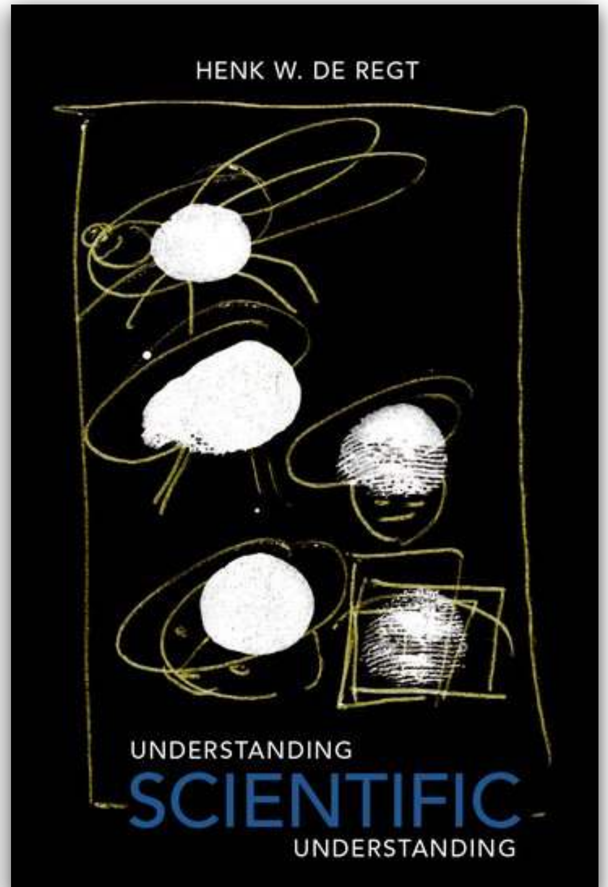


Henk de Regt and his collaborators have developed an account of understanding as:

- Intelligibility
- Empirical adequacy
- Consistency

Recent work on understanding





Henk de Regt and his collaborators have developed an account of understanding as:

- Intelligibility
- Empirical adequacy
- Consistency

These are contextual features of *disciplinary* or *professional* understanding, without reference to subjects

De Regt, Henk W. *Understanding Scientific Understanding.* Oxford Studies in Philosophy of Science. Oxford University Press, 2017.

De Regt, Henk W., Sabine Leonelli, and Kai Eigner. *Scientific Understanding: Philosophical Perspectives*. Pittsburgh: University of Pittsburgh Press, 2009.

John S Wilkins john@wilkins.id.au

Recent work on understanding

A scientific theory *T* (in one or more of its representations) is intelligible for scientists (in context *C*) if they can recognize qualitatively characteristic consequences of *T* without performing exact calculations.

Henk de Regt and his collaborators have developed an account of understanding as:

- Intelligibility
- Empirical adequacy
- Consistency

These are contextual features of *disciplinary* or *professional* understanding, without reference to subjects

De Regt, Henk W. *Understanding Scientific Understanding.* Oxford Studies in Philosophy of Science. Oxford University Press, 2017.

De Regt, Henk W., Sabine Leonelli, and Kai Eigner. *Scientific Understanding: Philosophical Perspectives*. Pittsburgh: University of Pittsburgh Press, 2009.

John S Wilkins john@wilkins.id.au

Formal mechanics

The *mechanism* of understanding is not generally considered in epistemology

John S Wilkins john@wilkins.id.au

Formal mechanics

The *mechanism* of understanding is not generally considered in epistemology

- Because we do not have the requisite neurobiological or neuropsychological knowledge yet (we do not understand understanding)

John S Wilkins john@wilkins.id.au

Formal mechanics

The *mechanism* of understanding is not generally considered in epistemology

- Because we do not have the requisite neurobiological or neuropsychological knowledge yet (we do not understand understanding)
- But also because there are objections to subjectivism in psychological approaches (I concur)

John S Wilkins john@wilkins.id.au

Formal mechanics

The *mechanism* of understanding is not generally considered in epistemology

- Because we do not have the requisite neurobiological or neuropsychological knowledge yet (we do not understand understanding)
- But also because there are objections to subjectivism in psychological approaches (I concur)

How to approach this?

John S Wilkins john@wilkins.id.au

Formal mechanics

The *mechanism* of understanding is not generally considered in epistemology

- Because we do not have the requisite neurobiological or neuropsychological knowledge yet (we do not understand understanding)
- But also because there are objections to subjectivism in psychological approaches (I concur)

How to approach this?

One domain in which such matters are dealt with mechanically (i.e., mechanisms are sought) is *machine learning* (ML, a subset of AI).

John S Wilkins john@wilkins.id.au

Formal mechanics

The *mechanism* of understanding is not generally considered in epistemology

- Because we do not have the requisite neurobiological or neuropsychological knowledge yet (we do not understand understanding)
- But also because there are objections to subjectivism in psychological approaches (I concur)

How to approach this?

One domain in which such matters are dealt with mechanically (i.e., mechanisms are sought) is *machine learning* (ML, a subset of AI).

- The mechanisms can be opaque “black boxes” in their operation (deep learning)

John S Wilkins john@wilkins.id.au

Formal mechanics

The *mechanism* of understanding is not generally considered in epistemology

- Because we do not have the requisite neurobiological or neuropsychological knowledge yet (we do not understand understanding)
- But also because there are objections to subjectivism in psychological approaches (I concur)

How to approach this?

One domain in which such matters are dealt with mechanically (i.e., mechanisms are sought) is *machine learning* (ML, a subset of AI).

- The mechanisms can be opaque "black boxes" in their operation (deep learning)
- There is no subjectivity required (but there is a "subject"; that is, a *learning system*)

John S Wilkins john@wilkins.id.au

Formal mechanics

The *mechanism* of understanding is not generally considered in epistemology

- Because we do not have the requisite neurobiological or neuropsychological knowledge yet (we do not understand understanding)
- But also because there are objections to subjectivism in psychological approaches (I concur)

How to approach this?

One domain in which such matters are dealt with mechanically (i.e., mechanisms are sought) is *machine learning* (ML, a subset of AI).

- The mechanisms can be opaque “black boxes” in their operation (deep learning)
- There is no subjectivity required (but there is a “subject”; that is, a *learning system*)

I will attempt to generalise ML and algorithmic information theoretic tools to apply to this problem of understanding within *knowing systems*

John S Wilkins john@wilkins.id.au

Kinematic versus dynamic

In physics there is a distinction between two kinds of description:

John S Wilkins john@wilkins.id.au

Kinematic versus dynamic

In physics there is a distinction between two kinds of description:

- **Kinematics**: *description of motions* without consideration of forces

John S Wilkins john@wilkins.id.au

Kinematic versus dynamic

In physics there is a distinction between two kinds of description:

- **Kinematics**: *description of motions* without consideration of forces
- **Dynamics**: the relations between *forces* and *motion*

Dunamis is an ordinary Greek word for possibility or capability. Depending on context, it could be translated "potency", "potential", "capacity", "ability", "power", "capability", "strength", "possibility", "force" [Wikipedia]

John S Wilkins john@wilkins.id.au

Kinematic versus dynamic

In physics there is a distinction between two kinds of description:

- **Kinematics**: *description of motions* without consideration of forces
- **Dynamics**: the relations between *forces* and *motion*

In brief: *phenomenal* accounts and *causal* explanations.

Dunamis is an ordinary Greek word for possibility or capability. Depending on context, it could be translated "potency", "potential", "capacity", "ability", "power", "capability", "strength", "possibility", "force" [Wikipedia]

John S Wilkins john@wilkins.id.au

Kinematic versus dynamic

In physics there is a distinction between two kinds of description:

- **Kinematics**: *description of motions* without consideration of forces
- **Dynamics**: the relations between *forces* and *motion*

In brief: *phenomenal* accounts and *causal* explanations.

We can explore the problem of understanding in knowing systems as the shift from *kinematic accounts* (what the behaviour of the system that understands is) to *dynamic explanations* (what forces the behaviour of the understanding system)

Dunamis is an ordinary Greek word for possibility or capability. Depending on context, it could be translated "potency", "potential", "capacity", "ability", "power", "capability", "strength", "possibility", "force" [Wikipedia]

John S Wilkins john@wilkins.id.au

Kinematic versus dynamic

In physics there is a distinction between two kinds of description:

- **Kinematics**: *description of motions* without consideration of forces
- **Dynamics**: the relations between *forces* and *motion*

In brief: *phenomenal* accounts and *causal* explanations.

We can explore the problem of understanding in knowing systems as the shift from *kinematic accounts* (what the behaviour of the system that understands is) to *dynamic explanations* (what forces the behaviour of the understanding system)

- Of systems that are the *subject* of understanding

Dunamis is an ordinary Greek word for possibility or capability. Depending on context, it could be translated "potency", "potential", "capacity", "ability", "power", "capability", "strength", "possibility", "force" [Wikipedia]

John S Wilkins john@wilkins.id.au

Kinematic versus dynamic

In physics there is a distinction between two kinds of description:

- **Kinematics**: *description of motions* without consideration of forces
- **Dynamics**: the relations between *forces* and *motion*

In brief: *phenomenal* accounts and *causal* explanations.

We can explore the problem of understanding in knowing systems as the shift from *kinematic accounts* (what the behaviour of the system that understands is) to *dynamic explanations* (what forces the behaviour of the understanding system)

- Of systems that are the *subject* of understanding
- Of systems that *understand*

Dunamis is an ordinary Greek word for possibility or capability. Depending on context, it could be translated "potency", "potential", "capacity", "ability", "power", "capability", "strength", "possibility", "force" [Wikipedia]

John S Wilkins john@wilkins.id.au

Kinematic versus dynamic

In physics there is a distinction between two kinds of description:

- **Kinematics**: *description of motions* without consideration of forces
- **Dynamics**: the relations between *forces* and *motion*

In brief: *phenomenal* accounts and *causal* explanations.

We can explore the problem of understanding in knowing systems as the shift from *kinematic accounts* (what the behaviour of the system that understands is) to *dynamic explanations* (what forces the behaviour of the understanding system)

- Of systems that are the *subject* of understanding
- Of systems that *understand*

These coincide: as we move from kinematic descriptions of things to dynamic explanations of knowing systems, we also move from considering knowledge to considering understanding

Dunamis is an ordinary Greek word for possibility or capability. Depending on context, it could be translated "potency", "potential", "capacity", "ability", "power", "capability", "strength", "possibility", "force" [Wikipedia]

John S Wilkins john@wilkins.id.au

Pattern recognition

Knowing systems have a limit to their **working memory** under realistic assumptions

John S Wilkins john@wilkins.id.au

Pattern recognition

Knowing systems have a limit to their **working memory** under realistic assumptions

Knowing systems also have to deal with noisy data and uncertainty

John S Wilkins john@wilkins.id.au

Pattern recognition

Knowing systems have a limit to their **working memory** under realistic assumptions

Knowing systems also have to deal with noisy data and uncertainty

Once the scale of the data set exceeds the capacity of the system's cognitive limits [working memory, processing power] pattern recognition must be off-sourced

John S Wilkins john@wilkins.id.au

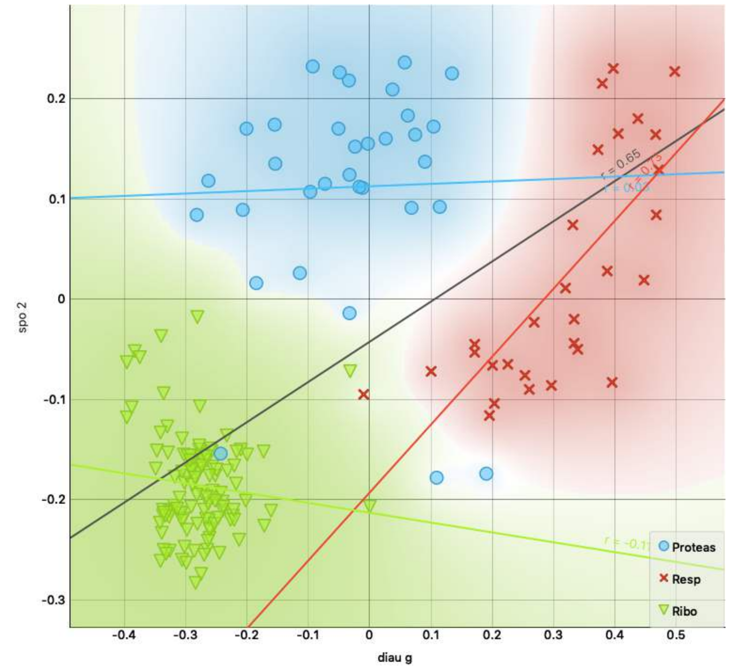
Pattern recognition

Knowing systems have a limit to their **working memory** under realistic assumptions

Knowing systems also have to deal with noisy data and uncertainty

Once the scale of the data set exceeds the capacity of the system's cognitive limits [working memory, processing power] pattern recognition must be off-sourced

- E.g., regression and other statistical analyses



John S Wilkins john@wilkins.id.au

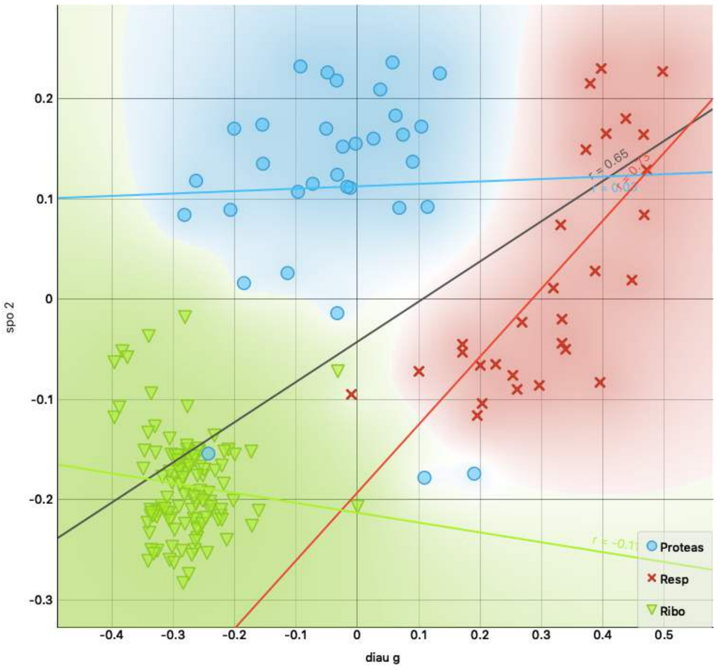
Pattern recognition

Knowing systems have a limit to their **working memory** under realistic assumptions

Knowing systems also have to deal with noisy data and uncertainty

Once the scale of the data set exceeds the capacity of the system's cognitive limits [working memory, processing power] pattern recognition must be off-sourced

- E.g., regression and other statistical analyses
- Machine Learning classifications (ANNs)



John S Wilkins john@wilkins.id.au

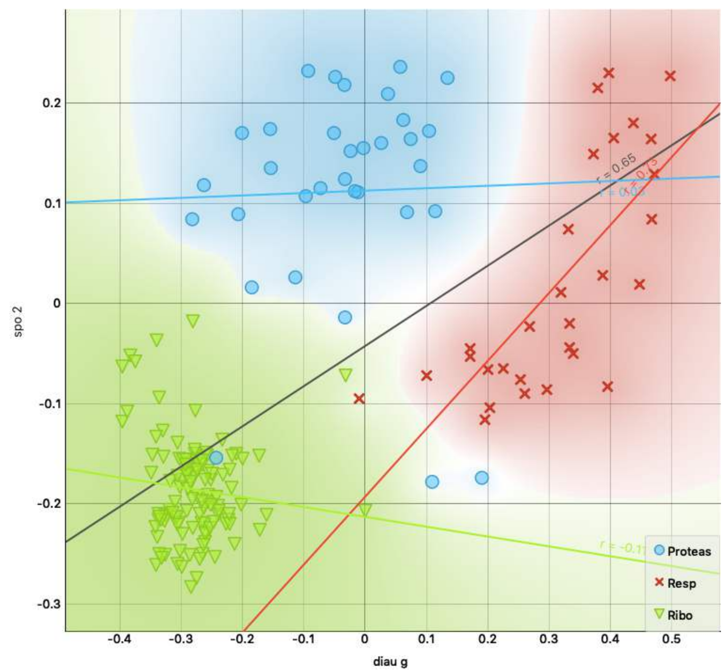
Pattern recognition

Knowing systems have a limit to their **working memory** under realistic assumptions

Knowing systems also have to deal with noisy data and uncertainty

Once the scale of the data set exceeds the capacity of the system's cognitive limits [working memory, processing power] pattern recognition must be off-sourced

- E.g., regression and other statistical analyses
- Machine Learning classifications (ANNs)
- Distance analyses (Hamming, Manhattan)



John S Wilkins john@wilkins.id.au

Pattern recognition

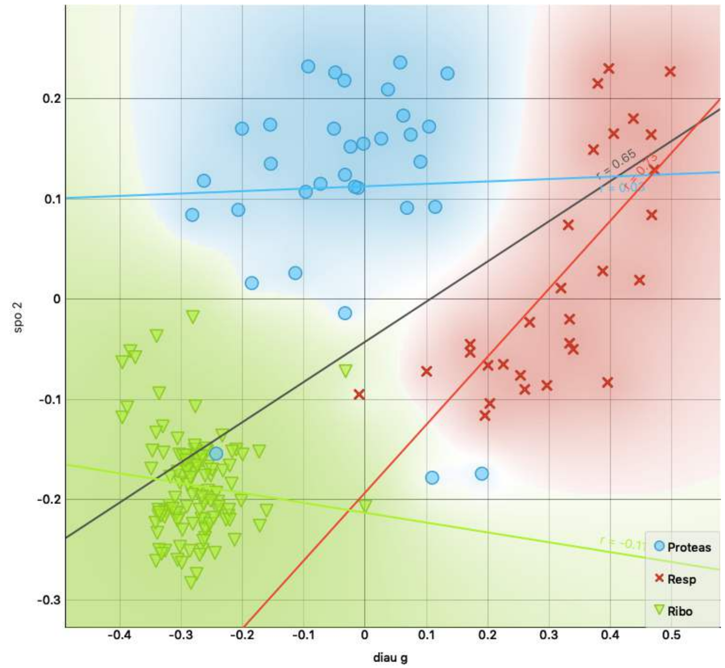
Knowing systems have a limit to their **working memory** under realistic assumptions

Knowing systems also have to deal with noisy data and uncertainty

Once the scale of the data set exceeds the capacity of the system's cognitive limits [working memory, processing power] pattern recognition must be off-sourced

- E.g., regression and other statistical analyses
- Machine Learning classifications (ANNs)
- Distance analyses (Hamming, Manhattan)

Less memory and cognitive resources are required to retrieve and interpret patterns



Pattern recognition

Knowing systems have a limit to their **working memory** under realistic assumptions

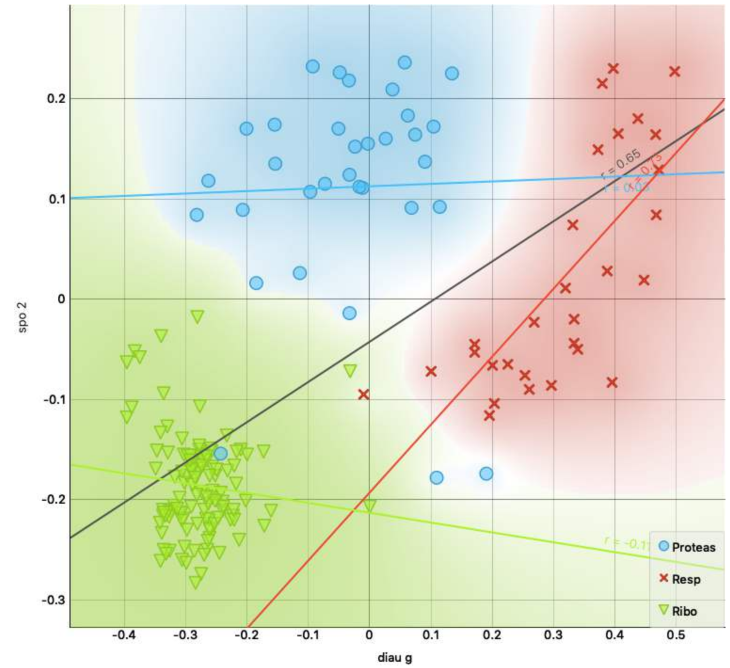
Knowing systems also have to deal with noisy data and uncertainty

Once the scale of the data set exceeds the capacity of the system's cognitive limits [working memory, processing power] pattern recognition must be off-sourced

- E.g., regression and other statistical analyses
- Machine Learning classifications (ANNs)
- Distance analyses (Hamming, Manhattan)

Less memory and cognitive resources are required to retrieve and interpret patterns

A pattern therefore has *information* the data does not



John S Wilkins john@wilkins.id.au

Pattern recognition

Knowing systems have a limit to their **working memory** under realistic assumptions

Knowing systems also have to deal with noisy data and uncertainty

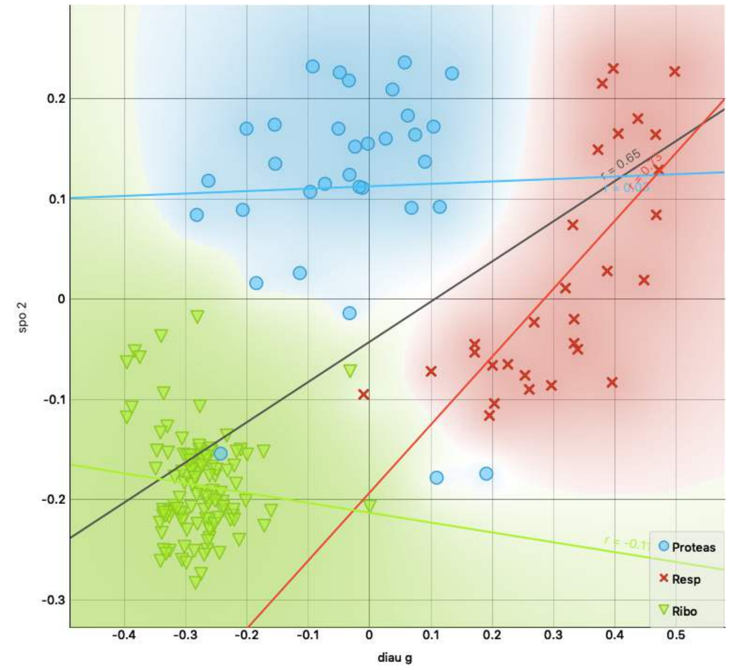
Once the scale of the data set exceeds the capacity of the system's cognitive limits [working memory, processing power] pattern recognition must be off-sourced

- E.g., regression and other statistical analyses
- Machine Learning classifications (ANNs)
- Distance analyses (Hamming, Manhattan)

Less memory and cognitive resources are required to retrieve and interpret patterns

A pattern therefore has *information* the data does not

[Here, information is a property of the state of a knowing system]



John S Wilkins john@wilkins.id.au

Machine learning and understanding

As any ML system is a Turing machine, it follows that, even when we do not know it, there is an algorithm *A* that generates the output pattern *P*

John S Wilkins john@wilkins.id.au

Machine learning and understanding

As any ML system is a Turing machine, it follows that, even when we do not know it, there is an algorithm *A* that generates the output pattern *P*

The *information content* of P is therefore the *complexity* of *A*, and the data set

John S Wilkins john@wilkins.id.au

Machine learning and understanding

As any ML system is a Turing machine, it follows that, even when we do not know it, there is an algorithm *A* that generates the output pattern *P*

The *information content* of P is therefore the *complexity* of *A*, and the data set

- If *P* is **lossy**, then its information content is lower in complexity than the data set

John S Wilkins john@wilkins.id.au

Machine learning and understanding

As any ML system is a Turing machine, it follows that, even when we do not know it, there is an algorithm *A* that generates the output pattern *P*

The *information content* of P is therefore the *complexity* of *A*, and the data set

- If *P* is **lossy**, then its information content is lower in complexity than the data set
- If *P* is **lossless**, then the complexity of *A* equals the complexity of *P*, by defn

John S Wilkins john@wilkins.id.au

Machine learning and understanding

As any ML system is a Turing machine, it follows that, even when we do not know it, there is an algorithm *A* that generates the output pattern *P*

The *information content* of P is therefore the *complexity* of *A*, and the data set

- If *P* is **lossy**, then its information content is lower in complexity than the data set
- If *P* is **lossless**, then the complexity of *A* equals the complexity of *P*, by defn

So *A* is a *compression* algorithm. By analogy, when we generate a pattern (such as a regression curve) from a large or noisy data set, we are *compressing* the data to generate information, and from that information we can fairly say we *know* the domain the data set samples or represents.

John S Wilkins john@wilkins.id.au

Machine learning and understanding

As any ML system is a Turing machine, it follows that, even when we do not know it, there is an algorithm *A* that generates the output pattern *P*

The *information content* of P is therefore the *complexity* of *A*, and the data set

- If *P* is **lossy**, then its information content is lower in complexity than the data set
- If *P* is **lossless**, then the complexity of *A* equals the complexity of *P*, by defn

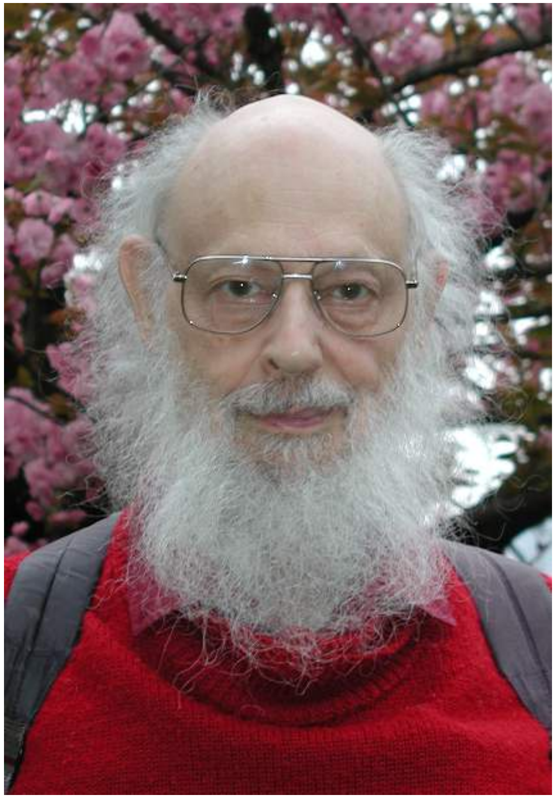
So *A* is a *compression* algorithm. By analogy, when we generate a pattern (such as a regression curve) from a large or noisy data set, we are *compressing* the data to generate information, and from that information we can fairly say we *know* the domain the data set samples or represents.

The complexity of *P* is specified in Algorithmic Information Theory (AIT) as the length (in bits) of the shortest program that generates *P*.

John S Wilkins john@wilkins.id.au

Kolmogorov compressed

AIT is based on the work of Andrey Kolmogorov (1903–1987) and Ray Solomonov (1926–2009)



John S Wilkins john@wilkins.id.au

Kolmogorov compressed

AIT is based on the work of Andrey Kolmogorov (1903–1987) and Ray Solomonov (1926–2009)

Kolmogorov Complexity (*K*) is the length of the shortest program in a language *L* that generates a string *w*

John S Wilkins john@wilkins.id.au

Kolmogorov compressed

AIT is based on the work of Andrey Kolmogorov (1903–1987) and Ray Solomonov (1926–2009)

Kolmogorov Complexity (*K*) is the length of the shortest program in a language *L* that generates a string *w*

- The most complex string in *L* is a fully random string

John S Wilkins john@wilkins.id.au

Kolmogorov compressed

AIT is based on the work of Andrey Kolmogorov (1903–1987) and Ray Solomonov (1926–2009)

Kolmogorov Complexity (*K*) is the length of the shortest program in a language *L* that generates a string *w*

- The most complex string in *L* is a fully random string
    - A fully random string is incompressible

John S Wilkins john@wilkins.id.au

Kolmogorov compressed

AIT is based on the work of Andrey Kolmogorov (1903–1987) and Ray Solomonov (1926–2009)

Kolmogorov Complexity (*K*) is the length of the shortest program in a language *L* that generates a string *w*

- The most complex string in *L* is a fully random string
  - A fully random string is incompressible
  - The working memory of a ML system is thus the largest random string that it can hold in memory while applying functions to it

John S Wilkins john@wilkins.id.au

Kolmogorov compressed

AIT is based on the work of Andrey Kolmogorov (1903–1987) and Ray Solomonov (1926–2009)

Kolmogorov Complexity (*K*) is the length of the shortest program in a language *L* that generates a string *w*

- The most complex string in *L* is a fully random string
  - A fully random string is incompressible
  - The working memory of a ML system is thus the largest random string that it can hold in memory while applying functions to it

Thus *K* gives the complexity of a string (of digits, or other structured data)

John S Wilkins john@wilkins.id.au

Kolmogorov compressed

AIT is based on the work of Andrey Kolmogorov (1903–1987) and Ray Solomonov (1926–2009)

Kolmogorov Complexity (*K*) is the length of the shortest program in a language *L* that generates a string *w*

- The most complex string in *L* is a fully random string
  - A fully random string is incompressible
  - The working memory of a ML system is thus the largest random string that it can hold in memory while applying functions to it

Thus *K* gives the complexity of a string (of digits, or other structured data)

If we conceive of ourselves as knowing systems analogous to ML systems, the information in data is thus the *K* "program" in our cognitive processes

Li, Ming, and Paul Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. 4th ed. Texts in Computer Science. Springer International Publishing, 2019.

Grünwald, Peter, and Paul M. B. Vitányi. "Shannon Information and Kolmogorov Complexity." *CoRR* cs.IT/0410002 (2004).

Wallace, C. S., and D. L. Dowe. "Minimum Message Length and Kolmogorov Complexity." *The Computer Journal* 42, no. 4 (January 1, 1999): 270–83.

Reducing the complexity of data

Note that while using complexity metrics is based on the *subject* – the ML or the knowing system – it is not *subjective*, as the experience felt by the knower is not relevant for understanding the information

John S Wilkins john@wilkins.id.au

Reducing the complexity of data

Note that while using complexity metrics is based on the *subject* – the ML or the knowing system – it is not *subjective*, as the experience felt by the knower is not relevant for understanding the information

[Nor is this an objectivist account, as it does not depend upon a truth operator]

John S Wilkins john@wilkins.id.au

Reducing the complexity of data

Note that while using complexity metrics is based on the *subject* – the ML or the knowing system – it is not *subjective*, as the experience felt by the knower is not relevant for understanding the information

[Nor is this an objectivist account, as it does not depend upon a truth operator]

Complexity of the information is inversely related to the informativeness for a limited knowing system

John S Wilkins john@wilkins.id.au

Reducing the complexity of data

Note that while using complexity metrics is based on the *subject* – the ML or the knowing system – it is not *subjective*, as the experience felt by the knower is not relevant for understanding the information

[Nor is this an objectivist account, as it does not depend upon a truth operator]

Complexity of the information is inversely related to the informativeness for a limited knowing system

- Put another way, the tractability of inferences increases as complexity reduces

John S Wilkins john@wilkins.id.au

Reducing the complexity of data

Note that while using complexity metrics is based on the *subject* – the ML or the knowing system – it is not *subjective*, as the experience felt by the knower is not relevant for understanding the information

[Nor is this an objectivist account, as it does not depend upon a truth operator]

Complexity of the information is inversely related to the informativeness for a limited knowing system

- Put another way, the tractability of inferences increases as complexity reduces

[Inversely, tractability reduces as data complexity increases, unless it is compressible]

John S Wilkins john@wilkins.id.au

Reducing the complexity of data

Note that while using complexity metrics is based on the *subject* – the ML or the knowing system – it is not *subjective*, as the experience felt by the knower is not relevant for understanding the information

[Nor is this an objectivist account, as it does not depend upon a truth operator]

Complexity of the information is inversely related to the informativeness for a limited knowing system

- Put another way, the tractability of inferences increases as complexity reduces

[Inversely, tractability reduces as data complexity increases, unless it is compressible]

- Compression of data to information is a **lossy** process; the structure/pattern in the data is simplified through regression, etc.

John S Wilkins john@wilkins.id.au

Reducing the complexity of data

Note that while using complexity metrics is based on the *subject* – the ML or the knowing system – it is not *subjective*, as the experience felt by the knower is not relevant for understanding the information

[Nor is this an objectivist account, as it does not depend upon a truth operator]

Complexity of the information is inversely related to the informativeness for a limited knowing system

- Put another way, the tractability of inferences increases as complexity reduces

  [Inversely, tractability reduces as data complexity increases, unless it is compressible]

- Compression of data to information is a **lossy** process; the structure/pattern in the data is simplified through regression, etc.
- The function of this information is to set the expectations of the knowing system [but that's another talk; contrastive explanation is why]

John S Wilkins john@wilkins.id.au

Reducing the complexity of data

Note that while using complexity metrics is based on the *subject* – the ML or the knowing system – it is not *subjective*, as the experience felt by the knower is not relevant for understanding the information

[Nor is this an objectivist account, as it does not depend upon a truth operator]

Complexity of the information is inversely related to the informativeness for a limited knowing system

- Put another way, the tractability of inferences increases as complexity reduces

  [Inversely, tractability reduces as data complexity increases, unless it is compressible]

- Compression of data to information is a **lossy** process; the structure/pattern in the data is simplified through regression, etc.
- The function of this information is to set the expectations of the knowing system [but that's another talk; contrastive explanation is why]
  - In short: the information derived from the data sets the prior probabilities for future data, highlighting anomalies and points of interest

John S Wilkins john@wilkins.id.au

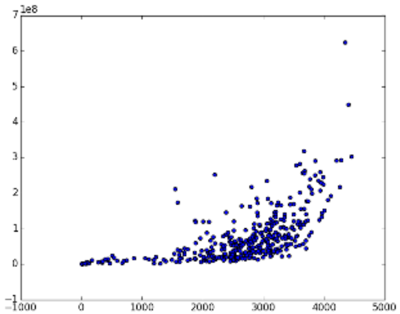
Compression and generalisation

Solomonoff held that all inference problems can be cast in the form of extrapolation from an ordered sequence of binary symbols.
[Li and Vintányi 2019,63]

John S Wilkins john@wilkins.id.au

Compression and generalisation

Measurement

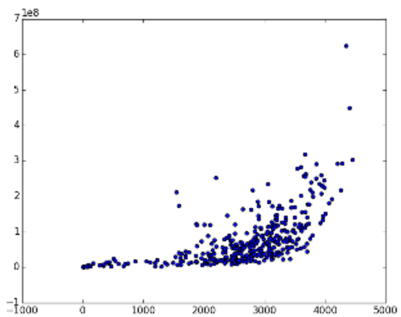


compresses to

Solomonoff held that all inference problems can be cast in the form of extrapolation from an ordered sequence of binary symbols.
[Li and Vintányi 2019,63]

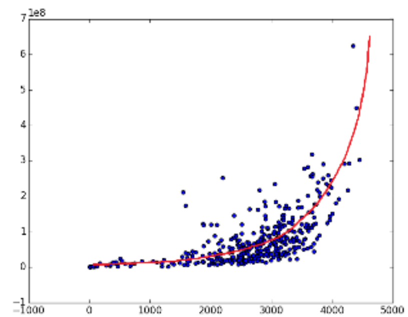
Compression and generalisation

Measurement



compresses to

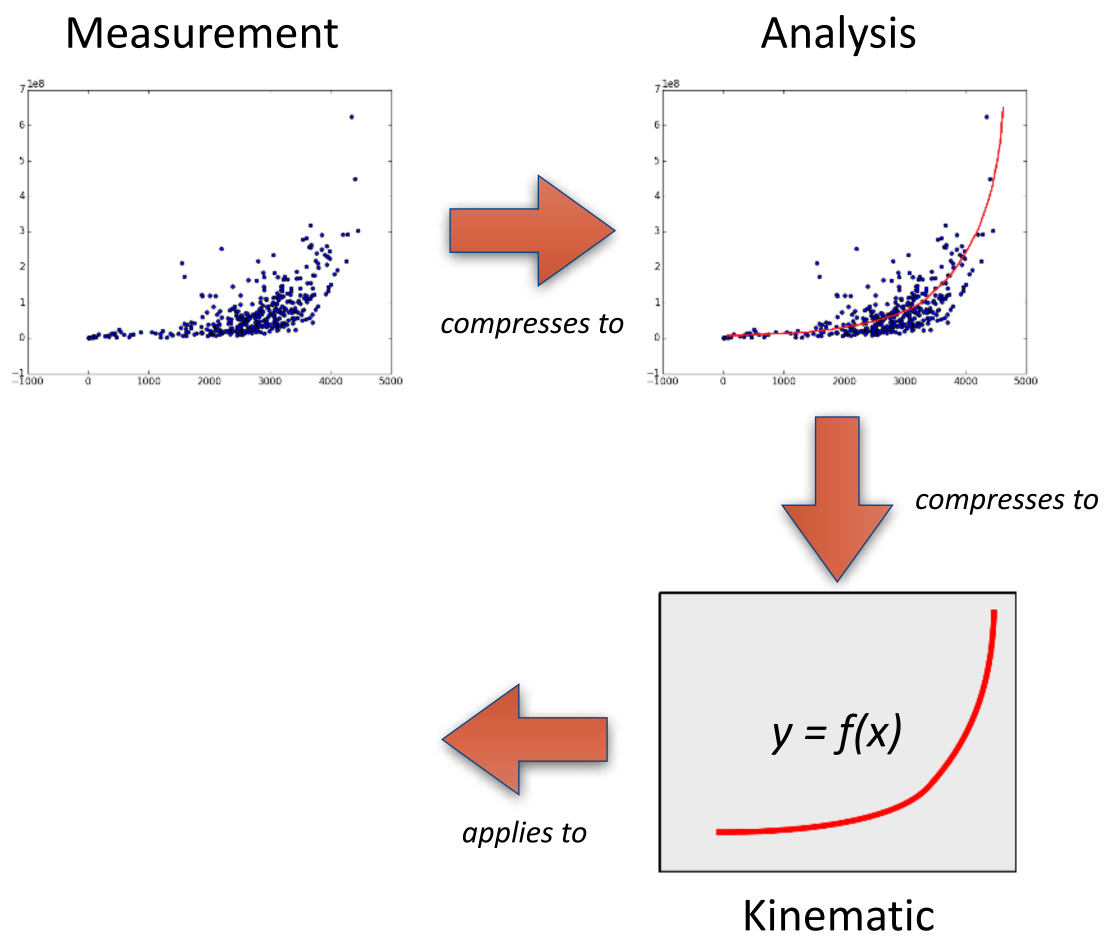
Analysis



compresses to

Solomonoff held that all inference problems can be cast in the form of extrapolation from an ordered sequence of binary symbols.
[Li and Vintányi 2019,63]

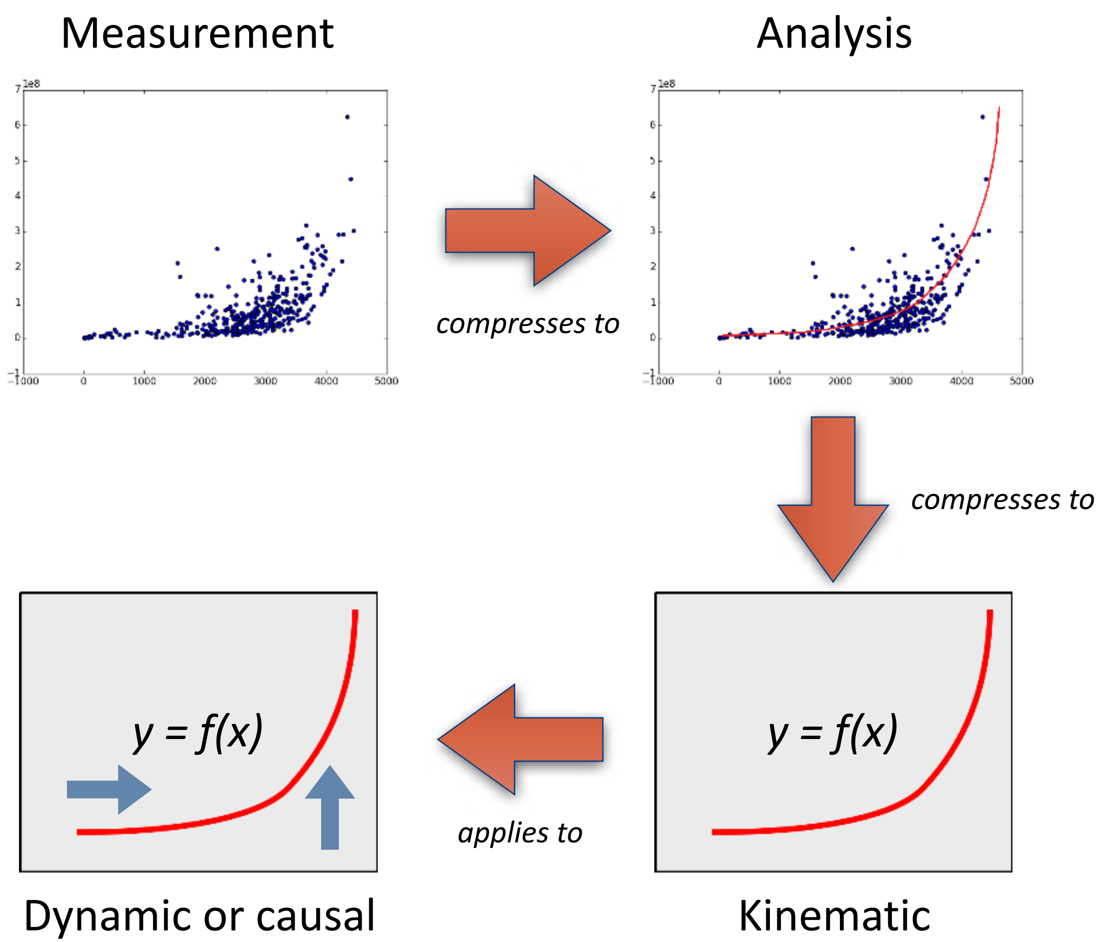
Compression and generalisation



Solomonoff held that all inference problems can be cast in the form of extrapolation from an ordered sequence of binary symbols.
[Li and Vitányi 2019,63]

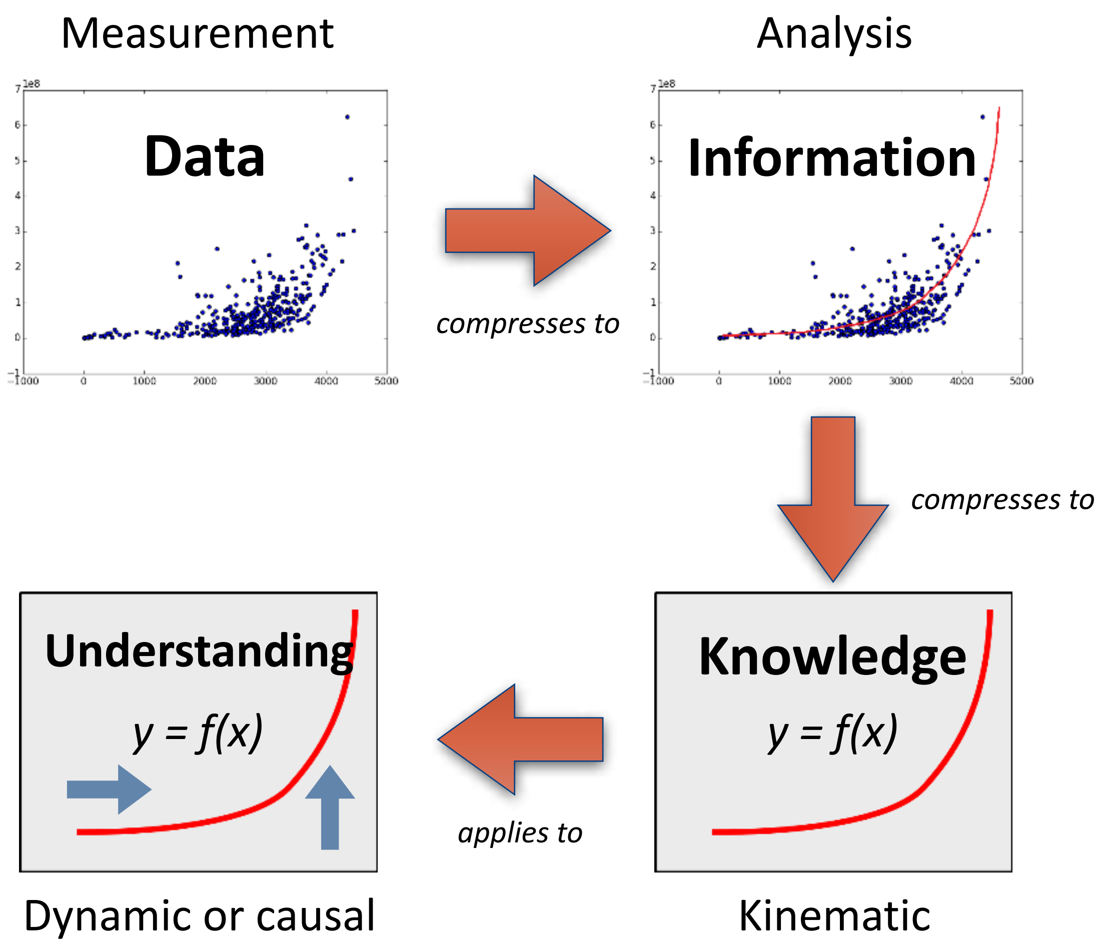
John S Wilkins john@wilkins.id.au

Compression and generalisation



Solomonoff held that all inference problems can be cast in the form of extrapolation from an ordered sequence of binary symbols.
[Li and Vintányi 2019,63]

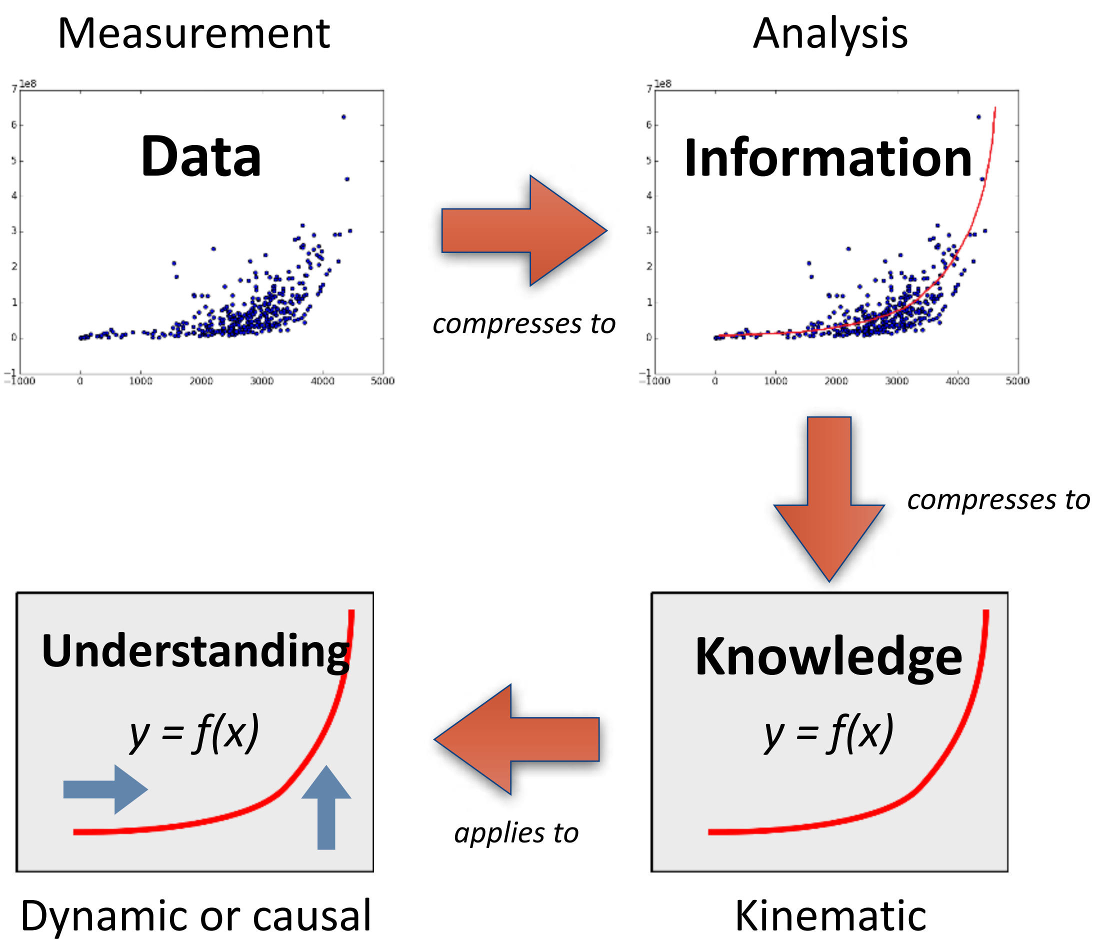
Compression and generalisation



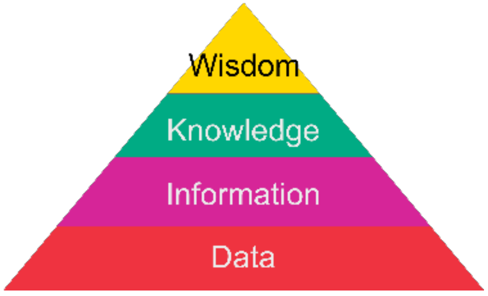
Solomonoff held that all inference problems can be cast in the form of extrapolation from an ordered sequence of binary symbols.
[Li and Vintányi 2019,63]

John S Wilkins john@wilkins.id.au

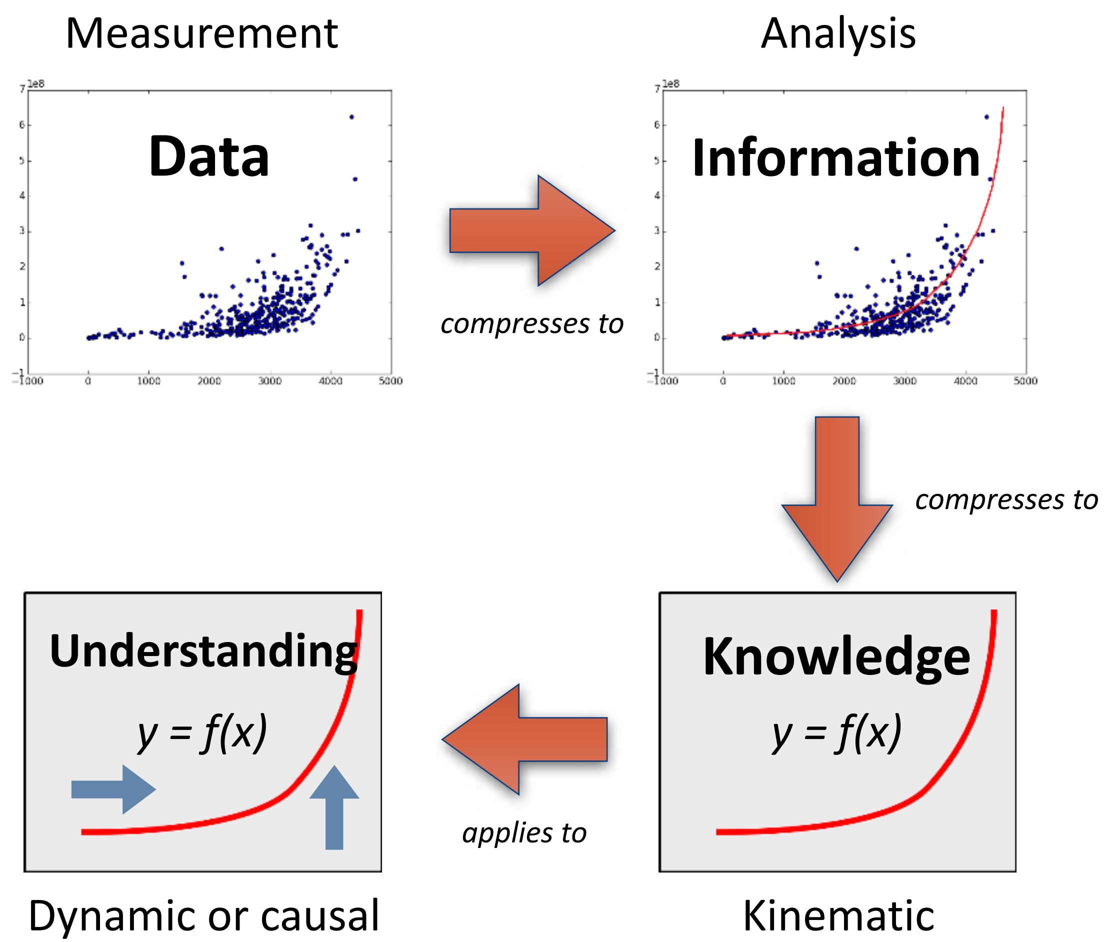
Compression and generalisation



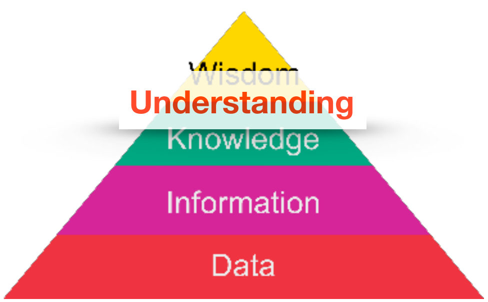
Solomonoff held that all inference problems can be cast in the form of extrapolation from an ordered sequence of binary symbols.
[Li and Vitányi 2019,63]



Compression and generalisation



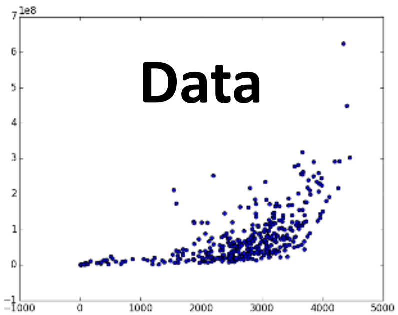
Solomonoff held that all inference problems can be cast in the form of extrapolation from an ordered sequence of binary symbols.
[Li and Vitányi 2019,63]



John S Wilkins john@wilkins.id.au

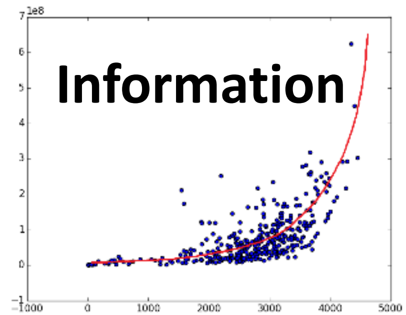
Compression and generalisation

Measurement

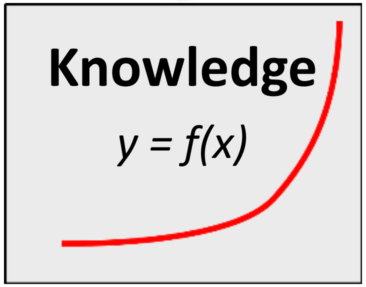


compresses to

Analysis

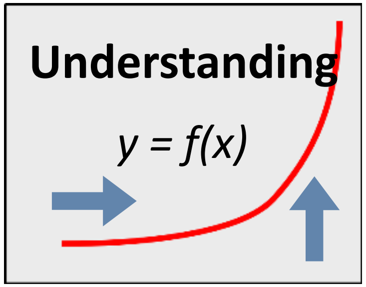


compresses to



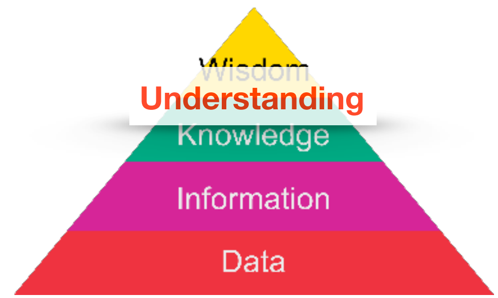
Kinematic

applies to



Dynamic or causal

Solomonoff held that all inference problems can be cast in the form of extrapolation from an ordered sequence of binary symbols.
[Li and Vintányi 2019,63]



What counts as wisdom is left to you to decide

John S Wilkins john@wilkins.id.au

Subjects understanding

So, how might this analogy work in the arena of science? Scientists seek to understand the universe, of course, as Hempel said.

John S Wilkins john@wilkins.id.au

Subjects understanding

So, how might this analogy work in the arena of science? Scientists seek to understand the universe, of course, as Hempel said.

- But what does “scientists” refer to? What are the knowing systems in science?

John S Wilkins john@wilkins.id.au

Subjects understanding

So, how might this analogy work in the arena of science? Scientists seek to understand the universe, of course, as Hempel said.

- But what does “scientists” refer to? What are the knowing systems in science?
    - *particular* scientists

John S Wilkins john@wilkins.id.au

Subjects understanding

So, how might this analogy work in the arena of science? Scientists seek to understand the universe, of course, as Hempel said.

- But what does “scientists” refer to? What are the knowing systems in science?
  - *particular* scientists
  - *disciplines* and *sub disciplines*

John S Wilkins john@wilkins.id.au

Subjects understanding

So, how might this analogy work in the arena of science? Scientists seek to understand the universe, of course, as Hempel said.

- But what does "scientists" refer to? What are the knowing systems in science?
  - *particular* scientists
  - *disciplines* and *sub disciplines*
  - *research groups*

John S Wilkins john@wilkins.id.au

Subjects understanding

So, how might this analogy work in the arena of science? Scientists seek to understand the universe, of course, as Hempel said.

- But what does “scientists” refer to? What are the knowing systems in science?
  - *particular* scientists
  - *disciplines* and *sub disciplines*
  - *research groups*

I would take this approach to apply equally well to both individual and collective knowing systems.

John S Wilkins john@wilkins.id.au

Subjects understanding

So, how might this analogy work in the arena of science? Scientists seek to understand the universe, of course, as Hempel said.

- But what does “scientists” refer to? What are the knowing systems in science?
  - *particular* scientists
  - *disciplines* and *sub disciplines*
  - *research groups*

I would take this approach to apply equally well to both individual and collective knowing systems.

We can say that <u>Professor *X*</u> understands some phenomenon, or that <u>Discipline Δ</u> understands some phenomenon independently (individuals can fail to understand what the discipline does, and *vice versa*)

John S Wilkins john@wilkins.id.au

Subjects understanding

So, how might this analogy work in the arena of science? Scientists seek to understand the universe, of course, as Hempel said.

- But what does “scientists” refer to? What are the knowing systems in science?
  - *particular* scientists
  - *disciplines* and *sub disciplines*
  - *research groups*

I would take this approach to apply equally well to both individual and collective knowing systems.

We can say that <u>Professor *X*</u> understands some phenomenon, or that <u>Discipline Δ</u> understands some phenomenon independently (individuals can fail to understand what the discipline does, and *vice versa*)

[Does this mean we have actual Turing machines in our heads? No, it’s an abstract way to consider the problem (just as neurones are not artificial neurones or *v.v.*)]

John S Wilkins john@wilkins.id.au

Insert case studies here

I do not have time here to give proper case studies; so I will handwave at some:

John S Wilkins john@wilkins.id.au

Insert case studies here

I do not have time here to give proper case studies; so I will handwave at some:

- **Gas laws** (Boyles' Law, Charles' Law, Gay-Lussac's Law, Avogadro's Law, etc.) leading to the Ideal Gas Law, to van der Waals' equation

John S Wilkins john@wilkins.id.au

Insert case studies here

I do not have time here to give proper case studies; so I will handwave at some:

- **Gas laws** (Boyles' Law, Charles' Law, Gay-Lussac's Law, Avogadro's Law, etc.) leading to the Ideal Gas Law, to van der Waals' equation
- **Periodic table** construction by empirical lab work, followed first by valency, then electron shell, then quantum mechanical causal accounts

John S Wilkins john@wilkins.id.au

Insert case studies here

I do not have time here to give proper case studies; so I will handwave at some:

- **Gas laws** (Boyles' Law, Charles' Law, Gay-Lussac's Law, Avogadro's Law, etc.) leading to the Ideal Gas Law, to van der Waals' equation
- **Periodic table** construction by empirical lab work, followed first by valency, then electron shell, then quantum mechanical causal accounts
- **Mendelian genetics** (phenomenal) through to population genetics (kinematic), through to molecular genetics (dynamic)

John S Wilkins john@wilkins.id.au

Insert case studies here

I do not have time here to give proper case studies; so I will handwave at some:

- **Gas laws** (Boyles' Law, Charles' Law, Gay-Lussac's Law, Avogadro's Law, etc.) leading to the Ideal Gas Law, to van der Waals' equation
- **Periodic table** construction by empirical lab work, followed first by valency, then electron shell, then quantum mechanical causal accounts
- **Mendelian genetics** (phenomenal) through to population genetics (kinematic), through to molecular genetics (dynamic)
- Arguably: observational to theoretical **ecology**, through to ecological thermodynamics and "ecological orbits" (Haynie 2001, Ginsburg and Colyvan 2004)

John S Wilkins john@wilkins.id.au

Insert case studies here

I do not have time here to give proper case studies; so I will handwave at some:

- **Gas laws** (Boyles' Law, Charles' Law, Gay-Lussac's Law, Avogadro's Law, etc.) leading to the Ideal Gas Law, to van der Waals' equation
- **Periodic table** construction by empirical lab work, followed first by valency, then electron shell, then quantum mechanical causal accounts
- **Mendelian genetics** (phenomenal) through to population genetics (kinematic), through to molecular genetics (dynamic)
- Arguably: observational to theoretical **ecology**, through to ecological thermodynamics and "ecological orbits" (Haynie 2001, Ginsburg and Colyvan 2004)
- **Tectonic drift** theory (Oreskes and le Grand 2003)

John S Wilkins john@wilkins.id.au

Insert case studies here

I do not have time here to give proper case studies; so I will handwave at some:

- **Gas laws** (Boyles' Law, Charles' Law, Gay-Lussac's Law, Avogadro's Law, etc.) leading to the Ideal Gas Law, to van der Waals' equation
- **Periodic table** construction by empirical lab work, followed first by valency, then electron shell, then quantum mechanical causal accounts
- **Mendelian genetics** (phenomenal) through to population genetics (kinematic), through to molecular genetics (dynamic)
- Arguably: observational to theoretical **ecology**, through to ecological thermodynamics and "ecological orbits" (Haynie 2001, Ginsburg and Colyvan 2004)
- **Tectonic drift** theory (Oreskes and le Grand 2003)
- *&c.*

Ginzburg, Lev R., and Mark Colyvan. *Ecological Orbits: How Planets Move and Populations Grow*. Oxford; New York: Oxford University Press, 2004.

Haynie, Donald T. *Biological Thermodynamics*. Cambridge; New York: Cambridge University Press, 2001.

Oreskes, Naomi, and Homer E. LeGrand. *Plate Tectonics: An Insider's History of the Modern Theory of the Earth*. Boulder, CO.: Westview Press, 2003.

John S Wilkins john@wilkins.id.au

Conclusions

Science is, in my view, an attempt to express the greatest amount of empirical consequences in the fewest terms. Understanding is thus:

John S Wilkins john@wilkins.id.au

Conclusions

Science is, in my view, an attempt to express the greatest amount of empirical consequences in the fewest terms. Understanding is thus:

- Empirical adequacy (van Fraassen)

John S Wilkins john@wilkins.id.au

Conclusions

Science is, in my view, an attempt to express the greatest amount of empirical consequences in the fewest terms. Understanding is thus:

- Empirical adequacy (van Fraassen)
- Generalisation/-ability

John S Wilkins john@wilkins.id.au

Conclusions

Science is, in my view, an attempt to express the greatest amount of empirical consequences in the fewest terms. Understanding is thus:

- Empirical adequacy (van Fraassen)
- Generalisation/-ability
- Causal explanation

John S Wilkins john@wilkins.id.au

Conclusions

Science is, in my view, an attempt to express the greatest amount of empirical consequences in the fewest terms. Understanding is thus:

- Empirical adequacy (van Fraassen)
- Generalisation/-ability
- Causal explanation

which are pretty much de Regt's three criteria restated.

John S Wilkins john@wilkins.id.au

Conclusions

Science is, in my view, an attempt to express the greatest amount of empirical consequences in the fewest terms. Understanding is thus:

- Empirical adequacy (van Fraassen)
- Generalisation/-ability
- Causal explanation

which are pretty much de Regt's three criteria restated.

The initial epistemic activities of **measurement**, **analysis** and **kinematic generalisation** involve compression so that we may *comprehend* the things needing explanation

John S Wilkins john@wilkins.id.au

Conclusions

Science is, in my view, an attempt to express the greatest amount of empirical consequences in the fewest terms. Understanding is thus:

- Empirical adequacy (van Fraassen)
- Generalisation/-ability
- Causal explanation

which are pretty much de Regt's three criteria restated.

The initial epistemic activities of **measurement**, **analysis** and **kinematic generalisation** involve compression so that we may *comprehend* the things needing explanation

- Compression of data is a kind of pattern matching, on the basis of which ampliative inferences rest.

Conclusions

Science is, in my view, an attempt to express the greatest amount of empirical consequences in the fewest terms. Understanding is thus:

- Empirical adequacy (van Fraassen)
- Generalisation/-ability
- Causal explanation

which are pretty much de Regt's three criteria restated.

The initial epistemic activities of **measurement**, **analysis** and **kinematic generalisation** involve compression so that we may *comprehend* the things needing explanation

- Compression of data is a kind of pattern matching, on the basis of which ampliative inferences rest.
- Big data is not, *ipso facto*, a good thing

John S Wilkins john@wilkins.id.au

Acknowledgements

John S Wilkins john@wilkins.id.au

Acknowledgements

For discussion:

- John Collier
- Malte Ebach
- Ward Wheeler
- Adam Ford

John S Wilkins john@wilkins.id.au

Acknowledgements

For discussion:

- John Collier
- Malte Ebach
- Ward Wheeler
- Adam Ford
- Marcus Hutter

For funding:

John S Wilkins john@wilkins.id.au

Acknowledgements

For discussion:

- John Collier
- Malte Ebach
- Ward Wheeler
- Adam Ford
- Marcus Hutter

For funding:

- ISHPSSSB conference organisers

For ear nibbles

- Clio, a muse

Thank you

John S Wilkins john@wilkins.id.au