

# ON INFORMATION: AN INTRODUCTION

---

jack w chen

department of east asian languages,  
literatures, and cultures

university of virginia

# What is “information”?

## **Oxford English Dictionary**

- I. The imparting of knowledge in general.
- II. The imparting of incriminating knowledge.
- III. The giving of form.

# “The imparting of knowledge in general”

1a. The shaping of the mind or character; communication of instructive knowledge; education, training; †advice (*obs.*). Now *rare*. (1387: Fyve bookes com doun from heven for informacioun of mankynde.)

1b. As a count noun: a teaching; an instruction; a piece of advice. *Obs.*

1c. Chiefly *Christian Church*. Divine influence or direction; inspiration, esp. through the Holy Spirit. (c1450: Crist was..First a prophete by holy informacion.)

1d. Capacity of informing; instructiveness. *Obs. rare.*

# “The imparting of knowledge in general”

2a. Knowledge communicated concerning some particular fact, subject, or event; that of which one is apprised or told; intelligence, news.

2b. As a count noun: a fact or circumstance of which a person is told; a piece of news or intelligence; (in early use) an account or narrative (*of* something). Now *rare*.

2c. As a mathematically defined quantity divorced from any concept of news or meaning; *spec.* one which represents the degree of choice exercised in the selection or formation of one particular symbol, message, etc., out of a number of possible ones, and which is defined logarithmically in terms of the statistical probabilities of occurrence of the symbol or the elements of the message.

2d. Separated from, or without the implication of, reference to a person informed: that which inheres in one of two or more alternative sequences, arrangements, etc., that produce different responses in something, and which is capable of being stored in, transmitted by, and communicated to inanimate things.

2e. Contrasted with *data*: that which is obtained by the processing of data.

# Information as measure

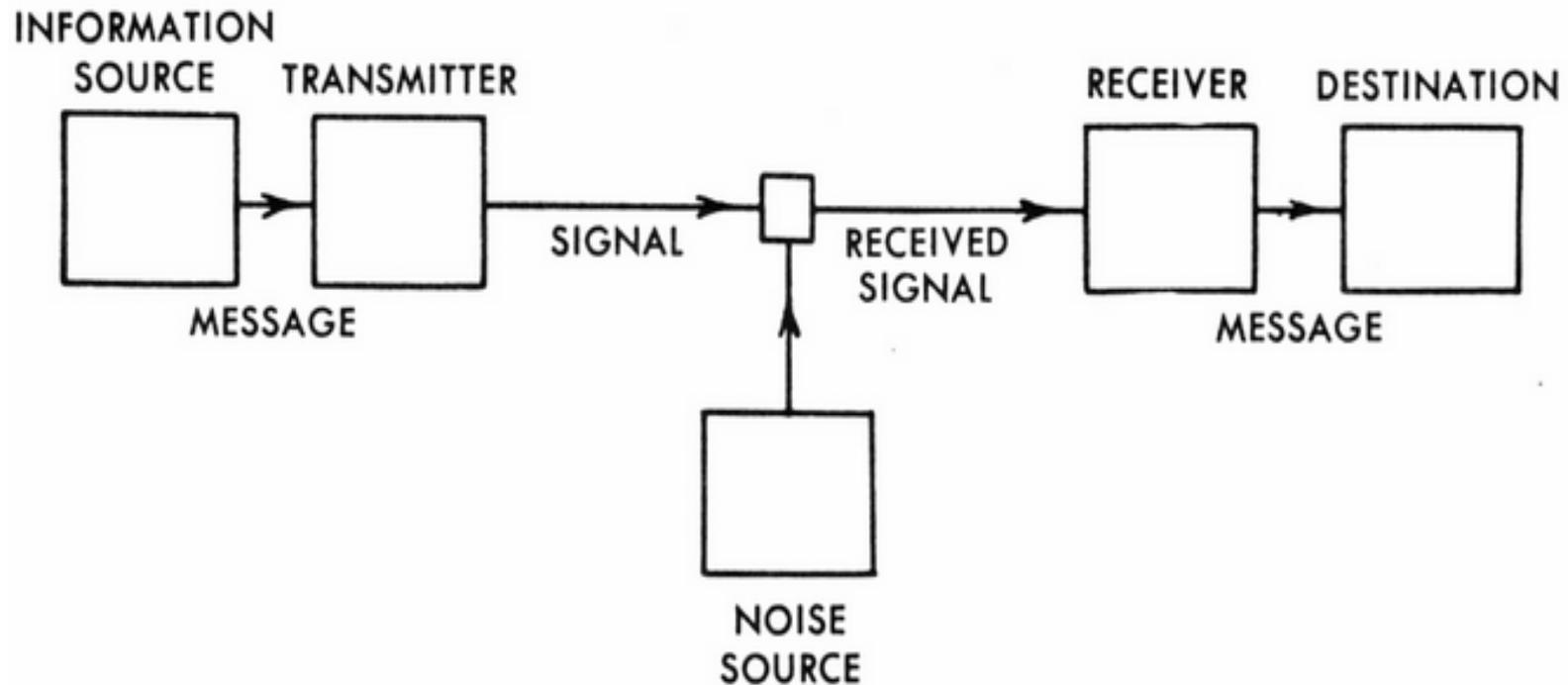
I.2c. As a mathematically defined quantity divorced from any concept of news or meaning; *spec.* one which represents the degree of choice exercised in the selection or formation of one particular symbol, message, etc., out of a number of possible ones, and which is defined logarithmically in terms of the statistical probabilities of occurrence of the symbol or the elements of the message.

# Information in “Information Theory”

Warren Weaver, “Recent Contributions to the Mathematical Theory of Information” (1949)

To be sure, this word information in communication theory relates not so much to what you *do* say, as to what you *could* say. That is, information is a measure of one’s freedom of choice when one selects a message. ... The concept of information applies not to the individual messages (as the concept of meaning would), but rather to the situation as a whole, the unit information indicating that in this situation one has an amount of freedom of choice, in selecting a message, which it is convenient to regard as a standard or unit amount.

# Claude Shannon and Warren Weaver Model of Communication



# Information and Entropy

- S-W uses a base-2 model so a unit of information is 1 bit (binary digit), consisting of 2 possible states (1 or 0, yes or no, on or off).
- Weaver: “That information be measured by entropy is, after all, natural when we remember that information, in communication theory, is associated with the amount of freedom of choice we have in constructing messages. Thus for a communication source one can say, just as he would also say it of a thermodynamic ensemble, ‘This situation is highly organized, it is not characterized by a large degree of randomness or of choice—that is to say, the information (or the entropy) is low.’”
- Shannon: “We shall call  $H = -\sum p_i \log_2 p_i$  the entropy of the set of probabilities  $p_1, \dots, p_n$ ” where the quantity  $H$  of information is the set ( $\Sigma$ ) of probable states of the message in bits ( $-\log_2 p_i$ ), or “a reasonable measure of choice or information.”

# Information and meaning

I.2d. Separated from, or without the implication of, reference to a person informed: that which inheres in one of two or more alternative sequences, arrangements, etc., that produce different responses in something, and which is capable of being stored in, transmitted by, and communicated to inanimate things.

(1937, *Discovery*: “The whole difficulty resides in the amount of definition in the [television] picture, or, as the engineers put it, the amount of information to be transmitted in a given time.”)

(1944, *Journal of Scientific Instruments*: “Information is conveyed to the machine by means of punched cards.”)

# Information vs. “Information”

Donald M. MacKay, *Information, Mechanism, and Meaning*

- “The trouble here appears to be due largely to a confusion of the concept of *information* with that of *information-content*—the confusion of a *thing* with the *measure of a thing*. Communication engineers have not developed a concept of information at all.” (1953)
- “Information makes a difference to *what we believe to be the case*. It is always information *about* something. Its effect is to change, in one way or another, the total of ‘all that is the case’ for us.” (1950)

# Information and Difference

Gregory Bateson, “Cybernetics of Self: A Theory of Alcoholism” (1971; rpt in *Steps to an Ecology of Mind*)

“A ‘bit’ of information is definable as a difference which makes a difference. Such a difference, as it travels and undergoes successive transformation in a circuit, is an elementary idea.”

## Bateson, continued

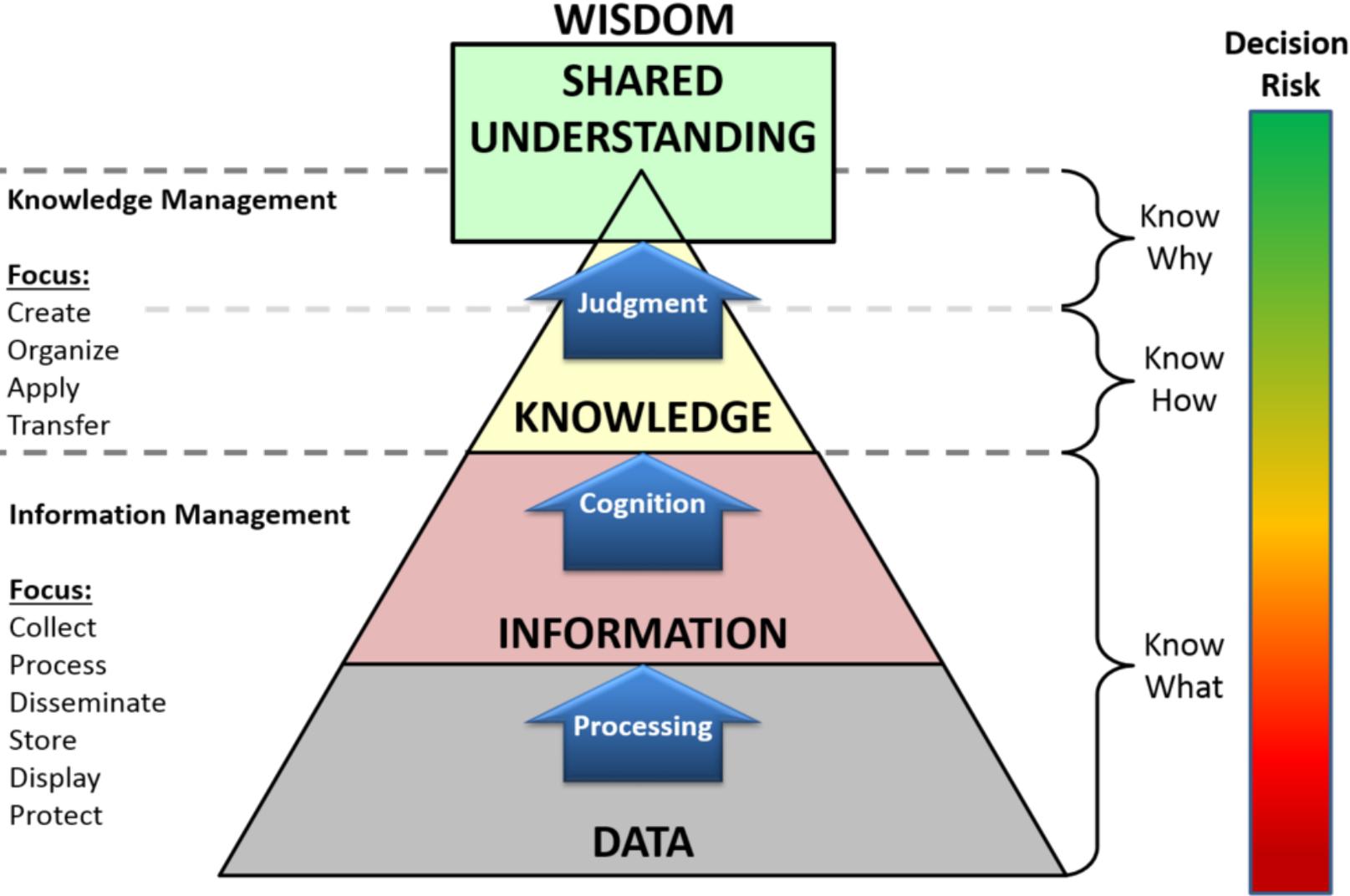
“Consider a man felling a tree with an axe. Each stroke of the axe is modified or corrected, according to the shape of the cut face of the tree left by the previous stroke. This self-corrective (i.e., mental) process is brought about by a total system, tree-eyes-brain-muscles-axe-stroke-tree; and it is this total system that has the characteristics of immanent mind. More correctly, we should spell the matter out as: (differences in tree) - (differences in retina) - (differences in brain) - (differences in muscles) - (differences in movement of axe) - (differences in tree), etc. What is transmitted around the circuit is transforms of differences. And, as noted above, a difference which makes a difference is an idea or unit of information.”

# Information in contrast to data

I.2e. Contrasted with *data*: that which is obtained by the processing of data.

(2007, *Information and Management* 44.1): “A common distinction within this domain is that data is raw numbers and facts, information is processed data, while knowledge has “the highest value, the most human contribution, the greatest relevance to decisions and actions, and the greatest dependence on a specific situation or context.”)

# Knowledge Management Cognitive Pyramid



graphic from the U.S. Army Knowledge Management Proponent

# So what is / are “data”?

Daniel Rosenberg, “Data before the Fact” (2013)

The word “data” comes to English from Latin. It is the plural of the Latin word *datum*, which itself is the neuter past participle of the verb *dare*, to give. A “datum” in English, then, is something given in an argument, something taken for granted. This is in contrast to “fact,” which derives from the neuter past participle of the Latin verb *facere*, to do, whence we have the English word “fact,” for that which was done, occurred, or exists. The etymology of “data” also contrasts with that of “evidence,” from the Latin verb *vidēre*, to see. There are important distinctions here: facts are ontological, evidence is epistemological, data is rhetorical.

# Luciano Floridi's General Definition of Information (Semantic Information)

Floridi, "Information" (2004, *Philosophy of Computing and Information*)

GDI)  $\sigma$  is an instance of information, understood as objective semantic content, if and only if:

GDI.1)  $\sigma$  consists of  $n$  data ( $d$ ), for  $n \geq 1$ ;

GDI.2) the data are *well-formed* (wfd);

GDI.3) the wfd are *meaningful* (mwfd =  $\delta$ )

(That is, information = data + meaning)

# Floridi's typology of data

δ.1) *primary data*. These are the principal data stored in a database, e.g. a simple array of numbers. They are the data an information-management system is generally designed to convey to the user in the first place.

δ.2) *metadata*. These are secondary indications about the nature of primary data. They describe properties such as location, format, updating, availability, copyright restrictions, and so forth.

δ.3) *operational data*. These are data regarding usage of the data themselves, the operations of the whole data system and the system's performance.

δ.4) *derivative data*. These are data that can be extracted from δ.1.–δ.3, whenever the latter are used as sources in search of patterns, clues, or inferential evidence, e.g., for comparative and quantitative analyses (ideometry).

# Floridi's definition of datum

A datum is reducible to just a lack of uniformity between two signs. So our definition of a datum ( $Dd$ ) is:

$Dd) d = ( x \neq y )$ , where the  $x$  and the  $y$  are two uninterpreted variables.

....

A datum is a relational entity.

....

No information without data representation.

(representation = physical implementation)

# Information and Representation

Rolf Landauer, “The Physical Nature of Information” (1996)

Information is not a disembodied abstract entity; it is always tied to a physical representation. It is represented by engraving on a stone tablet, a spin, a charge, a hole in a punched card, a mark on paper, or some other equivalent. This ties the handling of information to all the possibilities and restrictions of our real physical world, its laws of physics and its storehouse of available parts.

# Back to MacKay...

“The Nomenclature of Information Theory” (1950)

All the results could in a sense show or embody what we believe: they are what we may call *representations*: structures which have at least some *abstract features in common* with something else that they purport to represent. These abstract features of representations are what we want to isolate. ...information can be described as what we depend on for making statements or other representations. More precisely, we may define information in general as that which *justifies representational activity*.

# Edward Fredkin, prophet of information

“Discrete Theoretical Processes” from *A Computable Universe* (2012)

*“Information” State that has Meaning.* In this context *“Information”* is a *configuration* that has *meaning*. Thus *meaning* is a property of the combination of a *configuration* and a *process*. Each *meaning* of a *configuration* is given by the *process* that is interacting with it. A given *configuration* can have more than one *meaning*.

(reported speech from a seminar)

“The meaning of information is given by the processes that interpret it.”

# Norbert Wiener, *Cybernetics* (1948)

We have decided to call the entire field of control and communication theory, whether in the machine or in the animal, by the name *Cybernetics*, which we form from the Greek *χυβερνήτης* or *steersman*. In choosing this term, we wish to recognize that the first significant paper on feedback mechanisms is an article on governors, which was published by Clerk Maxwell in 1868, and that *governor* is derived from a Latin corruption of *χυβερνήτης*. We also wish to refer to the fact that the steering engines of a ship are indeed one of the earliest and best-developed forms of feedback mechanisms.

# N. Katherine Hayles, “Cybernetics” (2010)

Whereas first-order cybernetics was concerned with the flow of information in a system, and second-order cybernetics with interactions between the observer and the system, third-order cybernetics is concerned with how the observer is constructed within social and linguistic environments.

....

Third-order cybernetics redraws the boundary...to locate both the observer and the system within complex, networked, adaptive, and coevolving environments through which information and data are pervasively flowing, a move catalyzed by the rapid development of ubiquitous technologies and mixed reality systems.

# Branches and branchings

Information might be defined as a measure, that is, how much information is conveyed in a statement.

Information might be defined as meaning within a system, that is, as a differential between two quantities and thus a relationality.

Information might also be defined in terms of its embodiments, configurations, and representations. (And objects.)