ESJP
#5 | 2013

# Is Information Out There?

Jasper van den Herik



The concept of information is becoming a central category in the sciences and in society at large. Apart from the rise of information technology, information is used to shed light on all sorts of phenomena, ranging from physics, biology, cognition and perception, epistemology, ontology, ethics to aesthetics: some even argue that the universe itself is an information-processing device. The concept of information is thus changing the way we perceive and evaluate the world and ourselves. De Mul (1999) states that this results in an *informatisation of the worldview*, comparable to the mechanisation of the worldview in the seventeenth century. Yet, the sheer number of applications of the concept of information makes it a 'polysemantic concept' (Floridi, 2013) and a 'notoriously promiscuous term with a marked capacity for dulling critical capacities' (Timpson, 2006: 221).

In this paper, I argue that the failure to distinguish between *information* and *data* lies at the root of much confusion that surrounds the concept of information. Although data are 'out there', i.e. concrete, informational content is abstract and always co-constituted by *information agents* – a set which includes at least linguistically capable human beings. Information is thus not an intrinsic property of concrete data, but rather a relational property, which relies on the existence of information agents.

In part one, I take our ordinary, semantic, conception of language – as something that can inform us – as the explanandum of this paper. I therefore first delineate this concept from the technical notion of information as developed by Shannon. Thereafter, I introduce Floridi's (2013) *General Definition of Information*, wherein information is construed as well-formed meaningful data. Elaborating on this distinction between information and data, I argue, *pace* Floridi, that human-generated information can only be meaningful relative to an information agent who knows how to interpret the data, since the semantic value of the human-generated data is dependent on the horizon of experience of the information agent. The meaningfulness of data is therefore a relational property.

In part two, I broaden the scope and argue that besides human-generated information, environmental information also depends on information agents. Using Hutto and Myin's (2013) *Covariance Doesn't Constitute Content Principle*, I argue that it is not possible to speak of informational content 'out there' as existing independent of information agents. I argue that such a concept of *informational content 'out there'*, could not be causally efficacious, thereby making a description in terms of content superfluous.

In part three, I consider and reject two proposals that do take information to be an objective commodity. The first is Dretske's (1981), which I argue does not succeed in providing an information agent-independent concept of informational content. The second concerns foundational views of information, which make the ontological claim that information is the fundamental ingredient of reality (one can think for instance of Wheeler's 'it from bit', or certain positions in theoretical physics, such as Susskind's idea of the holographic universe). I argue that these accounts trivialise the concept of information by conflating the notions of data and information.

## 1. What is 'Information'?

As noted in the introduction, 'information' as a concept is notoriously polysemantic, pertaining to very different applications. In this section I introduce Floridi's (2005; 2013) data/information distinction, which allows us to get a grip on the slippery concept of information. Thereafter, I argue that human-generated data do not have a semantics independent of an information agent. But first of all, I explicate the difference between our ordinary conception of information and Shannon's technical notion of information.

### 1.1 Two Concepts of Information

When we talk about information, there are different kinds of phenomena we might be interested in. In our everyday use, information has both a *passive* and an *active* connotation. First, we can think of it as something that is 'out there', a commodity or stuff that can be stored and transmitted. For instance, there is information contained on the hard disk of my computer, but this information cannot do anything by itself – it patiently awaits processing. In this sense, information is used as an abstract mass noun (Adriaans, 2012), i.e. it is uncountable and not individuated, like the concrete mass noun 'water'. On the other hand, we also view information as having an informing relation to an information agent[1]. An agent thereby learns, or gets to know, something about the world through this information (De Mul, 1999). Moreover, this implies that information is always about something else, it describes a state of affairs and is hence intentional. In our everyday use of the concept of information, three features therefore seem crucial: '*agents* which represent and use the information, *dynamic events* of information change, and '*aboutness*': the information is always about some relevant described situation or world' (Adriaans & Van Benthem, 2008: 13). Viewed in this way, information has semantic or meaningful content, and allows us to come to know things about the world. Furthermore, it is a *qualitative* concept: it is about what we can come to know about the world, not how much.

Apart from this everyday use, there are rigorous mathematical definitions of information that *do* quantify information. Although these employ the word 'information', this concept of information is distinct from our everyday use of it. The most prominent of these mathematical definitions[2] is the one formulated in the Mathematical Theory of Communication (MTC) (Shannon, 1948). Using this theory, we can calculate the amount of information contained in a message that is transmitted from a sender to a receiver over a communication channel, based on the probabilities that are associated with the different messages that could have been sent. The underlying idea is that messages which are less likely to be sent contain more information. Consider a *unary information source*, which is a source capable of sending only one message. Receiving *this* message is not informative as nothing can be learnt from it[3]. As the possibilities increase, the informativeness of the message also increases. This process can be thought of as a reduction of uncertainty: if I tell you the outcome of a coin toss, supposing the coin is fair, the two possibilities (heads or tails) are reduced to one, namely the one I tell you. But if I tell you about the random placement of a marker on a chessboard, there is a much greater reduction of uncertainty: sixty-four possibilities get reduced to one[4].



It is important to realise that MTC does not specify what the *content* of a message is. It can only tell us about the quantity of information that is transmitted. As long as two possible outcomes are equally likely, just one bit of information is transmitted when we are told about the actual outcome, no matter what the content of this message is. MTC therefore deals with a technical meaning of information that is distinct from the ordinary meaning of the word (Floridi, 2013: 33). One counter-intuitive result of this is that – given the probabilities of the occurrence of letter combinations in English – a page of random letters contains more information than a page of well-formed English sentences, as the probability of the former is lower than that of the latter. Hence, whereas in colloquial speech information is explicitly linked to epistemic notions based on informational content, this is not the case in the more technical notions of information. For the rest of this paper I use information in the broader, everyday sense of the word, as having semantic properties.

Although there is no standard view on how these two notions of information relate, there is widespread agreement that 'MTC provides a rigorous constraint to any further theorising on all the semantic and pragmatic aspects of information' (Ibid.: 48). The *strength* of the constraint, however, is currently a matter of debate. Interpretations of this constraining relation differ from very strong, as for instance mechanical engineering is constrained by Newtonian physics, to very weak, somewhat as playing tennis is constrained by the same Newtonian physics (Ibid.). In the conclusion I briefly return to this constraining relation.

### 1.2 Information and Data

As we have seen, the ordinary notion of information is epistemically related to information agents, who can use information to learn about their world. Information therefore has semantic content: it is about something. But this tells us nothing about what information is and how it is manifested in the world around us. In this paper I follow the *General Definition of Information* (GDI) as expounded by Floridi (2005; 2013), according to which there cannot be information without data. In this section, I briefly introduce this GDI, and the accompanying definition of data.

The general idea behind the distinction between data and information is the formula *data + meaning = information*. Although this distinction is not universally accepted, 'a conceptual analysis must start somewhere' (Floridi, 2013: 3). The GDI is as follows (Ibid.: 7):

> $\sigma$ is an instance of information, understood as semantic content, iff
> *1.* $\sigma$ consists of one or more *data*;
> 2. the data in $\sigma$ are *well-formed*;
> 3. the well-formed data in $\sigma$ are *meaningful*.[5]

The last condition implies that the data under consideration must comply with the semantics of a chosen system, code or language. This meaning, however, does not have to be linguistic, i.e. symbolical, as the referencing relation can also be determined causally or iconically (De Mul, 1999). The condition of well-formedness is syntactical of nature. This syntax also does not have to be linguistic, but must be understood in a broader sense, as what determines the form or structure of something. One can for instance think of the correct ordering of pixels when the informational content is a picture.

The first condition states that information consists of at least one datum. To explain what a datum is, Floridi (2013: 9) gives a *Diaphoric* (from the Greek *diaphora*, 'difference') *Definition of Data* (DDD): 'A datum is a putative fact regarding some difference or lack of uniformity within some context'. This definition, which is very general in nature, can be applied at three levels:

1. Data as *diaphora de re*: as lacks of uniformity in the world out there. As 'fractures in the fabric of being' (Floridi, 2013: 9) they cannot be directly known or experienced, but they can be empirically inferred from experience. They thus serve as an ontological requirement not unlike Kant's *noumena*.
2. Data as *diaphora de signo*: as lacks of uniformity between (the perception of) at least two physical states.
3. Data as *diaphora de dicto*: as lacks of uniformity between two symbols.

Based on different assumptions, *diaphora de re* may be either identical with, or a precondition for *diaphora de signo*, which in turn form a prerequisite for *diaphora de dicto*. For instance, the text you are reading now is based on the *diaphora de dicto* between the letters of the alphabet (they have different shapes), which in turn is made possible by the perceivably different light-reflecting properties of the paper and the ink, which are *diaphora de signo*.

From these two definitions (GDI and DDD) it is evident that information must always be embodied as data, i.e. as lacks of (perceived) uniformity in some medium. Moreover, the DDD allows for a great diversity of classifications, logical types, and realizations of these differences. This means that Floridi's framework is very general in nature, which makes it compatible with different frameworks. This generality is apparent because, according to Floridi (2013: 10), the DDD underdetermines:

- the classification of data (*taxonomic neutrality*);
- the logical type to which the data belong (*typological neutrality*);
- the physical implementation of data (*ontological neutrality*), and
- the dependence of the data's semantics on a producer (*genetic neutrality*).

The fact that Floridi's DDD is neutral with regard to these respects means that the analysis given in this paper does not hinge on any particular view of what could constitute data. In the next section, I briefly introduce the taxonomic and typological neutrality, in which I concur with Floridi.

A more elaborate discussion is needed for the ontological neutrality, as I have to introduce the type/token distinction between data and information – which Floridi does not – in order to discuss the causal efficaciousness of information in the next part. In the last section of this part, I depart from Floridi's framework, when I argue against his idea that data can have a semantics *independently* of any informee (*genetic neutrality*).

### 1.3 The Taxonomic and Typological Neutrality of Data

First of all, the DDD is *taxonomically neutral*. This is because the difference which constitutes the datum is an extrinsic, or relational, property. An example can demonstrate this: take a short burst of sound in a silent context. This sound is only a datum in relation to the silence, which is not only a necessary condition for the burst of sound to be discernible as a datum, but is also constitutive for the [burst-of-sound-in-silence] datum. It is thus the difference between sound and silence that constitutes the datum, not merely the burst of sound itself. This implies that the silence could also be classified as a datum, for this is the other relatum in the [burst-of-sound-in-silence] datum. In other words, nothing is a datum per se. This point is captured in the slogan 'data are relata' (Floridi, 2013: 11). A further example might clarify. In Morse code, long and short beeps constitute the data which allow telegraph operators to send messages. However, it would be possible to have a continuous tone with long and short interruptions to transmit messages in Morse code. In the latter case, it would be the silences that are the data. Similarly, there could be data that are not classified as such, as would be the case if the beeps that are used to transmit Morse code differ in volume. Although there would be additional data in the message (differences in volume of the beeps), we need not classify these as data.

Secondly, the *typological neutrality* states that information can consist of different types of data as relata (Floridi, 2013: 11). Most of the time, when we talk about data we mean *primary data*. These are the data that an artefact is designed to convey. We could for example think of the position of the hands of a clock informing us about the time. But the absence of data may also be informative, for instance when you ask a person if she is sleeping, and she does not answer. The fact that you do not get a response could still answer your question. Floridi coins these *secondary data*. Furthermore, we can often infer a lot more from primary data than just what they are meant to convey. If I ask a person whether he knows the way to the park and he gives me an answer, I do not only learn the route to the park, but I also come to know that he speaks English. This is a form of *derivative data*, which are created accidentally when we try to convey primary data. Lastly, there is information that concerns other data. *Meta-data* are data about other data, informing us for instance of the type of data. *Operational data* are data regarding the operations of a data system. For example, when your computer tells you there is an error, this prevents you from taking the primary data it produces at face value.



### 1.4 Ontological Neutrality: Information as an Abstract Type

As we have seen, information relies on the existence of data. The ontological neutrality states that the DDD is neutral with respect to the ontological realization of the data. This confirms our common-sense intuition that the same sentence, whether written on paper or encoded in binary and stored on a computer, contains the *same* information. Therefore, the medium, format and language do not influence the information contained in a message. The differing realisations could of course convey different secondary or derivative data, but from the perspective of the primary data, the realisation does not matter.

The ontological neutrality thus further implies that there is a type/token distinction between the information and the data it is realised in (Timpson, 2006). To explain how this works, we consider sending a message in the vocabulary of MTC. In order to send a message, the sender has to select elements from a fixed alphabet, say $\{a_1, a_2, ..., a_n\}$, and transmit them over a communication channel. Now suppose we want to send the number '42' to a receiver. We can do this using many different media: we could send him a piece of paper, an electronic message, or simply tell him the numbers directly. Now it is easy to see that the tokens would be very different in each case, ranging from scribbly lines ('42'), to bits transmitted as voltage differences along a copper wire, to complex vibrations in the air.

For those of us who speak English and are accustomed to using Arabic numerals to denote numbers, the three messages would convey the same type, i.e. the same informational content. The information that is represented by the type is therefore abstract. This implies that, being an abstract entity, the information itself has no spatio-temporal location, nor is it part of the material contents of the world. The tokens which realise these types, on the other hand, *do* have a spatiotemporal location. *Prima facie*, this seems like a denial of the objective existence of information, especially if you do not like abstracta in your ontology. But any talk of abstracta can easily be 'paraphrased away as talk of obtaining facts about whether or not concrete types would or wouldn't be instances of types' (Timpson, 2006: 228). This does not entail that information has no objective existence, or cannot be an objective commodity. But it does suggest that any talk of information, rather than of data, causing anything, has to be worded carefully. For different tokens (data), although they might realise the exact same type (information), might have very different effects in the world around us. Dretske (1989) gives us a clear example of this: consider a soprano, who sings a high note, thereby shattering a glass. If the token would be altered only slightly, for instance by singing a semitone lower, the glass would not have broken, whereas the informational content (the meaning of the words that the soprano is singing) would be identical. It is therefore, from the viewpoint of information, a contingent property of the token that causes the glass to break. However, when we are asking what the soprano is singing *about*, we are not interested in these contingent properties, but in the semantic content of the sounds she is producing. In this case, what we are asking for is the type, not the token. When I ask someone the question: 'What number is written on this piece of paper?', I want to be informed about the type, that is the number, that is realised by this particular token. We can think of this kind of ostensive acts as *deferred ostension* (Quine, 1969).

*Prima facie*, this implies that in order for the informational content to be causally efficacious, there has to be an information agent that, in one way or another, recognises the type, rather than the token. Before I analyse how this view on informational content relates to information 'out there' in the following part, I first argue that the type/token distinction between informational content and the data by which this content is realised implies that the informational content cannot be thought to exist *independently* of an information agent who co-constitutes this content.

### 1.5 Against Genetic Neutrality: the Meaninglessness of Data in the Absence of Information Agents

Genetic neutrality is the idea that 'data (as relata) can have a semantics *independently* of any informee' (Floridi, 2013: 17). This is not meant to be a thesis about how data can acquire a meaning in a semiotic system, but rather about how data can be thought of as meaningful independent of an informee. The example that Floridi (2013: 18) gives are Egyptian hieroglyphs, that, before the discovery of the Rosetta Stone, were incomprehensible. Even though there was a time when we did not know what their meaning was, there was a meaning hidden in these symbols – if we are to take Floridi's thesis at face value. This example deserves further analysis, especially considering the important role that information agents play, as we have seen in the last section.



The first observation that is relevant here is that when we study ancient texts, 'we do not "see" the meaning as a feint [sic] aura around the characters' (Hansen, 1985: 492). It is not the case that Egyptian hieroglyphs contain an objective meaning hidden within them, which can be made visible by acquiring the ability to interpret hieroglyphs. For 'the semantic value of information is dependent on the horizon of experience – or speaking hermeneutically – the *world* of the user' (De Mul, 1999: 81). In trying to understand the meaning of the hieroglyphs, we are not engaged in a theoretical reconstruction, for this is an illusion which can only be a regulative idea or a methodological idealisation (De Mul, 1993: 13). This implies that meaning cannot be an objective property of data as relata. Although the information contained in the data might *prima facie* seem to be well-formed and meaningful, this does not imply that they are actually meaningful. An example might illustrate this point.

The Voynich Manuscript, a book carbon dated to the early fifteenth century, is written entirely in an as of yet undeciphered script. Although the script shares many informational characteristics with European languages (it has for instance about 20-30 characters and a word entropy[6] of 10 bits (Landini, 2001)), its resistance against deciphering makes the *attribution* of a semantics speculative. It remains unclear whether a 'Rosetta Stone' will, or even *could*, ever be found for this manuscript. So we are now in the same position with regard to the Voynich Manuscript that we were in with regard to Egyptian hieroglyphs before the discovery of the Rosetta Stone. Both texts surely seemed to be meaningful to us, but whether they actually *do* possess a semantics was unknown – and remains unknown for the Voynich Manuscript. We can thus only say that the script carries meaning, when we are able to decipher it. In other words, if the Rosetta Stone did not exist (assuming for now that there would be no other way of deciphering hieroglyphs), the meaning of the hieroglyphs would have been lost forever.

But examples of this can also be found closer to home. Think for instance of the data that are on your hard disk. These data are encoded in a very particular way, based on convention. For instance, text can be encoded in ASCII (American Standard Code for Information Interchange) format. In this format, the letter 'A' is represented by the binary code '1000001', whereas the 'a' is encoded as '1100001'. It should be clear that in the absence of the ASCII decoding manual, the strings of ones and zeros would be unintelligible to most English speakers. So if there were no way of decoding them, the strings of ones and zeros would contain no information. Consider for instance that a person comes up with his own version of ASCII code, randomly switching around the encodings for the different letters. If he were to leave us a short message which we only found after his death, the data would be meaningless to us. And since they were only meaningful to one person, who no longer exists, it seems unclear what it would mean to claim that the information is still in there. The information is lost forever, independent of the fact whether the message was intended to carry information or not.

These examples, however, do not show that certain data cannot seem to be meaningful to us before we can attribute meaning to it. The reason why a lot of people try to decipher the Voynich Manuscript, and before that, hieroglyphs, is that they *seem* to be meaningful. However, a distinction has to be made between merely seeming to be meaningful and actually being meaningful. A wonderful example of this can be found in the *Codex Seraphinianus* (Serafini, 1981), an illustrated encyclopedia of an imaginary, surreal world. Like the Voynich Manuscript, it is written in a strange script, and similarly, attracted a lot of attention from people, who tried to decipher it. However, in 2009 Serafini announced that the script was asemic (Stanley, 2010), so we can know for sure that the script does not carry meaning. Although it seems unlikely, the same could have been true for the Egyptian hieroglyphs. The hieroglyphs could have turned out to be *asemic*, i.e. have no semantic content – they could have been merely decorative, carrying no information. From this we can conclude that *seeming to be meaningful* does not imply *meaningfulness*, although of course it could warrant us to try to decipher a text.

The idea expressed in the two examples given is that having-a-semantics, just as being-a-datum, is a relational property. It is therefore unclear what the genetic neutrality is meant to express, as we would be unable to verify its correctness: either we can interpret the text, in which case the semantics is not independent of an informee but depends equally on the interpreted and (the horizon of experience of) the interpreter, or we cannot interpret the text, in which case we cannot know *whether* the data under consideration could have a semantics. Moreover, in the former case the actual semantics that is attributed to the data in question is constitutively dependent on the information agent. An illuminating example of this is given by De Mul (1999: 81): 'A symptom that provides the doctor with valuable information for the determination of a diagnosis can be meaningless, or have a very different meaning, to the patient'. De Mul concludes from this remark that 'the same information [better: data] can give rise to different forms of knowledge and action' (Ibid.). Here the distinction between informational content and data can help us make sense of this: although both the doctor and the patient have access to the same data (the symptom), the informational content it provides them with is surely different. It is true that the data provide the doctor with valuable information, but his medical background knowledge in this case is constitutive for the information. If we give up on the intrinsic meaningfulness of data as relata, we can see that the data are not meaningful for the patient, whereas they are meaningful for the doctor. As meaningfulness is the third condition for information in the GDI, the symptom thus has informational content *for* the doctor, that it does not have *for* the patient. This is consistent with saying that although the doctor is informed by the symptom about the patient's particular ailment, the very same symptom *does not* inform the patient about his ailment. I would therefore like to modify the definition of genetic neutrality in order to incorporate this necessary relation: data (as relata) can *seem* to have a semantics independently of any informee; but the informational content is always constituted in the relation between the data and an information agent.

## 2. The Agent-Dependency of Information Content Out There

In the first part I have considered human-generated information, and argued that informational content in those cases is dependent on information agents. In this part, I argue that the same applies to environmental data. Although cognition is often thought of as essentially information-processing, this view has recently come under attack by a new paradigm in the cognitive sciences. *Enactivism*, as introduced in by Varela, Rosch and Thompson (1991), is opposed to the cognitivist idea of the information-processing brain as being sufficient for cognition. In the introduction to the edited volume called *Enaction – Toward a New Paradigm for Cognitive Science*, which aims to collect these new lines of thought and show how they deal with numerous aspects of cognition, John Stewart (2010: vii) states that '[t]his program makes a radical break with the formalisms of information-processing and symbolic representations prevalent in cognitive science.'

In their *Radicalizing Enactivism* Hutto and Myin (2013) claim that this enactivist paradigm should be radicalised by denying that informational content can be an explanatory concept in studying basic cognition, which includes, *inter alia*, perceptual processes and their intentionality and phenomenality, and emotional responding. Starting from the idea that 'the vast sea of what humans do and experience is best understood by appealing to dynamically unfolding, situated embodied interactions and engagements with worldly offerings' (Hutto & Myin, 2013:ix), they develop an account of basic cognition which has no need for mental content, where they define content as truth-bearing properties or specified conditions of satisfaction. Moreover, they claim that any theorist who does claim that cognition necessarily involves content must face up to the *Hard Problem of Content*, which is to explain the existence of content in a naturalistically respectable way. For if there is no informational content in nature, then 'cognitive systems don't literally traffic in informational content'[7] (Ibid.: xv). If anything, cognition can be thought of as content-*creating* rather than content-*consuming* (Ibid.: 76).

### 2.1 Covariance and Content

Hutto and Myin (2013) start from the assumption that information as *covariance* is the only scientifically respectable notion of information. Floridi (2013) seems to agree when he talks about environmental information, although he already relates the information to an information agent. He states that environmental information can be defined as follows: '[t]wo systems *a* and *b* are coupled in such a way that *a*'s being (of type, or in state) *F* is correlated to *b* being (of type, or in state) *G*, thus carrying *for the information agent* the information that *b* is *G*' (Floridi, 2013: 19, emphasis added). But if we want to have an account of informational content that can get basic cognition up and running, the content has to exist independently of anyone using the content. The informational content has to be able to be 'retrieved, picked up, fused, bounded up, integrated, brought together, stored, used for later processing, and so on and so forth' (Hutto & Myin, 2013). This problem of defining content naturalistically is what Hutto and Myin call the Hard Problem of Content.

For content has to have special properties to be properly called *content*. It has to have truth-bearing properties. In order to have these properties, content has to 'say' or 'convey' something about something else. Take a simple example: the number of tree rings can covary with the age of the tree, but by themselves the tree rings do not say or convey anything about the age of the tree, i.e., we can not meaningfully say that the tree rings are 'false', if for one reason or another they do not covary with the age of the tree. This is the *Covariance Doesn't Constitute Content Principle*, which implies the Hard Problem of Content: if covariance does not constitute content, we need a more elaborate story that explains how cognition can come to be contentful. Hutto and Myin use a slightly different terminology to separate (informational) content from the processes underlying it than I have used so far[8]. Instead of making a distinction between data and information, they make a distinction between a vehicle and its content. They argue that, if we accept the Covariance Doesn't Constitute Content Principle, the vehicle/content distinction falls apart at this level, which means we would be left with just the vehicle (Hutto & Myin, 2013: 68). Or, if we use the data/information distinction, we would be left with just the data, as there would be no informational content. In the next Section, I argue that, even if we did allow covariance to constitute content, a description in terms of information would not further our explanation of causal processes in the absence of information agents.

### 2.2 The Causal Efficaciousness of Informational Content in the Absence of Information Agents

A first stab at thinking about the *causality* of informational content – and its relation to covariance – thus conceived can be formulated by using a very simple example: a thermostat. For simplicity, let us assume that there are only two possible states in the environment, either too cold ($E_c$), or warm enough ($E_w$). The bimetal in the thermostat can then be either in a bent state ($B_b$) when it is too cold, or in a straight state ($B_s$) when it is warm enough. If the bimetal is bent, it will close a circuit, thereby turning on the heater ($H_{on}$), whereas if the bimetal is not bent, the circuit will be open, thereby turning off the heater ($H_{off}$). Suppose we further allow – for now – that because of the lawful covariance between the bending of the bimetal and the ambient temperature, the bimetal contains information about the temperature, and thus that covariance *does* constitute content. Whether or not the bimetal is bent will serve here as the datum *de signo*, realizing the information. Call this information either $I_B(c)$ or $I_B(w)$, where the subscript serves to designate the datum (either $B_b$ or $B_s$) under consideration, and the value between brackets specifies the ambient temperature. The status of the heater can be said to covary in the same manner with the temperature in the room, realizing the information $I_H(c)$ and $I_H(w)$.

For reasons of simplicity, we have limited the number of states the total system can be in to two discrete states[9]. Now the two states of the system can be schematically visualised, with the horizontal arrows indicating causal relations, and the vertical arrows indicating the realising relation:

$$\begin{array}{lccccc}
\text{Information (abstract)} & & & I_B(c) & & I_H(c) \\
\cdots\cdots\cdots\cdots\cdots\cdots & & & \uparrow & & \uparrow \\
\text{Data (concrete)} & E_c & \rightarrow & B_b & \rightarrow & H_{on}
\end{array}$$

*Diagram 1: too cold*

$$\begin{array}{lccccc}
\text{Information (abstract)} & & & I_B(w) & & I_H(w) \\
\cdots\cdots\cdots\cdots\cdots\cdots & & & \uparrow & & \uparrow \\
\text{Data (concrete)} & E_w & \rightarrow & B_s & \rightarrow & H_{off}
\end{array}$$

*Diagram 2: warm enough*

From these diagrams, we can easily see that once the causal story has been told, the informational states that are assumed to be realised by the bimetal and the status of the heater – based on the covariance relation that obtains between them and the environment – are superfluous[10]. In other words, once the causal story at the level of the concrete data has been told, there is nothing left to explain[11]. The concept of information is simply not needed to explain the workings of the thermostat.

This analysis is further corroborated when we analyse a possible way in which the workings of the thermostat might be interrupted: suppose that some properties of the metals of which the bimetal is composed changed, thereby transforming its bending behaviour. This might lead to a situation in which the bimetal does not close the circuit when the ambient temperature is too cold, whilst it might – based on the idea that covariance *does* constitute informational content – still contain information about the temperature because the bending of the bimetal still covaries with the ambient temperature. It is therefore not the information-carrying role that allows the intended working of the thermostat, but the – from the viewpoint of the information – contingent physical properties of the token that realises that information. This implies that even if we were to allow that covariance does constitute content, the alleged content would be causally superfluous. In other words, covariance by itself suffices to explain the workings of the thermostat.

We can thus conclude that the assumption of causal efficaciousness of information in inanimate systems is problematic because of the abstract nature of information. In the absence of information agents, it seems not to be the information, i.e. the abstract type, but rather the data, or concrete tokens which realise the information, that are causally active. It is only in the case when an information agent *recognises* a particular token to be a token of a particular type, that the informational content comes into existence and can become causally active. As we have seen, when a human being would point to a piece of bent bimetal, given enough background knowledge, she would point at the type through an act of deferred ostension ('look how warm it is'). The crucial phrase in the last sentence is 'given enough background knowledge'. The bimetal-as-datum only contains the information that it is either too cold or warm enough in relation to an information agent that already knows about the covariance relation that obtains between the bimetal and the ambient temperature.

## 3. Possible Defences of Agent-Independent Informational Content

The above analysis leaves the defenders of content with three possible responses to the Hard Problem of Content. First, they might posit informational content as an extra element of reality, not unlike how Chalmers (e.g. 1995) tries to solve the problem of phenomenal experience in a functionalist philosophy of mind by positing the existence of qualia. This, however, changes the way we look at information radically, leaving naturalistic accounts the task of finding fundamental bridging laws between covariance relations in the world and informational content (Hutto & Myin, 2013: 69). Moreover, this move leaves defenders of informational content with additional problems to solve. If the informational content is indeed an extra element of reality this introduces (1) epistemic problems: how do we get to know these informational contents if they are ontologically distinct from the causal processes which affect us; and (2) overdetermination problems: if we were to think of the informational contents as extra elements of reality, we would have secured their objective existence, but then we would still need to explain how they can be causally efficacious, as we have seen in the last section. Although this manoeuvre might be the only way to solve the Hard Problem of Content (Ibid.), it is most certainly not a panacea, and the metaphysical costs will be too high.

Second, the notion of informational content might be thought of as meatier than covariance, whilst retaining naturalistic respectability. The most prominent proposal along these lines is given by Dretske (1981), who thinks of informational content as having an *indicating relation* to some state of affairs, thereby realizing truth-bearing properties – that is, content – in an objective world. In the next section I take a closer look at Dretske's account, arguing that it does not succeed in defending this objective, information-agent independent, existence of informational content.

Third, the distinction between information and data (or vehicles and content) might be denied, thereby reducing the concept of information to the concept of data. In the last section of this part, I argue that this trivialises the concept of information, thereby adding to the confusion that surrounds the concept of information.

### 3.1 Dretske on Information as an Objective Commodity

'In the beginning there was information. The word came later.' (Dretske, 1981: vii). These opening lines of Dretske's book on information clearly show his ambition. Information is to be thought of as an objective commodity, whose existence pre-dates, and is independent of, the existence of information agents. This ambition is further developed in the second paragraph of the book, where Dretske explicitly opposes the view that 'something only *becomes* information when it is assigned some significance, interpreted as a sign, by some cognitive agent' (Ibid.), a variant of which I am defending in this paper. But *prima facie*, this ambition is not visible in his definition of information, as the background knowledge of the information agent (denoted by the variable *k*) is explicitly mentioned in it: '*Informational content:* A signal *r* carries the information that *s* is *F* = [sic] The conditional probability of *s*'s being *F*, given *r* (and *k*), is 1 (but, given *k* alone, less than 1)' (Dretske, 1981: 65). That this background knowledge is constitutive of the informational content that a signal carries is further underlined in one of the examples that Dretske uses.

Dretske asks us to suppose that there are four shells, with a peanut located under one of them (Dretske, 1981: 78). Suppose further that person *a* knows that the peanut is not under either shell 1 or 2, whilst person *b* has no knowledge of the location of the peanut at all. If both person *a* and *b* now get the information that the peanut is not under shell 3, this observation of course allows person *a* to work out that the peanut is under shell 4, whereas person *b* is still unaware of the location of the peanut. After considering both the option that for person a the observation only carries the information that the peanut is not under shell 3, and the option that this observation additionally also carries the information for person a that the peanut is under shell 4, Dretske decides on the latter: 'the third observation supplies [person *a*] with the information that shell 3 is empty and the information that the peanut is under shell 4. The latter piece of information is (for [person *a*]) nested in the former piece of information. For [person *b*] it is not' (Dretske, 1981: 79). So the informational content contained in the same signal differs depending on the background knowledge of the person who receives that signal.

This seems to be in direct opposition to the idea that information is out there. Dretske's solution, which allows him to hold both that information is out there and that the informational content of a signal is dependent on the background knowledge of the information agent, is the recursive character of his definition. The background knowledge can be explained itself in terms of information received earlier, until 'eventually we reach the point where the information received does not depend on any prior knowledge' (Dretske, 1981: 87). At first sight, however, it is not obvious that all knowledge can be recursively based on these foundational cases (Alston, 1983). Moreover, Dretske does not provide a way in which the probability of these foundational cases of information extraction from the environment could be one, as is required by his own definition (Levi, 1983). So unless Dretske's account is supplemented with a valid description of how we, as *tabulae rasae*, might – based solely on a signal *r* – know that the conditional probability of *s* being *F* is 1, the informational content Dretske is talking about is always relative to the background knowledge of an information agent. In other words, an information agent has to know the probabilities attached to the possible signals that a source could send before she can know the informational content that a particular signal carries (Moor, 1982: 238). Moreover, even if this problem were to be solved, this would only prove the objective existence of these foundational cases of information. The majority of the informational content 'picked up' from the environment would still be co-constituted by the background knowledge. Barwise (1983: 65) acknowledges this point when he states that although 'information is out there, it *informs* only those attuned to the relations that allow its flow'. In the terminology of Dretske, we could translate this by saying that although the signals are out there, the informational content they carry is always relative to an information agent. And this just amounts to saying that *data* are out there, but *information* is always relative to an information agent.

### 3.2 Foundational Accounts of Information

We can use Shannon's Mathematical Theory of Communication (MTC) to calculate the average amount of information that a system transmits[12]. This measure is also called the entropy of the information source (Adriaans, 2012: 15). Entropy is a measure that, prior to the rise of MTC, was already widely used in thermodynamics, of which the second law states that the entropy of isolated systems can never decrease, because isolated systems evolve to a state of maximal entropy. Entropy is therefore often associated with *disorder*, although *randomness* would be a better term as it is a syntactical, not a semantic notion (Floridi, 2013: 37). The concept of entropy therefore connects thermodynamics to information theory. In the words of Adriaans and Van Benthem (2008: 8): 'information theory is the thermodynamics of code strings, while thermodynamics is the information theory of particles in space'. Because in quantum mechanics, information turns out to be discrete instead of continuous, any physical system could in principle be described by a finite amount of information. This analogy can be taken to the extreme, in the claim that the universe is ultimately a computational system, with information being the most basic ingredient.

According to theoretical physicist Susskind, for instance, the idea that information never disappears is *the* most fundamental principle of physics (Susskind & Lindesay, 2005). The concept of information he is referring to here is that of fundamental distinctions between things: 'Information means distinctions between things. A hydrogen atom is not an oxygen atom, an oxygen atom is not a hydrogen atom' (World Science Festival, 2011[13:30]). Physicist and mathematician Brian Greene states: 'Every object in some sense contains information, because it contains a very specific arrangement of particles' (World Science Festival, 2011[9:20]). From this kind of observations, one might conclude that information is the most basic ingredient of reality, and that space and time, matter and energy, are merely derivative notions[13]. Wheeler (1990) coined this idea 'it from bit' (see also Schmidhuber (1997) and Lloyd & Ng (2004) for similar accounts). I shall refer to accounts like these as foundational accounts of information.

*Prima facie*, if we take these accounts seriously, it seems that information is out there after all. But on second thought, this view on information is more akin to Floridi's DDD. It just states that the world ultimately consists of lacks of uniformity 'out there', the *diaphora de re* mentioned earlier. Floridi (2013: 16) can therefore state that the GDI is neutral with regard to these foundational accounts of information. What is important to realise here, is that these accounts do not give us any hints on how one state of affairs could carry information about another state of affairs. Strictly speaking, things would only carry information about themselves. Taking information to be fundamental in this way thus reduces the concept of 'information' to that of 'data'. Foundational accounts of information thus trivialise the concept of information. Quite literally everything becomes information if we regard information as *diaphora de re*. It should hardly come as a surprise that the world is full of differences. Everything, from a rock rolling down a hill, to a lone atom traversing the interstellar void, to the universe itself, becomes an information processing entity. Moreover, this conception of information actually negates the common-sense idea that information could be realised in different ways, for if two situations differ, so will their informational content. It would therefore no longer be possible to say that two different tokens of the same type would contain the same information.

Finally, on closer inspection, foundational accounts of information turn out to be irrelevant to the question asked in this paper, that is, what semantic information is. For the *diaphora de re* that these accounts take to be the fundamental ingredient of reality are not directly perceivable by information agents, whilst the data to which they *do* attribute semantic properties can only be the *diaphora de signo*, which are perceivable. And whether these *diaphora de signo* ultimately consist of diaphora de re, particles or fields of energy is simply irrelevant to the question of how we can attribute meaning to them. Even if we were to accept the view that information is foundational in this sense, we would need a new concept to differentiate our ability of information processing from any other physical process. It therefore seems better to take these foundational accounts of information to be talking about data as being fundamental, reserving the concept of information for the role specified in this paper.



## 4. Conclusion

In this paper, I argued that in the beginning there were data and information came later. This distinction between data and information can be helpful to differentiate between two concepts that are fundamentally different, but are now often conflated. Because the analysis of information given in this paper relies on Floridi's *General Definition of Information* and the accompanying *Diaphoric Definition of Data* – which is taxonomically, typologically and ontologically neutral – it is consistent with a large variety of theories about what these data could be. In relying on the formula *data + meaning = information*, the analysis in this paper therefore gives a general framework that could be adapted and worked out, for instance based on one's ontological views.

Much of the confusion that surrounds the concept of information can be traced, I think, to the fact that the use of the word ‘information’ carries connotations from our everyday, semantic use of the word to applications where these semantic properties do not exist. If the aim of a certain theory or field is not to talk about the semantic properties of data, the usage of ‘information’ can almost certainly be replaced with ‘data’. Because the concept of data does not carry these semantic connotations, this would clear some of the confusion. If we think back to Shannon’s Mathematical Theory of Communication, for instance, it seems that it would not lose any explanatory power if we take it to be about the communication of *data*, rather than of *information*. Rather than being about information per se, the MTC only weakly constrains theories about information because, as we have seen, information is always necessarily embodied as data.

Realizing that although data *are* out there, informational content is always co-constituted by information agents, therefore allows us to see that information cannot be the fundamental ingredient of reality, as it is a relational property that exists between the data (which might turn out to be foundational) and the informational agent. Only when data become meaningful for an agent – when they come to have informational content by acquiring conditions of satisfaction – can an explanation in terms of information add anything to a causal explanation. For only the abstract informational content can explain how an information agent might react similarly to different tokens which consist of concrete data, which could have very different physical properties.

If we were to reserve the word ‘information’ for informational content in this sense, and use the word ‘data’ when we mean differences that are ‘out there’, at least some of the confusion that surrounds the polysemantic concept of ‘information’ would dissolve.

## Acknowledgments

I would like to thank prof. dr. Jos de Mul for his insightful comments on the first version of this paper. Furthermore, the comments of three anonymous referees have contributed to substantial improvements of this paper.

*Jasper van den Herik (1986) is currently writing his master thesis on the role of language in Hutto’s radical enactivism. Earlier he finished a Bachelor’s degree in Psychology at the Erasmus University Rotterdam. His primary research interest is the intersection of post-cognitivist philosophy of mind and anti-representational views of language.*

*‘Is Information Out There?’ was written for the master course ‘Filosofie van de geesteswetenschappen: Actuele thema’s van de hermeneutiek’ taught by prof. dr. Jos de Mul.*



## Notes

1. Because of the distinction between data and information, it is not the case that any *agent* is necessarily also an *information agent*. For instance, simple organisms can be sensitive to and act upon data from their environment, whilst not relying on informational content for their agency (as is apparent from Hutto & Myin’s (2013) Hard Problem of Content which is discussed in section 2). If we follow Hutto and Myin (2013) this label is only reserved for creatures who have an enculturated, scaffolded mind, i.e. who have linguistic capabilities. Others might attribute these content-generating capabilities to much lower forms of cognition, as in for instance the teleosemantics of Millikan (1984). For this paper I assume that at least linguistically capable human beings are information agents. The question whether other agents can also be information agents will have to be answered, but falls outside the scope of this paper.

2. Apart from Shannon-information, there are also other mathematical definitions that quantify information, like Kolmogorov complexity, Fisher information and Quantum information (Adriaans, 2012). As Shannon-information is the most widely used conception in philosophy, and it focusses on information transfer, I will only discuss this particular technical notion in this paper.

3. It has to be noted that MTC presupposes that the possible messages and the associated probabilities are known in advance.

4. Shannon gives the amount of information contained in a single message, for reasons that I will not go in here, as the negative $\log_2$ of the probability of that message occuring. This implies that a fair coin toss generates one bit of information, while the random placement of a marker on a chessboard generates six bits of information. The bits can be thought of as the amount of yes/no questions that have to be answered before the answer is reached. In the case of the coin this is one question (‘is it heads?’), whereas the position of the marker on the chessboard can be determined with six yes/no questions.

5. Floridi (2005) argues that a fourth condition has to be added, according to which the well-formed, meaningful data have to be truthful. In this paper I will try to steer clear of issues concerning truth(fulness), so I will not include it in the definition. The argument in this paper would, I think, not change depending on whether or not truthfulness is a necessary condition for information.

6. The word entropy specifies the amount of information given by the occurrence of that word, based on the probability of the word occurring.

7. Althought I am sympathetic to their project, in this chapter I merely wish to argue that the existence of informational content is dependent on users of this content, that is, informational content only arises when cognitive processes are in play. The stronger claim, that basic cognition could be explained without any appeal to content, lies outside the scope of this paper. Some commentators think that Hutto and Myin are not radical enough, because they take linguistic cognition – or in their terms 'enculturated, scaffolded minds' (Hutto & Myin, 2013: vii) – to be contentful, without telling a convincing story of how this content arises from the basic cognitive processes that on which the linguistic mind is built atop. See for instance Roberts (2013) for this critique.

8. The distinction between data and information could however, I believe, strengthen the account of Hutto and Myin. After they have concluded that basic cognition is not contentful, they state that '[we] can still endorse the idea that organisms are informationally sensitive (i.e., that they exploit correspondences in their environments to adaptively guide their actions) while denying that it follows that they take in, store, or process informational content' (Hutto & Myin, 2013: 82). If they were to accept the information/data distinction, we would see that organisms would not be *informationally* sensitive, but rather be sensitive to data. They would thereby be able to fend off attacks on their position, which could state that this still implies that this informational sensitivity implies information-processing in a weaker sense.

9. Extending the example to more or continuous states does not change the conclusion reached here, but would needlessly complicate matters.

10. This argument is inspired by the objection based on causal closure and overdetermination that Jaegwon-Kim (1998) gives against non-reductive physicalist accounts of the mental.

11. At this point, it might be protested that the bimetal only *carries* the information about the temperature in virtue of being bent. Dretske puts forward a proposal along these lines: 'When, therefore, a signal carries the information that *s* is *F* in virtue of having property *F'* [that the room is too cold in virtue of being bent], when it is the signal's being *F'* that carries the information, then (and only then) will we say that the information that *s* is *F* causes whatever the signal's being *F'* causes' (Dretske, 1981: 87). However, this does not yet show that it is the information that is causally efficacious. In the words of Rundle (1983: 78): 'rather, it amounts to a proposal to speak as if the information has this role when its carrier does. However, since the latter does give us a genuine cause. there is no way of pressing the objection that confronts the usual causal theories'.

12. The formula for calculating this for a system of possible messages $A$ is $H(P) = - \sum_{(i \in A)} p_i \log_2 p_i$. This means that we take the average of the information contained in all messages that are a member of communication system $A$, i.e. the possible message that could be sent, by summing the amount of information contained in each message ($\log_2 p_i$), correcting for the chance of them occuring.

13. Timpson (2006) reminds us that the fact that a process in reality is accurately describable in terms of the information it contains, does not necessitate us to view this information as foundational. There might still be some material substrate that realises these fundamental differences. Both interpretations produce the same outcomes in experimental settings.

## References


Adriaans, P. (2012) 'Information'. In: *The Stanford Encyclopedia of Philosophy (Winter 2012 Edition)*, Edward N. Zalta (ed.), http://plato.stanford.edu/archives/win2012/entries/information/ [page numbers refer to the PDF version].

Adriaans, P. and Benthem, J. van (2008) 'Introduction: Information is what Information does'. In: P. Adriaans and J. van Benthem (Eds.) *Handbook of the Philosophy of Science. Volume 8: Philosophy of Information*. Elseviers Science Publishers, 5-28.

Alston, W.P. (1983) 'Dretske on knowledge'. In: *The Behavioral and Brain Sciences* 6, 63-64.

Barwise, J. (1983) 'Information and semantics'. In: *The Behavioral and Brain Sciences*, 6, 65.

Chalmers, D. (1995) 'Absent Qualia, Fading Qualia, Dancing Qualia'. In: T. Metzinger (Ed.) *Conscious Experience*. Exeter: Imprint Academic, 309-330.

Dretske, F. (1981) *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.

Dretske, F. (1989) 'Reasons and Causes'. In: *Philosophical Perspectives* 3, 1–15.

Floridi, L. (2005) 'Is Information Meaningful Data?'. In: *Philosophical and Phenomenological Research* 70 (2), 351-70.

Floridi, L. (2013) 'Semantic Conceptions of Information'. In: *The Stanford Encyclopedia of Philosophy (Spring 2013 Edition)*, Edward N. Zalta (ed.), http://plato.stanford.edu/archives/spr2013/entries/information-semantic/ [page numbers refer to the PDF version].

Hansen, C. (1985) 'Chinese Language, Chinese Philosphy, and "Truth"'. In: *Journal of Asian Studies* XLIV (3), 491-518.

Hutto, D. D. & Myin, E. (2013) *Radicalizing Enactivism: Basic Minds without Content*. Cambridge, MA: The MIT Press.

Kim, J. (1998) *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA.: MIT Press.

Landini, G. (2001) 'Evidence of linguistic structure in the Voynich manuscript using spectral analysis'. In: *Cryptologia* 25 (4), 275–295.

Levi, I. (1983) 'Information and error'. In: *The Behavioral and Brain Sciences* 6, 74-75.

Lloyd, S. and Ng, J. (2004) 'Black Hole Computers'. In: *Scientific American* 291 (5), 30-9.

Moor, J. (1982) 'Knowledge and the Flow of Information'. In: *Philosophical Books* 23 (4): 237-39.

Millikan, R. (1984) *Language, Thought and Other Biological Categories*. Cambridge, MA.: MIT Press.

Mul, J. de (1999) 'The Informatization of the Worldview'. In: *Information, Communication & Society* 2 (1), 69-94.

Quine, W. V. (1968) 'Ontological relativity'. In: *Journal of Philosophy*, 65, 185-212.

Stewart, J (2010) 'Introduction'. In: J. Stewart, O. Gapenne & E. Di Paolo (Eds.) *Enaction: Towards a New Paradigm for Cognitive Science*. Cambridge, Mass.: MIT Press, vii-xvii.

Roberts, T. (2013) 'Radicalizing Enactivism: Basic Minds Without Content'. In: *Notre Dame Philosophical Reviews*. Retrieved from: http://ndpr.nd.edu/news/40035-radicalizing-enactivism-basic-minds-without-content/ [5 July 2013]

Rundle, B. (1983) 'The sufficiency of information-caused belief for knowledge'. In: *The Behavioral and Brain Sciences* 6, 78.

Schmidhuber, J. (1997) 'A Computer Scientist's View of Life, the Universe, and Everything'. In: C. Freska, M. Jantzen and R. Valk (Eds.) *Lecture notes in computer science: Vol. 1337. Foundations of computer science: potential – theory – cognition*. Berlin: Springer, 201-8.

Serafini, L. (1981) *Codex Seraphinianus*. Milan: Franco Maria Ricci.

Shannon, C.E. (1948) 'A Mathematical Theory of Communication'. In: *Bell System Technical Journal* 27, 379–423; 623-56.

Shannon, C.E. and Weaver, W. (1949) *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.

Stanley, J. (2010) *To Read Images Not Words: Computer-Aided Analysis of the Handwriting in the Codex Seraphinianus*, (unpublished MSc dissertation). North Carolina State University at Raleigh. Retrieved from http://repository.lib.ncsu.edu/ir/bitstream/1840.16/6460/1/etd.pdf [15 november 2013].

Susskind, L. and Lindesay, J. (2005) *An introduction to Black Holes, Information and the String Theory Revolution. The Holographic Universe*. Singapore: World Scientific Publishing.



Timpson, C.G. (2006) 'Philosophical Aspects of Quantum Information Theory'. In: D. Rickles (Ed.) *The Ashgate Companion to Contemporary Philosophy of Physics*. Aldershot, UK; Burlington, USA: Ashgate, 197-261.

Varela, F.J., Thompson, E. & Rosch, E. (1991) *The Embodied Mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

Wheeler, J.A. (1990) 'Information, physics, quantum: The search for links'. In W. Zurek (Ed.) *Complexity, Entropy, and the Physics of Information*. Redwood City, CA.: Addison-Wesley, 309-336.

World Science Festival (2011). *A Thin Sheet of Reality: The Universe as a Hologram* [video file]. Retrieved from: http://worldsciencefestival.com/events/holographic_world [18 june 2013].