phuse·connect
connect·share·advance
10th–13th November
RAI Amsterdam
The Netherlands
EU 2019
The Clinical Data Science Conference

Paper PP20

Holon – What, Why and How

Catharina Dahlbo, Capish, Malmö, Sweden
Eva Kelty, Capish, Malmö, Sweden

ABSTRACT

A new ontology-based information model, structured as a graph, has been developed to make information stored in datasets more accessible. The model is constituted of small messages of information, called Holons, that can be defined as the minimum information a person knowledgeable within a field can understand and use for making a decision. By utilizing pre-defined relations these Holons create a network of connected information. Both data and metadata coexist in a Holon. When data is modelled in a graph with focus on Holons instead of relationships, new and powerful possibilities are created. The simplicity of the Holons enables the development of rapid search tools for handling and reusing the information they contain.

This poster will focus on explaining what a Holon is and how it can be used in order to have an impact of the way data can be evaluated.

BACKGROUND

The world is changing, and it is no longer enough for an organization to have excellent control of data concerning its internal processes and efficiency. In order to remain competitive, it is necessary to competently navigate in the new information age, which requires a much more holistic view built on a mixture of internal and external data. Information is king, and internal data silos must be broken down and integrated with other types of information regardless of where it comes from. The ability to understand the entire picture is what creates competitive advantage today.

The human brain is constantly and effortlessly conceptualizing everything we encounter. This is how we bring meaning to our experiences. Looking around there is not just shapes and colors, but also books, screens, lamps and keyboards. The ambiguous duck-rabbit figure published in Die Fliegende Blätter 1892 is a classic illustration of our brains' conceptualizing abilities: looking at the picture you see either a rabbit or a duck, but never both at the same time or nothing.

The human brain is constantly and effortlessly conceptualizing everything we encounter. This is how we bring meaning to our experiences. Looking around there is not just shapes and colors, but also books, screens, lamps and keyboards. The ambiguous duck-rabbit figure published in Die Fliegende Blätter 1892 is a classic illustration of our brains' conceptualizing abilities: looking at the picture you see either a rabbit or a duck, but never both at the same time or nothing.

Figure 1. Rabbit and Duck *from the Fliegende Blätter 1892. An early example of cognitive illusion. Your brain organizes incoming information into something meaningful and conceptualizes it into something known. This makes the picture switch between a rabbit and a duck even though you may never have seen such a rabbit or duck.*

The ability to classify things into concepts is extremely important for our brains. Concepts can be considered together with other concepts simultaneously which increases understanding and rapid decision making.

According to Wikipedia a holon is something that is simultaneously a whole and a part and the Capish® Holon is an artificial representation of this philosophical concept.

Capish has developed a way to take control of scattered, unorganized and non-integrated data and information by collecting all these pieces of information, taking them apart, standardizing them and re-integrating them into reusable chunks, Holons, that bring new meaning to, and understanding of, the information. Simple things like a date or more complicated entities like blood pressure are conceptualized and handled with the same ease. Concepts are also given "names" which are the fundamental building blocks of our language.

The definition of the Capish Holon (Holon) builds on and is inspired by a number of thinkers:

1. The Greek philosopher Parmenides for proposing an "ontological characterization of the fundamental nature of reality". [1]
2. The philosophers Arthur Koestler and Ken Wilber for defining and suggesting the holon as a stable element within a "school of thought". [2, 3]
3. The Library of Congress for maintaining the "Dewey Decimal Classification" to indicate the knowledge domain of a "school of thoughts", which assist in the definition of the terms used. [4]
4. The "father of modern linguistics" Noam Chomsky for the idea that language has a neurological base and that our brain in its turn correctly interprets the world, and for the semantic "grammar of term specifications". [5]
5. The cognitive psychologist Steven Pinker for clarifying the neurological basis for concepts in our brains which is the foundation of the Holons. [6]

In this paper medical and clinical study data is used to illustrate what Holons are and how they can be used. The medical field was chosen since it represents an extremely complex area full of laws and regulations and external standards as well as a multitude of data of different types, ranging from text messages to precise numerical measurement. And it is an area that is not adequately served by the current technologies. But it must be stated that there is a long list of other possible fields that handle a wide range of different types of complex information like e.g. government, construction, research, finance, development, or any manufacturing organization.

STORAGE OF INFORMATION

All information models must have some defined way of storing data and information. Long lists of structured information are often stored in tables (e.g. an address list in a spreadsheet), while information collected from different sources about something specific appears unstructured and often ends up in text (e.g. a Medical Record for a patient). Here, Holons are the fundamental information carriers used to care for both these situations.

CAPISH HOLON

The definition of a Capish Holon is

- a concept that can be understandable and meaningful by itself
- classified according to content
- an information carrier – including groups, fields, values and attributes
- related to other Holons - directly or indirectly
- a building block in the ontology
- system independent

Generally, concepts rely on atomized definitions of individual entities. This level of detail is difficult for a human mind to perceive; therefore, Capish is dividing the world into larger chunks i.e. Holons and composition of Holons. The Holons are smaller than ordinary database tables, but large enough to include a number of individual data points. This limits the complexity to something that both people and computers can handle. A Holon collects data corresponding to a common concept, which is something that is understandable within a given knowledge domain or discipline so persons trained within the same domain can understand it. A concept shall represent the notion or

thought of a concept by a human and shall aim to follow the most common way of using and representing information, as described in scientific literature and by field experts. One of the key features of a Holon is that it shall cover a single concept. We call the concepts Holons, and they contain both detailed content information in well-defined fields (content) as well as relationships to other Holons (context). Examples of concepts are Medication, Adverse Event and Blood Pressure. Every Holon is of a specific type that is designed to be large enough to convey a meaning that can be used in decision making (self-contained), but small enough to serve as a standardized piece of information.

When applied to the medical field the types of Holons are similar to the "archetypes" suggested by the semantic modeling initiative Open Electronic Health Records (openEHR). This is not surprising since medical science is a well-defined school of concepts being taught at medical schools and included in medical textbooks all over the world. It is these common concepts that make it possible for physicians to advance the medical science by participating in the never-ending discussions taking place between colleagues, within scientific publications and at international conferences. The Holons simply adhere to these universally used concepts.

The Dewey Decimal System (DDC) is a library classification system and was selected for defining a scientific area of a Holon. The reasons for this choice include that it: 1) Adheres to the disciplines as taught at universities; 2) Is globally spread (more than 135 countries) and widely used in schools and libraries (200,000 libraries); and 3) Is regularly updated to accommodate new scientific areas. Taken together the Dewey code and concept will identify a scientific discipline, or knowledge domain, in which all experts have a common understanding of the terms used.

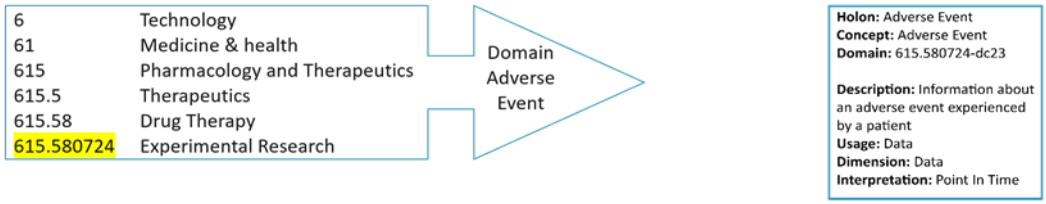

Figure 2. *The Holon, Adverse Event, has been connected to a scientific area by assigning a Dewey code and a concept.*

INFORMATION MODEL

When the Holons are related to each other they will create a mind map of information. In computer terminology such a mind map is called a graph and the Holons are called nodes. Having content rich Holons reduces the number of nodes and relationships in the data graph drastically without losing the information details contained within the Holons.

The information content of a Holon is designed to be big enough for the Holon to be read and understood by itself (self-contained), but small enough to serve as a reusable building block in a larger map of related information. This implies that a Holon may contain various amounts of information in specified fields and attributes depending on the concept at hand.

Minimum information on a Blood Pressure Holon is Systolic Blood Pressure and Diastolic Blood Pressure together with their units, Date and a relationship to the Patient on which it was measured. It is obvious that a Holon for a Medication will include totally other fields like Drug Name, Daily Dose, Route of Administration, Indication and other fields relevant for describing a medication. These fields are, of course, also concepts − but they are too small to be meaningful by themselves. A Blood Pressure Holon without a Date or relationship to a Patient is not meaningful. As is a Blood Pressure Holon containing only a Date. It is the combination of data that gives the Holon a meaning and keeping the fields together within a single Holon makes them easy to understand for a person, and possible to process for a computer. In addition to data fields, a Holon also contains one or more named relationships to other Holons. Hence, both the Blood Pressure and the Medication Holons can be related to a Patient that in turn is related to a Clinical Study that is related to a Sponsor etc. In this way, the model can be expanded into a huge map of information resembling a gigantic mind map.

A Holon is solidly based on current philosophies, sciences and standards. It must be noted that it is a semantic approach, implying that it is best at modeling things that we have concepts and terms for, and therefore can be discussed. It is not suitable for keeping a genome database but works very well for storing information about found biomarkers and studying their relationships to clinical data for the patients.

THE CAPISH HOLON IN CONTEXT

The figure below shows how the Capish Holon contains data corresponding to a common and well-known concept and that both data and metadata are stored in it which makes it understandable by itself. The Capish Ontology describes the structure of the metadata and contains all the rules for the information model, which is a graph of connected Holons. The Ontology defines how the Holons relate to each other, which fields can be included in each Holon, the datatypes of the fields, units to be used, etc.

The actual data is stored in instances, Holons, of the Ontology. The Holon is thereby an information carrier that consists of groups, fields, values, and attributes, i.e. data and metadata. Every Holon should be large enough to convey meaning for use in decision making and small enough to be like "Lego pieces" to make a complete description. Fields are defined in the same way as a Holon by placing it in a scientific area by assigning a Dewey code and a concept.

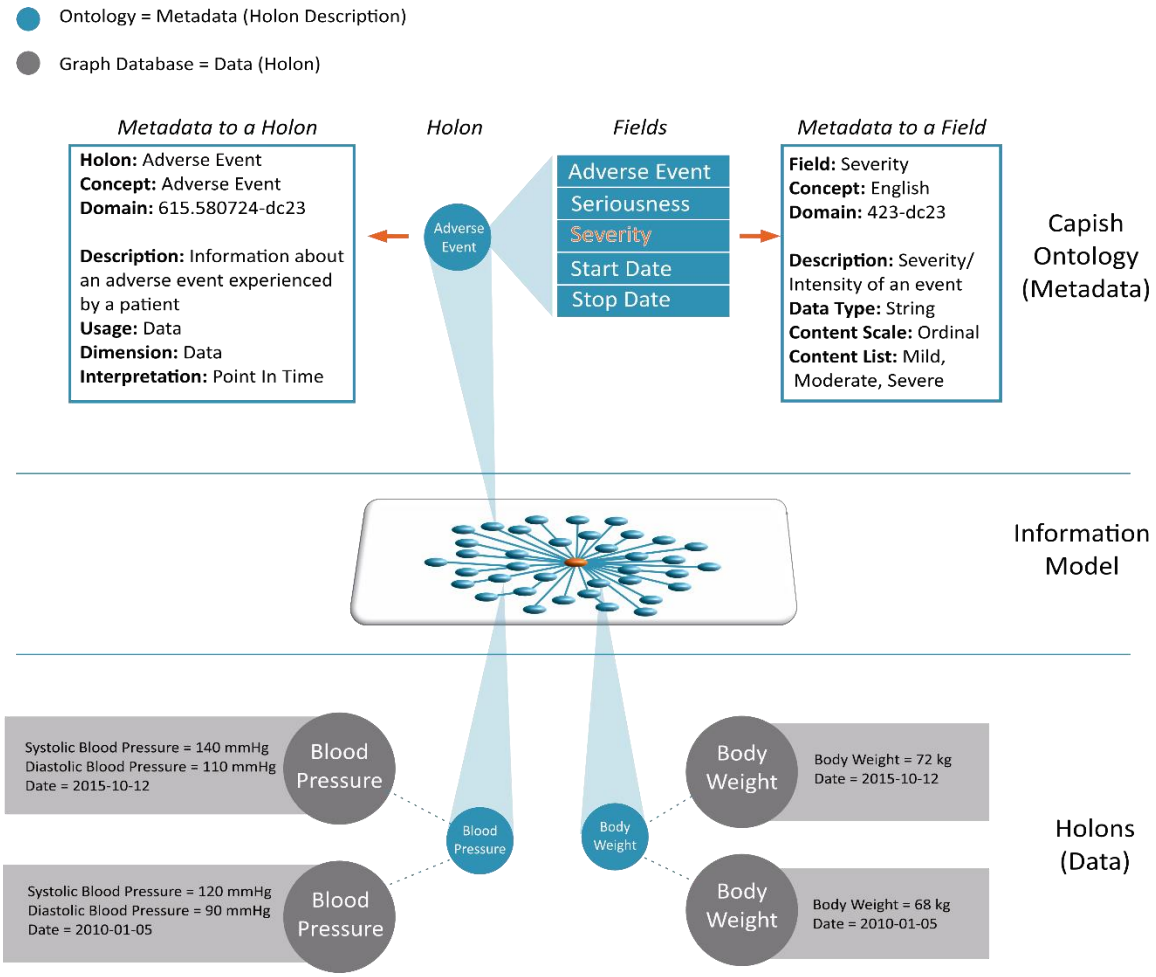

Figure 3. *The Capish Ontology describes the structure of the metadata and contains all the rules for the information model, the map of related Holons. The actual data is stored together with metadata in instances of the Ontology.*

THE PRACTICAL USE OF THE HOLON

The practical realization of the Capish information model has developed over many years and builds on, and is inspired by, standardization work within several organizations:

1. ***NASA's "Open Archival Information System"***, which is keeping context and metadata in a text format that can be read by both computers and humans. This is the basis for the semantic approach and extensive use of the markup language XML from the World Wide Web Consortium (W3C), which also assures that the information carriers (Holons) are future proof and digitally archivable. [7]
2. ***US Department of Defense's "CALS Initiative",*** which uses an unbiased modeling in an XML/SGML based infrastructure that makes it possible for vehicles and ships to keep their individual information and logbooks on board while still being able to pool them for overall analysis and maintenance. Capish provides similar functionality in that patients could keep and control their individual medical records on a card or a chip and still seamlessly include this data (after their permission) in longitudinal or cross-sectional research. [8, 9]
3. ***"ICAM" within the US Air Force***, which introduced the functional modeling method IDEF0, which is used for breaking down complex processes. [10]

STORAGE FORMAT

In practice, a Holon is like a document with content and relations to other documents and together they can create huge graphs of related information. Technically, the Holon is a simple structured text message that can be tagged in the XML syntax. It is generic in nature and has only six tags which are used in combinations to describe the Holon.

Header

```
<o
holonType="Body Weight.Results"
holonConcept="Body Weight"
dimension="Data"
domain="616.0754-dc23"
ontology="(G5.24.1.0)ClinSci.ontology@capish.se"
usage="Data"
>
<k>
<f>Body Weight.Results Key<v>BodyWeightKey…</v></f>
</k>
```

Message

```
<f>Date<v standard="ISO 8610">2010-01-05</v><f>
<f>Body Weight<v unit="kg">68</v></f>
<f>Body Weight.Original<v unit="pounds">150</v></f>
```

Relation

```
<r
reverse="Has" dimension="Data" holonType="Patient">
Of
<f>Patient Key<v> PatientKey…</v></f>
<r/>

</o>
```

Figure 4. *An example of a Holon stored in XML format.*

Having XML messages in flat files is also ideal for digital archiving since the information and its descriptions can be read by both computers and a human's "naked eye". Tagged flat files are also easy to secure both with respect to data integrity through check sums and digital signatures.

EFFECTS OF HOLONS

The foundation of this approach is the way the information is modeled and that different user interfaces can be designed depending on how the information is indexed. The simplicity of the Holons makes it possible to develop rapid search tools for handling and reusing the information they contain. The Holons are stored in a graph database where a patented index structure [11] enables rapid search and retrieval of a broader result set, as the direct and indirect relations between the Holons are utilized. [12, 13]

Having data in Holons makes it possible to include large amounts of data and text while keeping up both speed of processing and end user overview. The reason for this is, of course, that the system extremely fast locates the relevant Holons and then in the next step only needs to care about these Holons. Another nice feature is that the located Holons are easy to collect and sort in different ways. The Holon is an efficient information carrier that makes it possible to study the individual Holon as such, as well as aggregate all data across Holons. The ontology-based information model makes it possible to easily identify and graphically display the data of interest for a user. [14]

So far, the information graph contains a single layer of related Holons. But this is a simplification, an additional effect is that this approach allows different information graphs in parallel layers to each other, several dimensions, such as provenance, traceability and archiving.

CONCLUSION

The Capish Holon was developed to process and conceptualize complex data. The way to process a conceptual information model, derived from semantically stable concepts that constitute the common language within any specific knowledge area, makes it possible to include large amounts of data and still maintain speed of processing. Holons, information carriers linked by relations, create a graph that together with an advanced index structure enables end users to answer complex questions. New Holons and relationships can easily be incorporated without disturbing the existing graph. Once data has been properly curated and put in Holons, data can be exported as XML-files, e.g. one file per patient or concept. Data can then be reused repeatedly for many different purposes over many years as the information remains understandable.

REFERENCES

1. https://en.wikipedia.org/wiki/Parmenides
2. A. Koestler. *"The ghost in the machine",* 1967
3. K. Wilber, "*A Brief History of Everything*", 1st ed. 1996, 2nd ed. 2001: ISBN 1-57062-740-1
4. https://www.oclc.org/en/dewey.html
5. N. Chomsky, *"The Logical Structure of Linguistic Theory",* 1975 (1955)
6. S. Pinker, "*The Stuff of Thought: Language As a window Into Human Nature",* 2007
7. The Consultive Committee for Space Data Systems (CCSDS), *"Reference Model for An Open Archival Information System (Oais)"*, 2012
8. F. Dawson, and F. Nielsen, *"ODA and Document Interchange",* UNIX Review, vol. 8, no. 3, 1990, p. 50.
9. A. Hobgood, *"CALS Implementation--Still a Few Questions",* Advanced Imaging, April 1990, pp. 24-25.
10. Computer Systems Laboratory of the National Institute of Standards and Technology (NIST), *"FIPS Publication 183 released of IDEF0"*, December 1993
11. S. Gestrelius and H. Drews, "*Patent US20130246438 – Reflective logic unlocks knowledge in datasets*", 2013
12. C. Dahlbo, A. Workneh and H. Drews, *"PP04 – Unique Technology for the Future"*, presented at the PhUSE EU Connect, 2018.
13. C. Dahlbo and M.Rubison, *"PP01 - Make the Most of Your Data - Explore Different Perspectives",* presented at PhUSE US Connect 2018.
14. P.Tormay and A. Berg, *"DH07 - Harness the power of ontologies to build better information platforms"*, presented at PhUSE US Connect 2019

ACKNOWLEDGMENTS

We would like to acknowledge our colleagues Anna Berg, Peter Tormay and Henrik Drews for their feedback to the work presented in this paper.

RECOMMENDED READING

1. P.Tormay, "*SD06 – Selection in Action",* presented at PhUSE EU Connect 2019
2. P.Tormay and A. Berg, "*DH08 – Data Transparency – Breaking Down Data Silos for Improved Insight*", presented at PhUSE EU Connect 2018
3. E.Kelty, "*DV07 - Effective Data Modelling for Effective Data Visualization"*, presented at PhUSE EU Connect 2018

4. P. Tormay and H. Drews, *"TT06 – Reflect on your data"*, presented at the annual PhUSE conference, 2016.
5. A. Berg, H. Drews and C.Dahlbo, "*PP05 – New Approach to Graph Databases*", presented at PhUSE EU Connect 2018
6. C. Dahlbo, E.Kelty and M.Rubison, *"PP20 - Understanding the Power of Information Modeling for Efficient and Effective Visualization" ",* presented at PhUSE US Connect 2017.
7. C. Dahlbo, "*PP07 - Experiences of managing Quality Registry data for Effective Exploration*", presented at PhUSE conference 2017
8. C. Dahlbo, "*PP21 - Reflections on the Effects of Data Pooling*", presented at PhUSE conference 2016
9. C. Dahlbo, "*PP31 - Flexible Solutions for Insight and Data Driven Innovation*", presented at PhUSE conference 2015
10. A.Berg and C. Dahlbo, "*PP35 – The Capish Information Model - Simplify Access to Your Data*", presented at PhUSE conference 2014

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Catharina Dahlbo
Capish Nordic AB
Carlsgatan 3
211 20 MALMÖ, Sweden

+46 (0)40 10 88 80
catharina.dahlbo@capish.com

Eva Kelty
Capish Nordic AB
Carlsgatan 3
211 20 MALMÖ, Sweden

+46 (0)40 10 88 80
eva.kelty@capish.com

www.capish.com