Praise for *Introduction to Knowledge Organization*

'The book uses a comprehensive theoretical framework for examining the theories of knowledge organization, the structural principles in knowledge organization, and the impacts of these theories and principles in the construction of knowledge organization systems (KOSs). A very unique component of the book is its coverage of the Semantic Web and linked data in the context of the representation of knowledge organization structures, including convincing cases and examples, as well as the issues and progress of fundamental and advanced applications. The range and scope of the book present a state-of-the-art review of the whole domain of knowledge organization.'
Marcia Lei Zeng, School of Information, Kent State University

'This is a broad and concise introduction to the growing field of knowledge organization. It introduces the most important kinds of knowledge organization systems used, both traditionally in physical libraries and in the digital world. It covers practical as well as theoretical and historical issues. I particularly like the view that the underlying theoretical assumptions of all knowledge organization systems need to be detected.'
Birger Hjørland, Department of Communication, University of Copenhagen

'This book provides a clear but detailed overview of knowledge organization. Its remarkable depth and analytical approach enables it to highlight useful connections between philosophical and theoretical approaches and practical applications. The author's classic interdisciplinary and encyclopedic knowledge makes it accessible both to a specialized and a general public. Searching first in its historical, philosophical and epistemological depths, Gnoli proposes a work that reads like a novel in spite of the details and technical terms. *Introduction to Knowledge Organization* will be suitable for established researchers, progressing and advanced students.'
Widad Mustafa El Hadi, Department of Information Science, University of Lille SHS

'In clear and accessible writing, for a broad group of readers regardless of their scientific field, Claudio Gnoli's new book introduces knowledge organization in a diachronic approach and presents the theoretical foundations and the practice in various contexts.

Introduction to Knowledge Organization constitutes an exceptional contribution to the learning of knowledge organization, targeted at students, both undergraduates and postgraduates, and also at researchers.'
Olívia Pestana, Department of Communication and Information Sciences, University of Porto

'This well-written book is an excellent and timely account of the theories, practical applications and discipline of knowledge organization. This book quickly dispels any possible misconceptions that this area of library and information science is merely about arranging books on shelves. One of the book's many strengths lies in its very clear explanations of theories and topics. It also situates knowledge organization within its broader context, articulating the broader philosophical arguments at work. The inclusion of contemporary developments in knowledge organization research, as well as discussions about digital knowledge organization, make this book especially useful to those currently studying or working with knowledge organization. This book provides a highly informative account of the field of knowledge organization, and offers a powerful argument as to the

importance, breadth and sheer usefulness of this significant area of library and information practice and research.

This book would be invaluable as a textbook for those studying library and information science (LIS), and would also be of interest to those working in knowledge organization and who want to enhance and deepen their understanding of knowledge organization.'
Deborah Lee, Department of Library and Information Science, City, University of London

Introduction to Knowledge Organization

Every purchase of a Facet book helps to fund CILIP's advocacy, awareness and accreditation programs for information professionals.

Introduction to Knowledge Organization

Claudio Gnoli

facet publishing

Published by Facet Publishing
7 Ridgmount Street, London WC1E 7AE
www.facetpublishing.co.uk

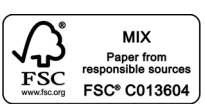
Facet Publishing is wholly owned by CILIP: the Library and Information Association.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

ISBN 978-1-78330-465-3 (paperback)
ISBN 978-1-78330-466-0 (hardback)
ISBN 978-1-78330-467-7 (e-book)

First published 2020

Text printed on FSC accredited material.



Typeset from author's files in 10/13pt Revival 565 and Frutiger by Flagholme Publishing Services.
Printed and made in Great Britain by CPI Group (UK) Ltd, Croydon, CR0 4YY.

For Sara Iommi (1983–2016), explorer of audiovisual archives, folk culture and knowledge at large, anthropologist, intellectual, militant, dearest friend.

Acknowledgments

I am grateful to David Bawden for his trust when proposing to produce this work. In revising its first version I have benefitted from knowledgeable comments and suggestions by Carlo Bianchini, Riccardo Ridi (all chapters), Ceri Binding, Amélie Daloz, Aida Slavic and Douglas Tudhope (individual chapters). My thanks go to my copy-editor Judith Oppenheimer for providing this non-native author with much help on English stylistic issues.

Contents

List of abbreviations

ASIS&T	Association for Information Science and Technology
BC2	Bliss Bibliographic Classification, 2nd edition
BT	broader term
CRG	Classification Research Group
DDC	Dewey Decimal Classification
ILC	Integrative Levels Classification
ISBD	International Standard Book Description
ISKO	International Society for Knowledge Organization
KO	knowledge organization
KOS	knowledge organization system
LCC	Library of Congress Classification
LCSH	Library of Congress Subject Headings
LIS	library and information science
MARC	Machine-Readable Cataloging
NKOS	networked knowledge organization systems
NLP	natural language processing
NT	narrower term
OPAC	online public access catalog
OWL	Web Ontology Language
RDA	Resource Description and Access
RDF	Resource Description Framework
RT	related term
SKOS	Simple KOS
UDC	Universal Decimal Classification
UF	used for
URI	uniform resource identifier
XML	Extended Markup Language
W3C	World Wide Web Consortium

CHAPTER 1

Focusing the field

1.1 What is knowledge?

We all are aware that we are living in the Information Age, just as there have been ages of hunting and gathering, of agriculture and of industry in the preceding centuries and millennia. Information and knowledge shape our lives and our society in many ways.

This is now increasingly reflected in scholarship: terms such as *information architecture*, *knowledge management* and *knowledge representation* are increasingly common in the titles of conferences, books and educational programs. This book contributes yet another tile to this mosaic, by dealing with the conceptual arrangement of knowledge content.

However, before we get down to details we need to get a clear idea of what we are talking about. Obviously, this is a requirement in any domain, but it is even more necessary in the present case, as terms such as those used above, although suggestive, cover a particularly wide and fuzzy semantic space. When people from outside our field ask me about my research domain and I answer 'knowledge organization', their faces often express a mix of polite deference and perplexity about what exactly am I referring to.

Let us start with the first half of the term. What do we mean by *knowledge*?

1.1.1 Knowledge as representation of networks

Intuitively, we are aware that knowledge often refers to someone knowing something – that is, having developed some internal representation of certain external objects. That Maria knows the story of Christopher Columbus means that her mind has some representation of a sequence of events involving the navigator and his actions which led to the rediscovery of the Americas by Europeans.

The *representation* consists in a model of some set of entities and of links among them. While Columbus, his sailors, his maps and his ships were made of material substances, the notions in Maria's mind are immaterial ideas, recorded in a complex system of neural connections. Still, the material entities and the ideas modeling them are similar in terms of the network of relationships between

their parts – they are isomorphic. Knowledge is thus a (usually simplified) reproduction of relationship networks in a different substrate. (This description differs from traditional philosophical definitions of knowledge as 'true and justified belief', although 'truth' corresponds in our description to isomorphism; a model that does not correspond to the represented object is a 'false' one.)

The exact relation between objects and their representation in knowledge has been the matter of philosophical discussion for millennia. Six centuries ago Nicholas of Cusa compared reality to the curvature of a circle, which human knowledge can try to approximate by designing polygons with an increasing number of sides, although without ever being able to achieve perfect circularity. Much more recently, philosopher of information Luciano Floridi (2019) has claimed that knowledge should be conceived of as being like cooking its objects – rather than just like depicting or photographing them. The resulting dish is very different from the original ingredients, although it does depend on them for its existence. The efforts of science are continually trying to get closer to a faithful representation.

In the above example the contents of Maria's mind exist in the brain of a living person. However, if we are talking about the knowledge contained in a locked library on a Sunday night, even though there is no living person there we can probably agree that there still is some knowledge. This seems to suggest that knowledge can exist even externally to people. So, can any entity be a carrier of knowledge?

Biologists have found that the DNA contained in cell nuclei transmits to the next generation some 'knowledge' about how an organism can develop, grow and work. Konrad Lorenz, a founder of the science of animal and human behavior, entitled his epistemological masterpiece *Behind the Mirror: a search for a natural history of human knowledge*. In chapter 1, on 'Life as a knowing process', he illustrates how the adaptation of an organism to its environment can be seen as a form of knowledge of the environment that has indeed been transmitted through its genes. The genes of a horse (through various biochemical stages) triggered the development of flat hooves, and that flat shape is adapted to the flat environment of the steppe where a wild horse will have to live, just as a sort of mold. So those genes are a form of knowledge of the flatness of the steppe (Lorenz, 1973).

1.1.2 Data, information, knowledge, understanding, wisdom

What Lorenz in this effective description calls 'knowledge' is nowadays more often described as *information*. These two terms are often used with overlapping meanings, and it has been found that the term 'information organization' is, to all intents and purposes, equivalent to 'knowledge organization' (Hjørland, 2012).

Their connection is better modeled in the metaphor of the *DIKW (data-information-knowledge-wisdom) pyramid* (Rowley, 2007; Frické, 2018): at the

base of the pyramid, sets of integrated *data* form information; further integration of information gives rise to the upper layer of knowledge; and integration of knowledge can give rise to wisdom. Bawden and Robinson (2015) also proposed an additional level of *understanding* between knowledge and wisdom. In another metaphorical model (Marshall, 2013) information consists in isolated dots, while knowledge is a network of lines connecting certain dots and wisdom is the deletion of most dots and lines, keeping only those few elements that have proved to be most important.

As to the particular transition that interests us, that between information and knowledge, we can consider the representation of specific facts, such as 'Columbus sailed from Palos', or 'the steppe is flat', as information. The integration of many information elements into a connected, wider system gives rise to knowledge, such as 'the history of North America', or 'life in a steppe environment'.

If we hear about a dinner happening in Europe in 1453, this is one piece of information; but we can then compare this information with the broader network of our knowledge, which includes the notions that 1453 is before 1492, when Europeans first came into contact with native Americans, and that foods such as potatoes or tomatoes were imported into Europe only after the rediscovery of the Americas: from all this, we will *know* that that European dinner in 1453 could not have consisted of potatoes or tomatoes. Knowledge is a network of integrated information that allows us to process further information and set it in a broader picture.

The further integration of knowledge into understanding or wisdom is also very relevant, but concerns personal and spiritual development rather than techniques for arranging representations of reality, hence it is beyond the scope of this book.

1.1.3 Informational systems at various levels

As we have seen, information and its arrangement into networks of 'knowledge' (in a broad sense) can occur at several levels: not just cultural and cognitive, but also organic or even material.

Indeed, it is increasingly acknowledged that information in the most basic sense is a physical property that can be observed in any system, as it just accounts for how the system is and what its state is. The combination of the quantum numbers of an electron determines the range of its possible positions and movements around an atom's nucleus. Some are coming to think that information, in this sense, is even more fundamental than matter: before particles or waves, the world is comprised primarily of information and its structures (French, 2014).

These ideas are stimulating the development of a general 'philosophy of information' that can also be relevant (although not exclusively) to the field of

library and information science (Floridi, 2002; Bawden and Robinson, 2017). A somewhat related conception is Charles Sanders Peirce's idea of 'thirdness' as a general ubiquitous relationship between any three entities. A typical example of thirdness is the relationship of 'semiosis', occurring between a sign, its object and its 'interpretant' (Peirce, 1934). According to Peirce, semiosis is not exclusive to human consciousness, but can occur among any three entities in the universe: in his conception, 'minds' exist at all levels of reality and the universe itself 'thinks'.

Such occurrence of signs at various levels reminds us of what we have said about representations occurring at several levels of organization. At the level of life, atoms and molecules are organized in ways that are specified by information contained in DNA and RNA to form organelles, cells, tissues, organs, organisms, colonies and populations. As we have seen, information plays a key role in genes, on which more complex biological systems depend.

At the next level, that of the mind (now meant in the typically human sense), information consists in notions that are acquired through the sense organs and learning, compared to innate logical faculties and processed in thought. Neural networks of information allow for the development of personal knowledge in individual organisms.

A further level is that of cultures developed by social beings, especially humans: in cultures, information is processed by gestures, languages and other symbolic systems that are transmitted through imitation and social learning. These can crystallize into artefacts, such as tools or buildings, and into 'mentefacts' (Gnoli, 2018b) – that is, intellectual products of human creativity, such as symphonies or theories.

As can be seen, each of these information levels can be based on the representation of the previous levels in some new form: organisms represent their material environment, such as the flatness of the steppe, by their adaptations as recorded in genetic information; minds represent an organism's internal and external situation to themselves; cultures represent mental, organic and material facts in their languages and symbols (Table 1).

Table 1 *Informational systems at different levels*

Informational systems		**Varieties of information**
cultures	mentefacts artefacts communities	public knowledge design gestures and words
minds		personal notions
life		adapted characters
matter		stable configurations
forms		logico-mathematical structures