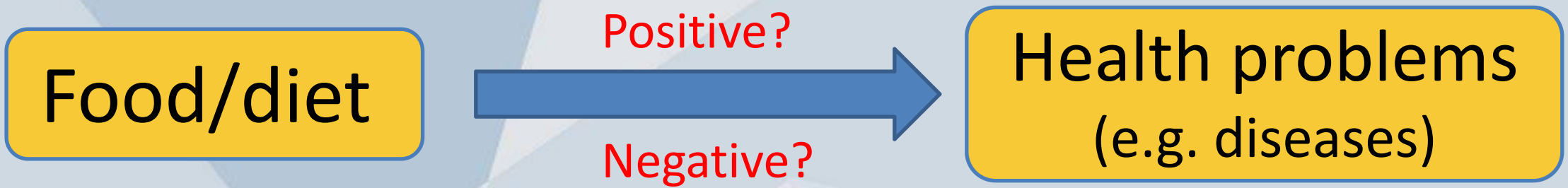
The Ontology for Nutritional Epidemiology (ONE)

" Current achievements and future perspectives

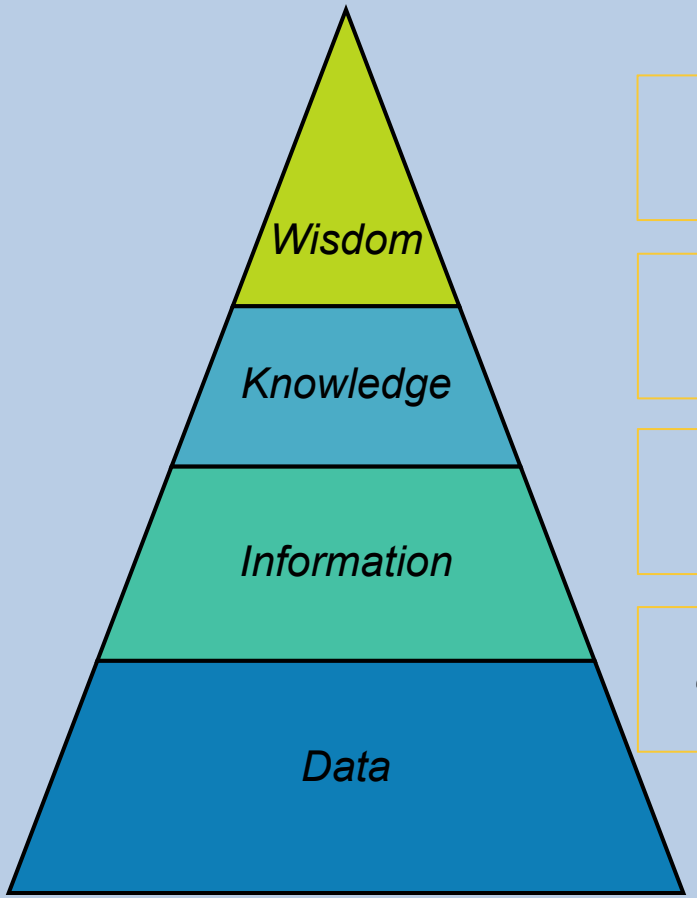
Chen Yang

23 September, 2020

/ What is Nutritional Epidemiology?



/ Information classification (IEEE bigdata 2019, LA, USA)



DIKW pyramid

1	Dietary guidelines and nutrition policies
2	“Diet – health status” relationship
3	Observational/intervention study designs, results…
4	Dietary intake data, anthropometry data, Socio-demographic data, physical activity data, Biomarkers, etc.

/ Ontology for Nutritional Epidemiology (ONE)

nutrients

MDPI

Article

An Ontology to Standardize Research Output of Nutritional Epidemiology: From Paper-Based Standards to Linked Content

Chen Yang [1], Henry Ambayo [1], Bernard De Baets [2], Patrick Kolsteren [1], Nattapon Thanintorn [3], Dana Hawwash [1], Jildau Bouwman [4], Antoon Bronselaer [5], Filip Pattyn [6,†] and Carl Lachat [1,*,†]

[1] Department of Food Technology, Safety and Health, Ghent University, 9000 Ghent, Belgium; Chen.Yang@UGent.be (C.Y.); henryambayo@gmail.com (H.A.); Patrick.Kolsteren@UGent.be (P.K.); Dana.Hawwash@UGent.be (D.H.)

[2] KERMIT, Department of Data Analysis and Mathematical Modelling, Ghent University, 9000 Ghent, Belgium; Bernard.DeBaets@UGent.be

[3] Department of Pathology and Anatomical Sciences, University of Missouri, Columbia, MO 65203, USA; nthanintorn@hotmail.com

[4] Netherlands Organization for Applied Scientific Research, NL-2509 Zeist, The Netherlands; jildau.bouwman@tno.nl

[5] Department of Telecommunications and information processing, Ghent University, 9000 Ghent, Belgium; Antoon.Bronselaer@UGent.be

[6] ONTOFORCE, 9052 Ghent, Belgium; filip.pattyn@ontoforce.com

[†] These authors contributed equally to this work.

[*] Correspondence: Carl.Lachat@UGent.be; Tel.: +32-(0)9-264-93-77

Received: 13 May 2019; Accepted: 6 June 2019; Published: 8 June 2019

”

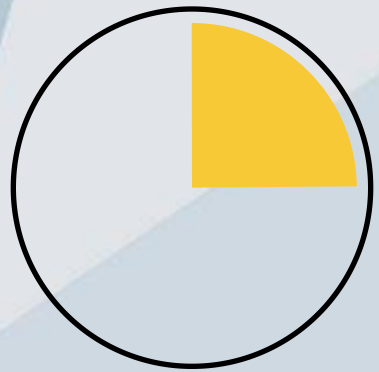
79 new classes to annotate datasets

24 new classes to annotate manuscripts.

/ ONE includes terms about…

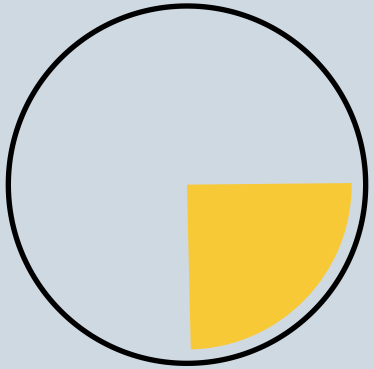
50%

Study designs & reporting



25%

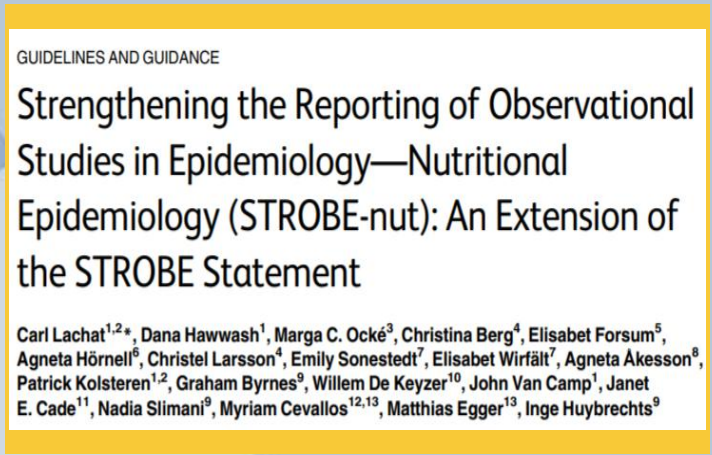
Dietary assessment methods (based on surveys)



25%

Anthropometry measurement methods (based on surveys)

/ Standards used in ONE (2016-2018)



GUIDELINES AND GUIDANCE

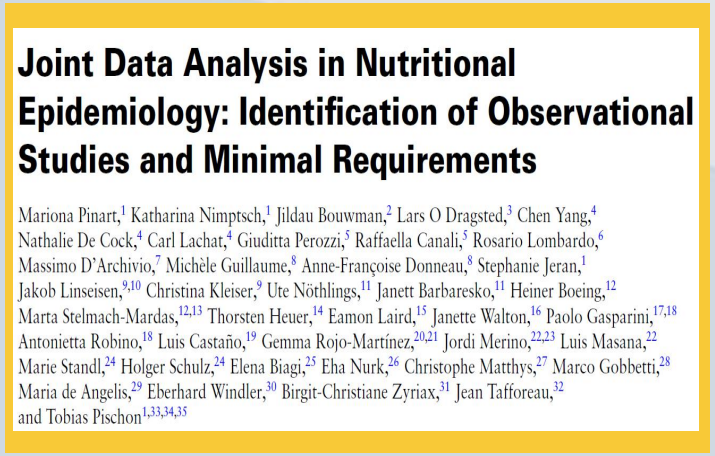
Strengthening the Reporting of Observational Studies in Epidemiology—Nutritional Epidemiology (STROBE-nut): An Extension of the STROBE Statement

Carl Lachat[1,2]*, Dana Hawwash[1], Marga C. Ocké[3], Christina Berg[4], Elisabet Forsum[5], Agneta Hörnell[6], Christel Larsson[4], Emily Sonestedt[7], Elisabet Wirfält[7], Agneta Åkesson[8], Patrick Kolsteren[1,2], Graham Byrnes[9], Willem De Keyzer[10], John Van Camp[1], Janet E. Cade[11], Nadia Slimani[9], Myriam Cevallos[12,13], Matthias Egger[13], Inge Huybrechts[9]

Perspective: Essential Study Quality Descriptors for Data from Nutritional Epidemiologic Research

Chen Yang,[1] Mariona Pinart,[5] Patrick Kolsteren,[1] John Van Camp,[1] Nathalie De Cock,[1] Katharina Nimptsch,[5] Tobias Pischon,[5–8] Eamon Laird,[9] Giuditta Perozzi,[10] Raffaella Canali,[10] Axelle Hoge,[11] Marta Stelmach-Mardas,[12,13] Lars Ove Dragsted,[14] Stéphanie Maria Palombi,[14] Irina Dobre,[14] Jildau Bouwman,[15] Peter Clarys,[16] Fabio Minervini,[17] Maria De Angelis,[17] Marco Gobbetti,[18] Jean Tafforeau,[19] Oscar Coltell,[20,21] Dolores Corella,[21,22] Hendrik De Ruyck,[23] Janette Walton,[24] Laura Kehoe,[24] Christophe Matthys,[25] Bernard De Baets,[2] Guy De Tré,[3] Antoon Bronselaer,[3] Angela Rivellese,[26] Rosalba Giacco,[27] Rosario Lombardo,[28] Sofian De Clercq,[29] Niels Hulstaert,[4,30] and Carl Lachat[1]

Departments of [1]Food Safety and Food Quality, [2]Mathematical Modelling, Statistics and Bioinformatic, [3]Telecommunications and Information Processing, and [4]Biochemistry, Ghent University, Ghent, Belgium; [5]Molecular Epidemiology Research Group, Max Delbrück Centre for Molecula



Joint Data Analysis in Nutritional Epidemiology: Identification of Observational Studies and Minimal Requirements

Mariona Pinart,[1] Katharina Nimptsch,[1] Jildau Bouwman,[2] Lars O Dragsted,[3] Chen Yang,[4] Nathalie De Cock,[4] Carl Lachat,[4] Giuditta Perozzi,[5] Raffaella Canali,[5] Rosario Lombardo,[6] Massimo D'Archivio,[7] Michèle Guillaume,[8] Anne-Françoise Donneau,[8] Stephanie Jeran,[1] Jakob Linseisen,[9,10] Christina Kleiser,[9] Ute Nöthlings,[11] Janett Barbaresko,[11] Heiner Boeing,[12] Marta Stelmach-Mardas,[12,13] Thorsten Heuer,[14] Eamon Laird,[15] Janette Walton,[16] Paolo Gasparini,[17,18] Antonietta Robino,[18] Luis Castaño,[19] Gemma Rojo-Martínez,[20,21] Jordi Merino,[22,23] Luis Masana,[22] Marie Standl,[24] Holger Schulz,[24] Elena Biagi,[25] Eha Nurk,[26] Christophe Matthys,[27] Marco Gobbetti,[28] Maria de Angelis,[29] Eberhard Windler,[30] Birgit-Christiane Zyriax,[31] Jean Tafforeau,[32] and Tobias Pischon[1,33,34,35]

STROBE-nut

Improve reporting of observational studies with a focus on diet and health

(PLOS Med, 2016, IF=11.862)

Data quality items

facilitate interoperability of data repositories

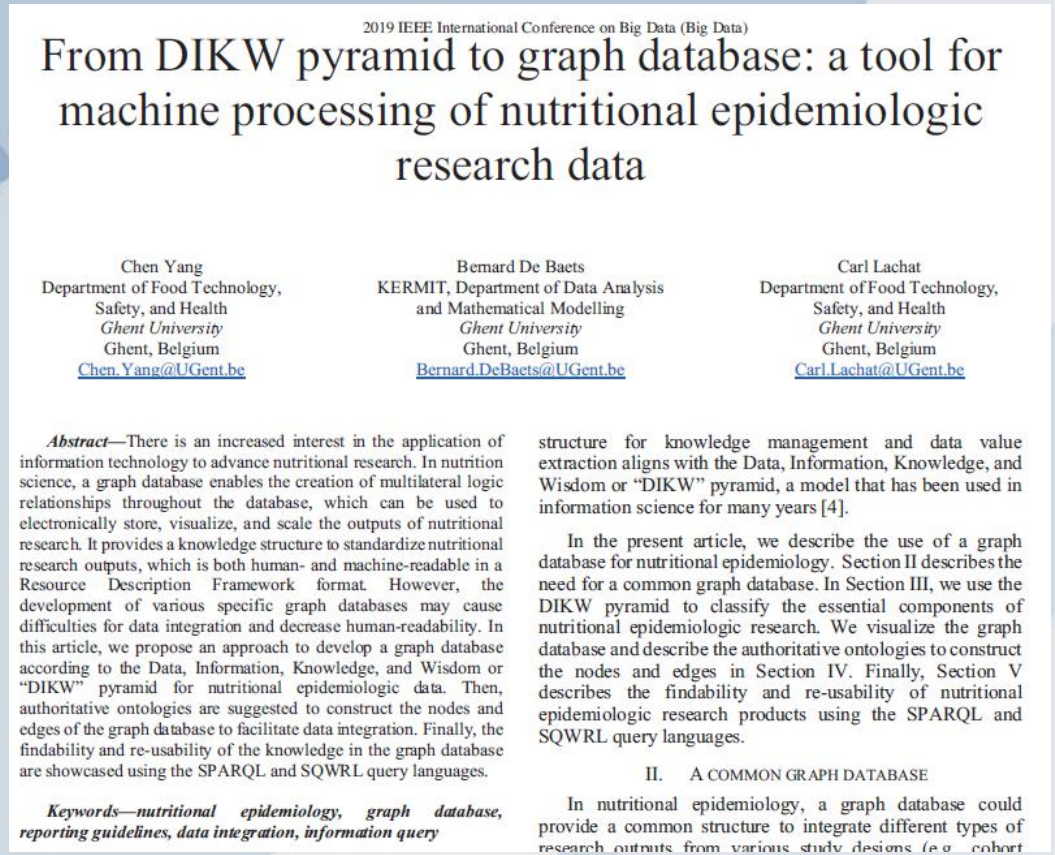
(Advances in Nutrition, 2017, IF=6.853)

Minimal data information

Provide sufficient information for researchers to draft future multicenter research proposals in nutrition

(Journal of Nutrition, 2018, IF=4.416)

/ Applications of ONE (2019-2020)



2019 IEEE International Conference on Big Data (Big Data)

From DIKW pyramid to graph database: a tool for machine processing of nutritional epidemiologic research data

Chen Yang
Department of Food Technology, Safety, and Health
Ghent University
Ghent, Belgium
Chen.Yang@UGent.be

Bernard De Baets
KERMIT, Department of Data Analysis and Mathematical Modelling
Ghent University
Ghent, Belgium
Bernard.DeBaets@UGent.be

Carl Lachat
Department of Food Technology, Safety, and Health
Ghent University
Ghent, Belgium
Carl.Lachat@UGent.be

***Abstract*—There is an increased interest in the application of information technology to advance nutritional research. In nutrition science, a graph database enables the creation of multilateral logic relationships throughout the database, which can be used to electronically store, visualize, and scale the outputs of nutritional research. It provides a knowledge structure to standardize nutritional research outputs, which is both human- and machine-readable in a Resource Description Framework format. However, the development of various specific graph databases may cause difficulties for data integration and decrease human-readability. In this article, we propose an approach to develop a graph database according to the Data, Information, Knowledge, and Wisdom or "DIKW" pyramid for nutritional epidemiologic data. Then, authoritative ontologies are suggested to construct the nodes and edges of the graph database to facilitate data integration. Finally, the findability and re-usability of the knowledge in the graph database are showcased using the SPARQL and SQWRL query languages.**

Keywords—nutritional epidemiology, graph database, reporting guidelines, data integration, information query

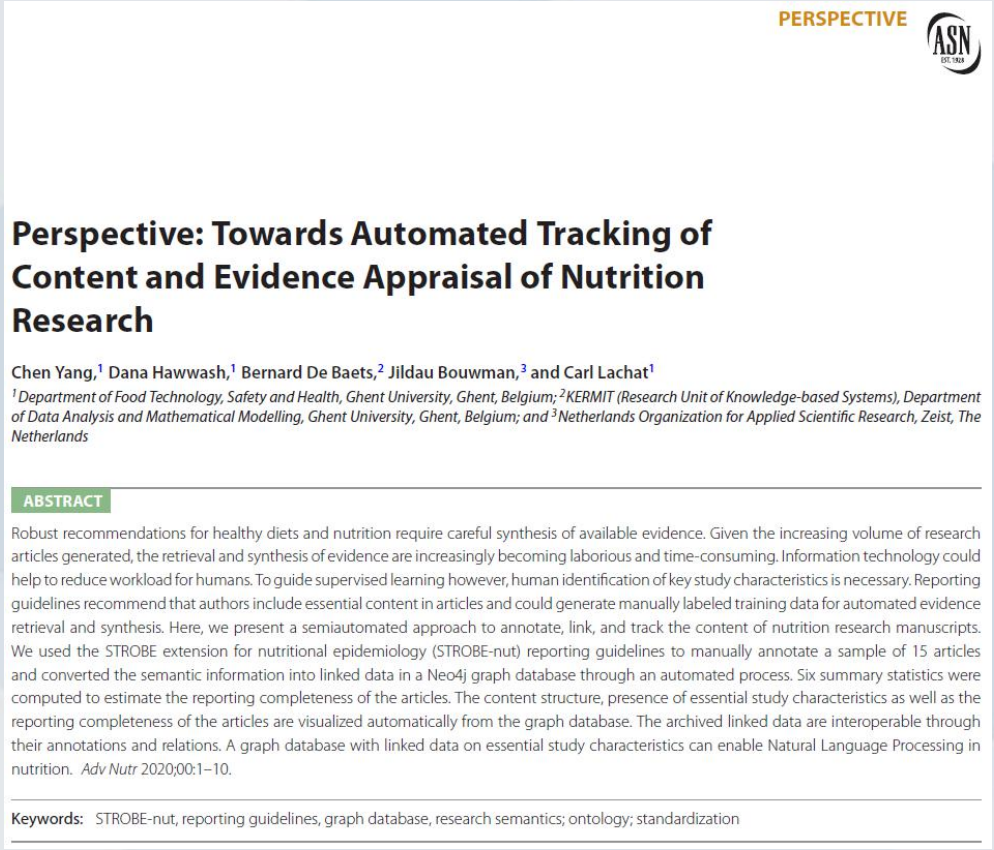
structure for knowledge management and data value extraction aligns with the Data, Information, Knowledge, and Wisdom or "DIKW" pyramid, a model that has been used in information science for many years [4].

In the present article, we describe the use of a graph database for nutritional epidemiology. Section II describes the need for a common graph database. In Section III, we use the DIKW pyramid to classify the essential components of nutritional epidemiologic research. We visualize the graph database and describe the authoritative ontologies to construct the nodes and edges in Section IV. Finally, Section V describes the findability and re-usability of nutritional epidemiologic research products using the SPARQL and SQWRL query languages.

II. A COMMON GRAPH DATABASE

In nutritional epidemiology, a graph database could provide a common structure to integrate different types of research outputs from various study designs (e.g., cohort

(IEEE bigdata 2019, Los Angeles)



PERSPECTIVE

Perspective: Towards Automated Tracking of Content and Evidence Appraisal of Nutrition Research

Chen Yang,[1] Dana Hawwash,[1] Bernard De Baets,[2] Jildau Bouwman,[3] and Carl Lachat[1]

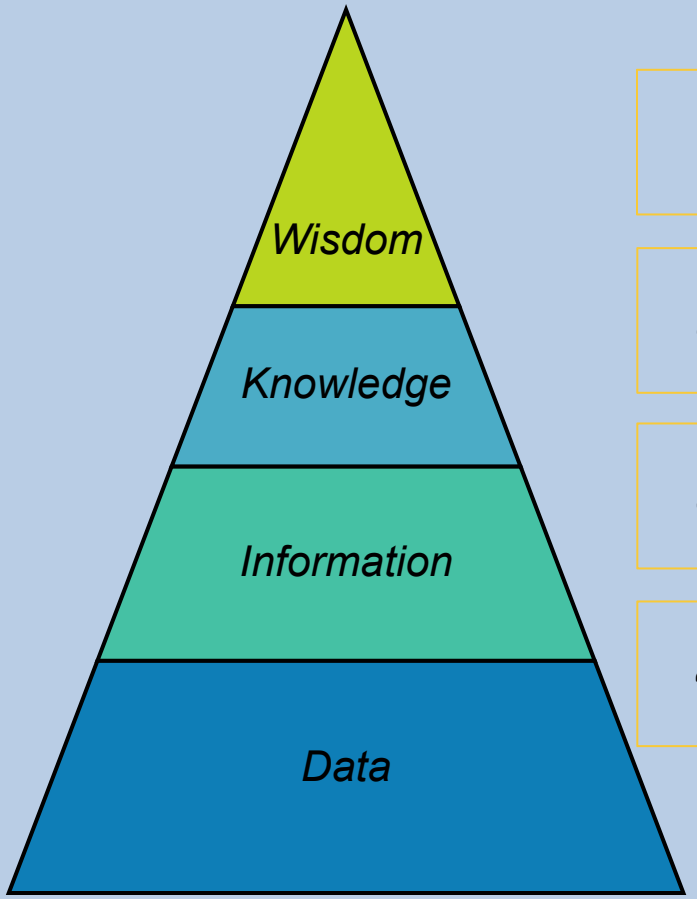
[1] Department of Food Technology, Safety and Health, Ghent University, Ghent, Belgium; [2] KERMIT (Research Unit of Knowledge-based Systems), Department of Data Analysis and Mathematical Modelling, Ghent University, Ghent, Belgium; and [3] Netherlands Organization for Applied Scientific Research, Zeist, The Netherlands

ABSTRACT

Robust recommendations for healthy diets and nutrition require careful synthesis of available evidence. Given the increasing volume of research articles generated, the retrieval and synthesis of evidence are increasingly becoming laborious and time-consuming. Information technology could help to reduce workload for humans. To guide supervised learning however, human identification of key study characteristics is necessary. Reporting guidelines recommend that authors include essential content in articles and could generate manually labeled training data for automated evidence retrieval and synthesis. Here, we present a semiautomated approach to annotate, link, and track the content of nutrition research manuscripts. We used the STROBE extension for nutritional epidemiology (STROBE-nut) reporting guidelines to manually annotate a sample of 15 articles and converted the semantic information into linked data in a Neo4j graph database through an automated process. Six summary statistics were computed to estimate the reporting completeness of the articles. The content structure, presence of essential study characteristics as well as the reporting completeness of the articles are visualized automatically from the graph database. The archived linked data are interoperable through their annotations and relations. A graph database with linked data on essential study characteristics can enable Natural Language Processing in nutrition. *Adv Nutr* 2020;00:1–10.

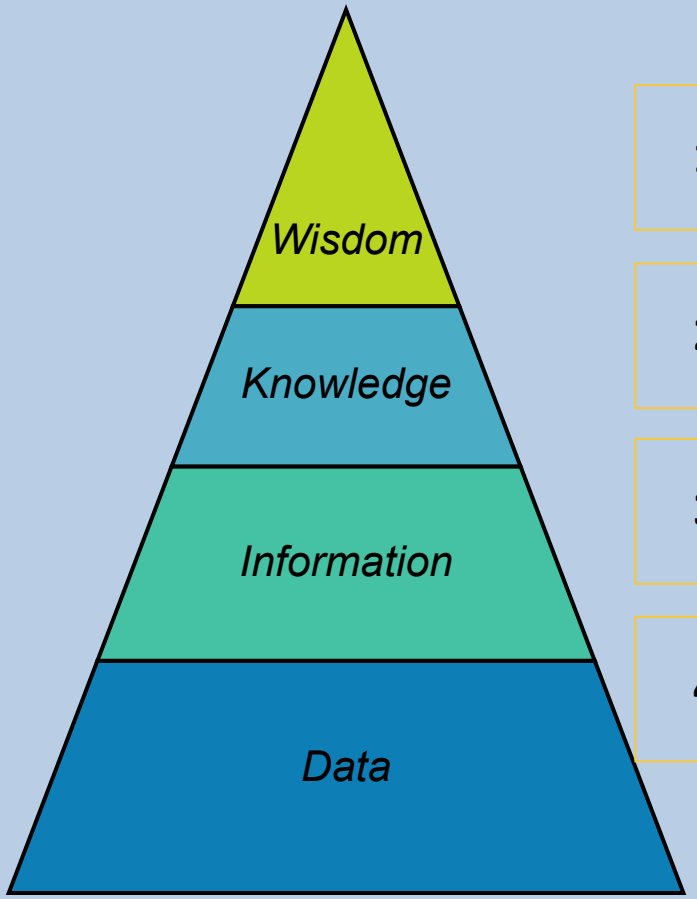
Keywords: STROBE-nut, reporting guidelines, graph database, research semantics; ontology; standardization

(Advances in Nutrition, IF=7.265)



DIKW pyramid

1	Dietary guidelines and nutrition policies
2	“Diet – health status” relationship
3	Observational/intervention study designs, results…
4	Dietary intake data, anthropometry data, Socio-demographic data, physical activity data, Biomarkers, etc.



Use existing ontologies

1	**ONE** (Ontology for Nutritional Epidemiology)
2	**FOODON** (Food Ontology); **DOID** (Human Disease Ontology);
3	**ONE** (Ontology for Nutritional Epidemiology)
4	**FOODON** (Food Ontology); **FOBI** (Food Biomarker Ontology); **OR** (OBO Relations Ontology); **DOID** (Human Disease Ontology); **ONS** (Ontology for Nutritional Studies)

/ A web crawler (Advances in Nutrition, IF=7.265)

BMC Part of Springer Nature

International Journal of Behavioral Nutrition and Physical Activity

Home About Articles Submission Guidelines

Research | Open Access | Published: 17 August 2017

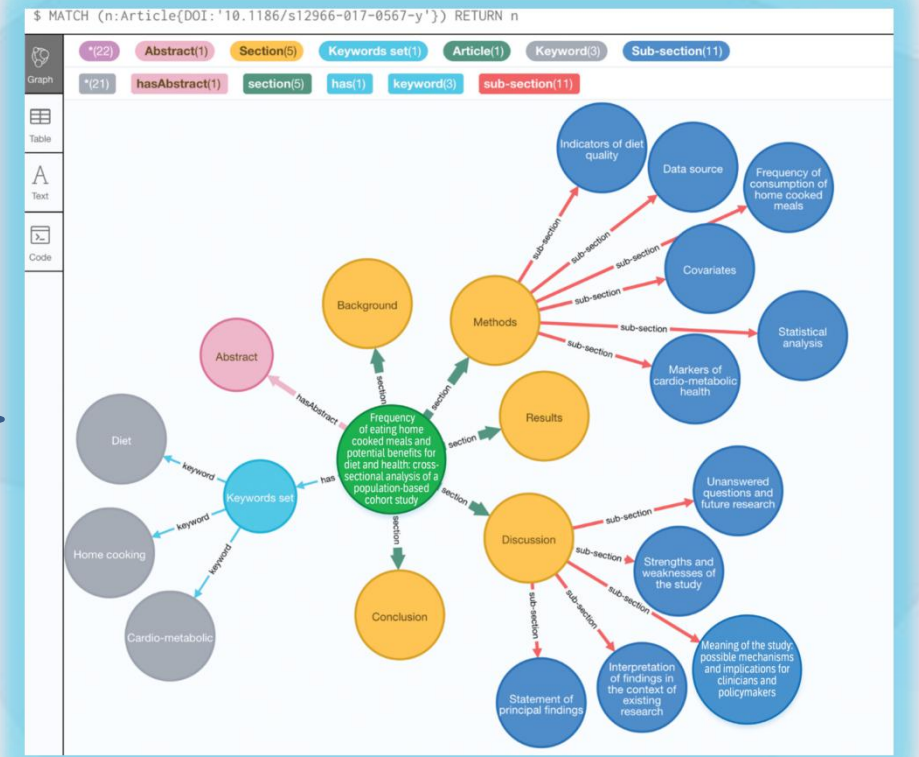
Frequency of eating home cooked meals and potential benefits for diet and health: cross-sectional analysis of a population-based cohort study

Susanna Mills, Heather Brown, Wendy Wrieden, Martin White & Jean Adams

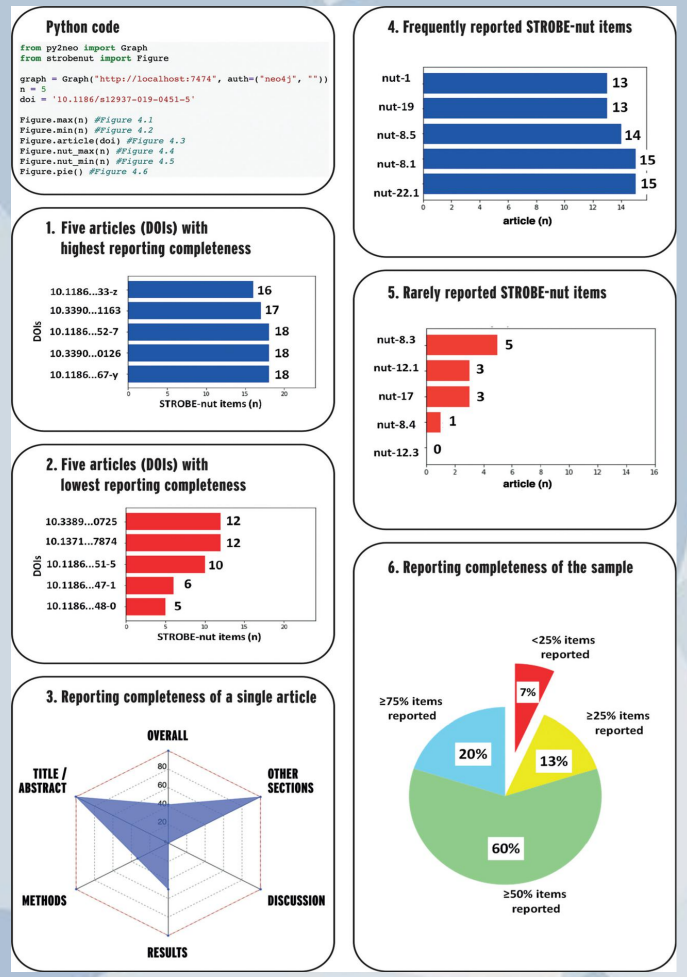
International Journal of Behavioral Nutrition and Physical Activity **14**, Article number: 109 (2017) | Cite this article

16k Accesses | **307** Altmetric | Metrics

Abstract



/ information visualization (Advances in Nutrition, IF=7.265)



A Python module for ONE

A Python module was developed to process content and STROBE-nut annotations of nutritional epidemiologic papers in XML format:

1.Annotate.py: annotate the reporting completeness of papers according to the STROBE-nut reporting guidelines;

2.Figure.py: visualize the statistics of reporting completeness of papers

/ Future perspectives

Release a new version of ONE

ONE will be updated by December 2020.

2020

Ontology harmonization

Work together with FOODON, ONS as well as other food & nutrition ontologies

2020

Be a OBO foundry member

ONE will be evolved according to the OBO foundry rules.

2020

Regular update of ONE

Regular update of ONE is necessary.

2021

Annotating dietary guidelines

New application of ONE will be explored.

2021

/ contact us! Chen.Yang@UGent.be or Carl.Lachat@UGent.be

Thank you!

> a culture of linked data in nutrition research needs to be fostered.

/ Yang et al. 2020

> Basic knowledge regarding the use of ontologies, open science, and FAIR data needs to be integrated in the curriculum of students and researchers.

/ Yang et al. 2020