

عنوان: شناسایی شاخص‌های کیفیت داده برای حکمرانی داده با رویکرد فراترکیب و دلفی فازی
نام نویسندگان: سحر بنیادی: دانشجوی دکتری علم اطلاعات و دانش‌شناسی؛
دانشگاه آزاد اسلامی؛ واحد علوم و تحقیقات؛
تهران، ایران sahar_bonyadi2005@yahoo.com

نجلا حریری*: دکتری علم اطلاعات و دانش‌شناسی؛ استاد؛
دانشگاه آزاد اسلامی؛ واحد علوم و تحقیقات؛
تهران، ایران nadjlahariri@gmail.com

سید مهدی طاهری: دکتری علم اطلاعات و دانش‌شناسی؛ دانشیار؛
دانشگاه علامه طباطبائی؛
تهران، ایران Taherismster@gmail.com

رویا پورنقی: دکتری علم اطلاعات و دانش‌شناسی؛ استادیار
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)؛
تهران، ایران pournaghi@irandoc.ac.ir

مقدمه: یکی از موضوعات مهم در مدیریت داده‌ها، بحث کیفیت داده است. داده خوب و باکیفیت مبنای رسیدن به نتایج مطلوب از سیستم‌ها است. پژوهش‌های مربوط به کیفیت داده، طبقه‌بندی کاملی از ابعاد کیفیت داده‌ها را فراهم می‌کنند؛ با این حال برخی اختلافات در تعریف مهم‌ترین ابعاد با توجه به ماهیت تعریف شده از کیفیت وجود دارد. شش طبقه مهم از ابعاد کیفیت توسط واند و وانگ^۱ (۱۹۹۶)؛ وانگ و استرانگ^۲ (۱۹۹۶)؛ ردمن (۱۹۹۶)؛ جارک^۳ و دیگران (۱۹۹۵)؛ بووی^۴ و دیگران (۲۰۰۳)؛ نیومن^۵ (۲۰۰۲) ارائه شده است. با تجزیه و تحلیل این طبقه‌بندی، تعریف مجموعه‌ی اولیه‌ای از ابعاد کیفیت داده‌ها امکان‌پذیر می‌شود، از جمله آن‌ها دقت، جامعیت، ثبات و به‌موقع بودن است که مورد تمرکز اکثر نویسندگان می‌باشد.

1. Wand & Wang
2. Wang & Strong
3. Jarke
4. Bovee
5. Naumann

با این حال، هیچ توافق کلی بر روی تعریف مجموعه‌ای از ابعاد کیفیت داده‌ها و یا معنای دقیق هر یک از ابعاد وجود ندارد (ارشادی و دیگران، ۱۳۹۶). به دلیل کثرت ابعاد کیفیت داده و تنوع آن‌ها در استفاده، این پژوهش به بررسی شاخص‌های کیفیت داده از دیدگاه متخصصان ایرانی می‌پردازد

روش‌شناسی: در مرحله اول روش فراترکیب با جست و جوی کلید واژه‌های کیفیت داده، مدیریت داده، حکمرانی داده و مدیریت کیفیت داده در پایگاه‌های ایرانداک، Google Scholar، Science Direct، ACM و IEEE، Springer در بین سال‌های (۱۹۹۵-۲۰۲۲) انجام شد و تعداد ۲۶۸ مقاله مرتبط تشخیص داده شد و در مراحل مطالعه و ارزیابی دقیق ۶۲ مقاله مورد تایید قرار گرفت. با بررسی و مطالعه کامل این مقالات ۸ مفهوم شناسایی شد. بعد از اعمال نظر اساتید و سه نفر از متخصصان داده در نهایت ۵۵ مولفه برای سوال پژوهش استخراج گردید. در مرحله دوم به منظور دریافت نظر خبرگان از روش دلفی فازی استفاده شد. به همین منظور از خروجی فراترکیب گویه‌های لازم برای طراحی پرسشنامه دلفی فازی تامین شد و این فرایند تا جایی ادامه یافت که نظرات خبرگان در مورد پاسخ به سوالات به اجماع رسید. ۲۱ نفر از خبرگان که حداقل دارای یک مقاله پژوهشی در حوزه‌ی کیفیت داده بودند، انتخاب شدند و در نهایت ۱۴ پرسشنامه تکمیل شده برگردانده شد

یافته‌های اصلی: در پاسخ به سوالات ابعاد کیفیت داده برای حکمرانی داده تعداد ۴۶ شاخص: کامل بودن، زمان‌مند بودن، ارتباط، دسترس پذیری، مطابقت با قوانین و استانداردها، محرمانه بودن، تفسیرپذیری، افزونگی (قابلیت اضافه کردن)، شهرت و اعتماد پذیری، قابلیت ردیابی، ارزش (مقدار)، سادگی، بروزرسانی، مفهوم، نظم پذیری، پیوند پذیری، یکپارچگی مرجع، منحصر بفرد بودن، عینیت، امکان ممیزی، صحت، مقایسه پذیری، ثبات، امنیت دسترسی، رایج و شایع بودن، تمامیت، انطباق فراداده، مقبولیت، اعتبار، موجز بودن، کاربردپذیری، (قابلیت استفاده)، باور پذیری، قابلیت درک، قابلیت اطمینان، منطقی بودن، سازگاری، قابلیت بازیابی، تکرارپذیری، قابلیت نمایش مقادیر پوچ، تناسب، وضوح، ارزش افزوده، جامعیت، توسعه‌پذیری، الگوپذیری و تغییر پذیری شناسایی شدند

بحث و نتیجه‌گیری: بیش از ۸۰ درصد شاخص‌های حاصل از فراترکیب از نظر خبرگان پذیرفته شدند بنابراین سازمان‌ها و کسب و کارهای داده محور می‌توانند این شاخص‌ها را در اولویت سنجش کیفی داده‌های خود قرار دهند.

کلیدواژه‌ها: کیفیت داده، حکمرانی داده، ابعاد کیفیت داده، فراترکیب، دلفی فازی

Title: Identifying data quality indicators for data governance with a metasynthesis and fuzzy delphi approach.

Names of Authors: Sahar Bonyadi: Doctoral student of Information Science and Knowledge management; Islamic Azad University; Science and Research Unit; Tehran, Iran

sahar_bonyadi2005@yahoo.com

Najla Hariri: PhD in information science and Knowledge management; Professore; Islamic Azad University; Science and Research Unit; Tehran, Iran
nadjlahariri@gmail.com

Seyed Mehdi Taheri: PhD in information science and Knowledge management; Associate Professor; Allameh Tabatabai University; Tehran, Iran

Taherismster@gmail.com

Roya Pournaghi: Ph.D. in information science and Knowledge management ; assistant professor Research Institute of Information Science and Technology of Iran (Irandoc); Tehran, Iran

pournaghi@irandoc.ac.ir

Introduction: One of the most important issues in data management is data quality. Good and high-quality data is the basis for achieving the desired results from the systems. Researches related to data quality provide a complete classification of data quality dimensions; However, there are some differences in the definition of the most important dimensions according to the defined nature of quality. Six important categories of quality dimensions by Wand and Wang (1996); Wang and Strong (1996); Redman (1996); Jarek and others (1995); Bowie et al. (2003); Newman (2002) presented. By analyzing this classification, it becomes possible to define an initial set of data quality dimensions, including accuracy, comprehensiveness, stability, and timeliness, which are the focus of most authors.

However, there is no general agreement on the definition of a set of dimensions of data quality or the exact meaning of each dimension (Arshady et al., 2016). Due to the multiplicity of data quality dimensions and their diversity in use, this research examines data quality indicators from the perspective of Iranian experts.

Methodology: In the first stage of the metasynthesis method, searching for the key and keywords of data quality, mod. Data management, data governance and

data quality management in Air&Dec databases, Science Direct, Google Scholar, Springer, IEEE and ACM were conducted between the years (1995-2022) and the number of 268 related articles were identified and in the stages A detailed study and evaluation of 62 approved articles was decided. By reviewing and studying these articles fully, 8 concepts were identified. After applying the opinions of professors and three data experts, finally 55 components were extracted for the research question. In the second stage, fuzzy Delphi method was used to get the opinion of experts. For this purpose, the necessary items for the design of a fuzzy Delphi questionnaire were provided from the output of meta, and this process continued until the experts' opinions on the answers to the questions reached a consensus. 21 experts who had at least one research paper in the field of data quality were selected and finally 14 completed questionnaires were returned.

Main findings: In response to the questions of data quality dimensions for data governance, there are 46 indicators: completeness, timeliness, communication, accessibility, compliance with laws and standards, confidentiality, interpretability, redundancy (ability to add), reputation and trustworthiness, ability Traceability, value (value), simplicity, update, concept, regularity, linkability, integrity of reference, uniqueness, objectivity, auditability, correctness, comparability, consistency, security of access, commonality, completeness, metadata compliance , acceptability, validity, conciseness, applicability, (usability), believability, comprehensibility, reliability, reasonableness, consistency, retrievability, reproducibility, ability to display null values, appropriateness, clarity, added value, comprehensiveness, development Adaptability, model adaptability and changeability were identified

Discussion and conclusions: More than 80% of the indices obtained from the meta-combination were accepted by the experts, so our organization and data-oriented businesses can use these indicators as a priority for measuring the quality of their data.

Keywords: Data quality, data governance, data quality dimensions, metasyntheses, fuzzy delphi