

Provide Data Quality Management Model for Data Governance using Meta Synthesis

Sahar Bonyadi

PhD Candidate in Knowledge and Information Science;
Department of Knowledge and Communication Sciences;
Science and Research Branch; Islamic Azad University; Tehran, Iran;
Email: sahar_bonyadi2005@yahoo.com

Nadjla Hariri*

PhD in Knowledge and Information Science; Professor;
Department of Knowledge and Communication Sciences;
Science and Research Branch; Islamic Azad University;
Tehran, Iran Email: nadjlahariri@gmail.com

Seyed Mahdi Taheri

PhD in Knowledge and Information Science; Associate Professor;
Department of Knowledge and Information Science; Allameh
Tabataba'i University; Tehran, Iran Email: Taherismster@gmail.com

Roya Poornaghi

PhD in Knowledge and Information Science; Assistant Professor;
Iranian Research Institute for Information Science and Technology
(IranDoc); Tehran, Iran Email: poornaghi@irandoc.ac.ir

Received: 19, Dec. 2021 Accepted: 17, Aug. 2022

Abstract: All organizations use data to execute their processes. Organizational data is either generated by the organization itself or created and provided by other organizations. Data and information play an essential role in the decisions and functions of organizational processes. For this reason, data is part of the organization's resources and perhaps the most important source of organizations. The purpose of this study was to provide a data quality model for data governance. The research method was used in a meta-combination. Therefore, out of 268 sources found, during the meta-combination process, 62 articles were used using keywords such as; data, data management, data governance and data quality management were selected in IRANDOC, Science Direct, Google Scholar, Springer, IEEE and ACM databases between 1995-2022. In this study, first a code was considered for all factors extracted from previous studies and then, considering the concept of each code, they were classified in a similar concept. In this

**Iranian Journal of
Information
Processing and
Management**

**Iranian Research Institute
for Information Science and Technology
(IranDoc)**

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 38 | No. 4 | pp. ??-??

Summer 2023

<https://doi.org/jipm.38.4>



* Corresponding Author

way, the concepts of research were identified. Two coders were used to control the extracted codes and categories, and the desired index in this field is the Kappa index. Based on the analysis performed using the content analysis method, a total of 12 main categories (data attribute, data, data fable, data value, initial data value, data pattern, data set, data access, data composition, data formatting, metadata and data objectivity) and 47 sub-categories for data quality management for data governance were identified.

Keywords: Data Quality, Data Governance, Quality Management, Meta Synthesis

ارائه مدل مدیریت کیفیت داده برای حکمرانی داده با استفاده از روش فرا ترکیب

سحر بنیادی

دانشجوی دکتری علم اطلاعات و دانش شناسی؛
واحد علوم و تحقیقات؛ دانشگاه آزاد اسلامی؛ تهران، ایران؛
sahar_bonyadi2005@yahoo.com

نجلا حریری

دکتری علم اطلاعات و دانش شناسی؛ استاد؛
واحد علوم و تحقیقات؛ دانشگاه آزاد اسلامی؛ تهران، ایران؛
nadjlahariri@gmail.com

سید مهدی طاهری

دکتری علم اطلاعات و دانش شناسی؛ دانشیار؛
دانشگاه علامه طباطبائی؛ تهران، ایران؛
Taherismster@gmail.com

رؤیا پورنقی

دکتری علم اطلاعات و دانش شناسی؛ استادیار؛
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)؛
تهران، ایران pournaghi@irandoc.ac.ir



مقاله برای اصلاح به مدت ۳ ماه نزد پدیدآوران بوده است.

پذیرش: ۱۴۰۱/۰۵/۲۶

دریافت: ۱۴۰۰/۰۹/۲۸

نشریه علمی | رتبه بین المللی
پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

شاپا (جایی) ۲۲۵۱-۸۲۲۳

شاپا (الکترونیکی) ۲۲۵۱-۸۲۳۱

نمایه در SCOPUS، ISI، و LISTA

jipm.irandoc.ac.ir

دوره ۳۸ | شماره ۴ | صص ۱۵۳۳-۱۵۶۴

تابستان ۱۴۰۲

<https://doi.org/jipm.38.4>



چکیده: تمام سازمان‌ها برای اجرای فرایندهایشان از داده استفاده می‌کنند. داده‌های سازمانی یا توسط خود سازمان تولید شده‌اند یا سازمان‌های دیگری آن‌ها را ایجاد کرده و در اختیارشان قرار داده‌اند. داده‌ها و اطلاعات نقش اساسی در تصمیم‌گیری‌ها و عملکردهای فرایندهای سازمانی دارند. به همین دلیل، داده‌ها بخشی از منابع سازمان و شاید مهم‌ترین منبع سازمان‌ها محسوب می‌شوند. از این‌رو، هدف از انجام این پژوهش ارائه مدل کیفیت داده برای حکمرانی داده است. در این پژوهش روش فرا ترکیب به کار گرفته شد. از ۲۶۸ منبع یافت‌شده طی مراحل فرا ترکیب، ۶۲ مقاله با استفاده از کلیدواژه‌هایی از جمله داده، مدیریت داده، حکمرانی داده و مدیریت کیفیت داده در پایگاه‌های «ایرانداک»، Science Direct، Google، Springer، Scholar و ACM در بین سال‌های (۱۹۹۵-۲۰۲۲) انتخاب گردید. در این پژوهش ابتدا برای تمام عوامل استخراج‌شده از مطالعات پیشین، کدی در نظر گرفته و سپس، با در نظر گرفتن مفهوم هر یک

از کدها، در یک مفهوم مشابه دسته‌بندی شدند. به این ترتیب، مفاهیم پژوهش مشخص شدند. برای کنترل کدها و مقوله‌های استخراجی از دو کدگذار استفاده شد و شاخص مورد نظر در این زمینه، شاخص «کاپا»ست. بر اساس تحلیل‌های صورت گرفته به کمک روش تحلیل محتوا، در مجموع، تعداد ۱۲ مقوله اصلی (ویژگی داده، فایل داده، ارزش داده، ارزش اولیه داده، الگوی داده، مجموعه داده، دسترسی به داده، ترکیب داده، قالب‌بندی داده، فراداده و عینی بودن داده) و ۴۷ مقوله فرعی برای مدیریت کیفیت داده برای حکمرانی داده کشف و مشخص شدند.

کلیدواژه‌ها: کیفیت داده، حکمرانی داده، مدیریت کیفیت، فراترکیب

۱. مقدمه

اهمیت دادن به کیفیت داده در سازمان‌ها نشانگر دید عمیق مدیران به این مسئله و استقرار حاکمیت داده در یک سازمان محسوب می‌شود. مطالعات وسیعی جهت شناخت ابعاد کیفیت داده و ارزیابی آن‌ها در نظام‌های اطلاعاتی صورت گرفته است. این مطالعات نشان می‌دهد که ارزیابی ابعاد مختلف کیفیت داده با استفاده از طراحی و تبیین مدل‌ها، شیوه‌ها، ابزار و چارچوب‌های مختلف انجام شده است (رحیمی ۱۳۹۵؛ Liaw et al. 2013). کیفیت داده‌های سازمانی، اصلی مهم برای رسیدن به بهره‌وری سازمان‌هاست؛ چرا که داده‌ها به‌عنوان بخشی از دارایی‌های سازمان‌ها به حساب آمده و مبنایی برای تصمیم‌گیری‌های آن‌ها هستند. با سازماندهی و مدیریت اطلاعات می‌توان تصمیمات صحیحی در جهت استفاده بهینه از موجودیت‌های سازمانی گرفت. این دارایی‌های سازمانی برای مدیریت بهینه، به نوعی حکمرانی نیاز دارند تا از این رهگذر این دارایی‌ها را با اهداف راهبردی سازمان همراستا نمایند. به همین دلیل، سازمان‌ها به سراغ برنامه‌های حکمرانی داده می‌روند. یک سازمان، فعالیت‌های حکمرانی داده^۱ را جهت تکمیل مدیریت اطلاعات استفاده می‌کند. این فعالیت‌ها شامل طراحی و قوانین اجرایی برای کنترل فعالیت مدیریت داده‌ها به همراه تخصیص حقوق تصمیم‌گیری و پاسخگویی است؛ یعنی حکمرانی داده‌ها سیستمی متشکل از حقوق تصمیم و قابلیت پاسخگویی با استفاده از فرایندهای مرتبط با اطلاعات است که بر اساس مجموعه‌ای از مدل‌های توافق‌شده تشریح می‌کند که چه کسی می‌تواند چه اقداماتی را با چه اطلاعاتی، در چه زمانی، با چه شرایط و مقتضیاتی و با استفاده از چه روشی انجام دهد. حکمرانی داده، مجموعه‌ای

از فرایندهای موجود در یک سازمان است که دو هدف را دنبال می‌کند: تعیین اهداف داده‌های سازمان، و تعیین یک چارچوب تصمیم‌گیری مناسب که برای رسیدن به این اهداف مناسب است. چارچوب، شامل وظایف، نقش‌ها و مسئولیت‌هایی است که همگی در محدوده‌های دارایی‌های داده، کیفیت داده، ابرداده، دسترسی به داده‌ها و چرخه حیات داده انجام می‌شوند. جنبه‌های دیگری نیز در بعضی از تعاریف حکمرانی وجود دارد، مانند مدیریت ریسک یا برنامه‌های آموزش سازمانی (Loshin 2010). سیاست‌های بالا منجر به شکل‌گیری رویه‌ای خاص می‌گردد که از آن تحت عنوان حاکمیت داده یاد می‌کنند. حاکمیت داده، تشریح‌کننده فرایندهایی برای تعریف سیاست‌های حوزه داده در سازمان، فرایندهای مشخص‌کننده روش بهره‌برداری از سیاست‌های مزبور، نوع ساختار سازمانی دربرگیرنده کنسول‌های حاکمیت داده و ناظران داده جهت کنترل و اطمینان از انطباق بین سیاست‌ها و داده‌هاست. حاکمیت داده یک برنامه مدیریتی چندوظیفه‌ای است که مهم‌ترین هدف آن برخورد با داده به‌عنوان یک سرمایه سازمانی است. این کار از طریق به‌کارگیری مجموعه‌ای از سیاست‌ها، استانداردها، فرایندها، افراد و فناوری که برای مدیریت داده حیاتی است، دنبال می‌شود. به‌طور کلی، اهداف کلی حاکمیت داده عبارت‌اند از: به جریان انداختن داده‌های باکیفیت از طریق ایجاد مکانیسم‌ها، فرایندها و محیط مناسب برای تبدیل شدن آن‌ها به بینش تجاری؛ حفاظت از نیازهای ذی‌نفعان داده؛ آموزش مدیران و کارکنان به‌منظور تطبیق با رویکردهای کلی مسائل داده؛ ایجاد فرایندهای استاندارد و قابل تکرار؛ کاهش هزینه‌ها و افزایش اثربخشی از طریق هماهنگی بین فعالیت‌های در ارتباط با داده؛ استاندارد کردن تعاریف داده در تمامی سازمان؛ حصول اطمینان از شفافیت فرایندها؛ امکان اخذ تصمیمات بهتر؛ تضمین بقا از طریق توجه به ریسک‌ها و آسیب‌پذیری‌ها؛ تطابق؛ امنیت؛ و حریم خصوصی با کمک داده‌ها. اما از کاربردهای حاکمیت داده می‌توان به موارد زیر اشاره کرد:

۱. در سامانه‌های بانکی (در چرخه حیات راهکارهای نرم‌افزارهای بانکی): تجمع داده‌ها در پایگاه داده‌های بانک‌ها، دارایی‌های بانک‌ها محسوب می‌شود که آن‌طور که باید و شاید مورد بهره‌برداری قرار نمی‌گیرند. انباشت داده‌های تراکشی‌های الکترونیکی، انباشت داده‌های مشتریان اعتباری، انباشت داده‌های تسهیلاتی و طرح‌های سرمایه‌گذاری و انباشت داده حساب‌ها و اوراق از دلایل عمده این تجمع داده هستند (Alhassan, Sammon & Daly 2019)، و چون استفاده از روش‌های تحلیلی و تصمیم‌سازی

در بانک‌ها بسیار کم است و بسیاری از سامانه‌های مفید در امر بانکداری وابسته به داده‌های یکپارچه و دارای کیفیت داده هستند و نیز کنترل کیفیت داده فاقد فرایندهای مناسب و دقیق در بانک‌هاست، در حال حاضر راهکارهای مواجهه با حجم بزرگی از داده‌ها یکی از نیازهای جدی در بانکداری است. بنابراین، بانک‌ها باید فرایند جامعی برای حاکمیت داده و مدیریت اطلاعات داشته باشند و روش‌های کنترل کیفیت خود را تقویت نموده و منابع انسانی خود را در این زمینه‌ها تقویت نمایند و زمینه و بستر مواجهه با حجم بزرگ داده‌ها را فراهم آورند.

۲. تصمیم‌گیری داده‌محور در هر کسب و کار: در یک کسب و کار، باید یک تفکر اجتماعی تکنیکی به موضوع جریان داده‌ها داشت. موضوع داده‌شناسی یکی از موضوعات کلیدی در حکمرانی داده‌ها در یک کسب و کار است. حکمرانی داده‌ها باید منجر به سود مالی شود و به فرایند پول‌سازی از داده‌ها منجر گردد.

۳. در حوزه سلامت: برنامه حکمرانی داده برای اطلاعات نیروی کار مراقبت‌های بهداشتی در مورد برنامه‌های مراقبتی که روی خط‌مشی ملی مراقبت‌های بهداشتی و دستاورد کلی خدمات بهداشتی در یک کشور تأثیر می‌گذارند، مورد استفاده قرار می‌گیرد که می‌تواند جهت حفظ کیفیت داده‌ها، کسب رضایت مشتری، اطمینان یافتن از کنترل و امنیت اطلاعات و دستیابی به کارایی عملکردی، به‌عنوان پیشرانان حکمرانی داده در مقامات نظارتی مورد استفاده قرار گیرد (Were & Moturi, 2017). همچنین نیاز به جمع‌آوری، بازیابی و استفاده از اطلاعات رکوردهای بهداشتی در قالب الکترونیک برای ارائه انعطاف‌پذیری بیشتر در حوزه سلامت بسیار مهم است، زیرا این امر می‌تواند بازیابی و نمایش داده‌ها را در مکان‌ها و فرمت‌های مختلف بدون توجه به جایی که داده‌ها در آن جمع‌آوری شده‌اند، امکان‌پذیر سازد. استفاده از داده‌های نگهداری‌شده از سوابق تولیدشده در سیستم‌های بالینی می‌تواند به کاهش زمان لازم برای به‌دست آوردن دانش و به‌عنوان نیاز به اطلاعات خاص تحقیق منجر شود. این تنها در صورتی امکان‌پذیر است که اصول حاکمیت داده اجرایی شود و در این حالت، حاکمیت داده‌ها، نمایش محتوای سیستم‌ها و استفاده از آن اطلاعات را تحت تأثیر قرار خواهد داد.

امروزه، داده‌ها اغلب با سیستم‌های اطلاعاتی مرتبط هستند و طی انجام فرایندهای کاری که به‌وسیله سیستم‌های اطلاعاتی مدیریت می‌شوند، ایجاد می‌گردند. سودآوری

بیشتر و پرهیز از هزینه‌های غیرضروری یکی از اهداف اصلی سازمان‌های تولیدکننده است و این مهم محقق نمی‌شود مگر با تولید محصولات باکیفیت که با استفاده از سیستم‌های کارا و برنامه‌ریزی‌های صحیح قابل دستیابی هستند. لازمه تحقق این موارد داشتن داده‌های باکیفیت است؛ داده‌ای باکیفیت که برای استفاده در فرایندها و تصمیم‌گیری‌ها مناسب باشد و الزامات مورد نیاز برای استفاده مورد نظر را برآورده کند.

مدیریت داده‌ها با استفاده از ابزارهای مرتبط با فناوری، پیشرفت چشمگیری داشته است؛ اما کیفیت داده‌ها به‌عنوان بخش مهمی از مقوله حکمرانی داده‌ها در سازمان‌ها زیاد مورد توجه قرار نگرفته است. ازاین‌رو، این تحقیق به بررسی مدیریت کیفیت داده برای حکمرانی داده در یک سازمان بر اساس مدل «داما»^۱ می‌پردازد. مهم‌ترین رسالت یک برنامه مدیریت کیفیت داده برنامه‌ریزی، پیاده‌سازی و کنترل فعالیت‌هایی جهت سنجش ارزیابی، بهبود و اطمینان از مناسب بودن داده برای سازمان با توجه به اهداف استراتژیک آن است. اگر نتوان داده حیاتی یک سازمان را شناسایی کرد، نمی‌توان ادعای حکمرانی داده را در آن سازمان داشت و در مسیر بهبود آن گام برداشت.

مدیریت کیفیت داده‌ها^۲ در سازمان شامل تمام شیوه‌ها، روش‌ها و سیستم‌های تجزیه و تحلیل، بهبود و حفظ کیفیت داده‌هاست. مدیریت کیفیت داده‌ها اساساً به دنبال به حداکثر رساندن ارزش داده (به‌عنوان مثال داده مشتری، داده کارپردازی و داده مواد) است (DAMA 2008). اگرچه کیفیت داده به‌طور گسترده‌ای به‌عنوان عامل موفقیت استراتژیک شناخته شده است، مدیریت کیفیت داده‌ها در اکثر شرکت‌ها در مراحل اولیه بلوغ به سر می‌برد (Ofner, Otto & Österle 2013).

مدیریت کیفیت داده‌ها یک مدل بدنه دانش درباره مدیریت داده را منتشر کرده است که شامل ۱۰ مقوله است. در قلب این مدل حکمرانی، داده‌ها قرار دارد که از کلیه مقوله‌ها پشتیبانی می‌کند. یکی از مهم‌ترین مقوله‌های بدنه دانش مدل «داما»، مدیریت کیفیت داده است که به دلیل اهمیتش در تصمیم‌گیری سازمان‌ها بسیار مورد توجه قرار گرفته است (DAMA 2008).

با توجه به مطالب پیش گفته، مدیریت داده‌ها با استفاده از ابزارهای مرتبط با

1. DAMA

2. data quality management (DQM)

فناوری، پیشرفت چشمگیری داشته است؛ اما کیفیت داده‌ها به‌عنوان بخش مهمی از مقوله حکمرانی داده‌ها در سازمان‌ها، زیاد مورد توجه قرار نگرفته است. با توجه به اینکه، در این پژوهش به بررسی حکمرانی داده از طریق مدیریت داده پرداخته می‌شود، بنابراین، برای رسیدن به یک مدل و الگوی مناسب، به بررسی سیستماتیک مرور پژوهش‌های مختلف در زمینه حکمرانی داده نیاز هست. از این‌رو، این تحقیق به بررسی مدیریت کیفیت داده برای حکمرانی داده در یک سازمان بر اساس روش فراترکیب می‌پردازد. مهم‌ترین رسالت یک برنامه مدیریت کیفیت داده برنامه‌ریزی، پیاده‌سازی و کنترل فعالیت‌هایی جهت سنجش ارزیابی، بهبود و اطمینان از مناسب بودن داده برای سازمان با توجه به اهداف استراتژیک آن است. اگر نتوان داده حیاتی یک سازمان را شناسایی کرد، نمی‌توان ادعای حکمرانی داده را در آن سازمان داشت و در مسیر بهبود آن گام برداشت.

۲. پیشینه پژوهش

«درزی خلردی، رضوی و قیاسی» در پژوهشی به بلوغ حکمرانی داده در کتابخانه‌های دانشگاهی استان مازندران پرداخته‌اند. روش تحقیق به‌صورت آمیخته انتخاب شده و بر اساس یافته‌های پژوهش و بررسی کدها و رتبه‌بندی شاخص‌های پژوهش، کد سیستم‌های سخت‌افزاری و نرم‌افزاری در رتبه اول قرار دارد که بیانگر اهمیت سیستم‌های سخت‌افزاری و نرم‌افزاری بوده و از شرایط لازم در بلوغ کتابخانه‌های دانشگاهی است. نظام‌های اطلاعاتی، کاربران و کیفیت داده، مجموعه‌سازی و سازماندهی، مدیریت، اینترنت و شبکه، سازمان و کارکنان به‌ترتیب، در اولویت‌های بعدی اهمیت قرار دارند. همچنین بر اساس یافته‌ها سطح بلوغ کتابخانه‌های دانشگاهی استان مازندران در سطح ۲ قرار داشته و تا رسیدن به سطح ایده‌آل (سطح ۵) فاصله نسبتاً زیادی دارند (۱۴۰۰).

«درودی، جلیلی و فامیل روحانی» در پژوهشی به بررسی نگرش کتابداران در پیاده‌سازی و اجرای راهبردهای حاکمیت داده در برنامه‌ریزی سازماندهی منابع اطلاعاتی در کتابخانه‌ها پرداخته‌اند. روش پژوهش به‌صورت پیمایشی تحلیلی انتخاب شده است. یافته‌های پژوهش حاکی از آن است که نگرش کتابداران در خصوص مؤلفه‌های هفت‌گانه کیفیت داده، کاربرد داده، فرایند داده، استاندارد داده، یکپارچگی داده، امنیت داده، و مدل‌سازی داده درباره سازماندهی اطلاعات در کتابخانه‌های مرکزی دانشگاه آزاد اسلامی، به‌جز عنصر یکپارچگی داده با سطح معناداری ۰/۰۸ و بقیه مؤلفه‌ها در ارتباط با سطح تحصیلات،

رشته تحصیلی و سابقه کاری مدیران و کتابداران ارتباط معناداری ندارد. بالا بردن کیفیت سازماندهی اطلاعات، آموزش نیروی انسانی، بهره‌گیری از ابزارهای استاندارد در کنار برنامه‌ریزی راهبردی کتابخانه از عوامل مهم پیاده‌سازی حاکمیت داده است (۱۴۰۰).

«غفاری، آذر و شباک» در مطالعه‌ای به بررسی نگاشت علی مدیریت محصول آماری با رویکرد کیفیت داده پرداختند. یافته‌های این مطالعه حاکی از آن بود که عناصر تضمین انسجام، تضمین درستی و جامعیت از شدت اثر بالایی برخوردار بوده و بر سایر عناصر بعد مدیریت محصول آماری از اثرگذاری بالایی برخوردار هستند. این است که مدیران مراکز آماری جهت افزایش و ارتقای سطح کیفیت داده‌ها لازم است عناصر مذکور را در اولویت برنامه‌ریزی خود قرار دهند (۱۳۹۷).

«ارشادی، نصیری و شیرانی» در مطالعه‌ای به بررسی طراحی مدل کیفیت فراداده در سامانه ثبت پایان‌نامه/ رساله‌های دانش‌آموختگان داخل و خارج کشور پرداختند. محققان در مطالعه خود بحث می‌کنند که سامانه ثبت پایان‌نامه و رساله‌های دانش‌آموختگان کل کشور به عنوان یک سامانه‌های حیاتی در ثبت و اشاعه مدارک علمی بوده و نقشی انکارناپذیر در توسعه پژوهش‌ها دارد. از این رو، سطح کیفیت مطلوب سامانه تأثیر ویژه‌ای در ارتقای کیفیت پژوهش‌های محققان خواهد داشت. نام، کد ملی، مقطع تحصیلی، رشته دانشگاهی، مشخصات استاد راهنما، مشخصات پیشنهاد و پایان‌نامه از مهم‌ترین اقلام اطلاعاتی هستند. به منظور بهبود سطح کیفیت در این سامانه تعیین وضعیت مطلوب هر یک از اقلام اطلاعاتی سامانه ثبت، اقدامی بنیادین محسوب می‌گردد (۱۳۹۶).

«خلیلی جعفرآباد» در مطالعه‌ای به بررسی تغییرات حوزه کیفیت داده با استفاده از تحلیل کلمات کلیدی پرداخت. بر اساس این مطالعه، که جدیدترین حوزه مرتبط با کلان‌داده است، مباحث یکپارچه‌سازی و داده گمشده در این حوزه از اولویت بالاتری نسبت به بقیه حوزه‌ها برخوردار است (۱۳۹۶).

«سلطانی‌نژاد و گودرزی» در مطالعه‌ای به بررسی فناوری اطلاعات و تحول در مفهوم حکمرانی خوب پرداختند. نتایج، نشان‌دهنده تحول مفهوم حکمرانی خوب در سایه ایجاد و تأسیس ابزارها و روش‌های نوین خدماتی با استفاده از فناوری اطلاعات و ارتباطات بود (۱۳۹۶).

«کارگر و محمدی» در مطالعه‌ای به بررسی مدلی به منظور ارزیابی کیفیت بر روی کلان‌داده پرداختند. محققان در مطالعه خود بحث می‌کنند که در عصر اطلاعات، داده‌های

دیجیتال نقش مهمی در همه ابعاد زندگی بازی می‌کنند. از ارسال یک پیام تا بزرگ‌ترین تصمیمات سازمانی، همه و همه وابسته به داده‌های ثبت‌شده در سامانه‌های اطلاعاتی هستند. بنابراین، کیفیت داده به‌عنوان تضمین‌کننده صحت و دقت این عملیات و تصمیم‌ها از اهمیت بسیار بالایی برخوردار است. در بعد سازمانی، کیفیت پایین داده می‌تواند منجر به زیان‌های چشمگیر و حتی شکست سازمان در فضای رقابتی امروز گردد. کیفیت پایین داده موجب کاهش بهره‌وری، اخذ تصمیم‌های اشتباه، افزایش هزینه‌ها و ... می‌گردد. کلان‌داده یک اصطلاح برای مجموعه داده است؛ بدین معنا که حجم عظیمی از داده‌های غیرساخت‌یافته، نیم‌ساخت‌یافته و ساخت‌یافته از منابع مختلف باید جمع‌آوری و ذخیره شوند. در حال حاضر، تجزیه و تحلیل جامع و پژوهشی در خصوص استانداردهای کیفی و روش‌های ارزیابی کیفیت برای کلان‌داده وجود ندارد. در این پژوهش پس از پیش‌پردازش بر روی مجموعه داده، خوشه‌بندی بر روی داده‌های پاک‌سازی شده صورت گرفت (۱۳۹۶). در ادامه، معیارهایی جهت سنجش کیفیت جامعه آماری ارائه گردید. بنابراین، ابتدا باید کیفیت را اندازه‌گیری کرد. در این راستا ابتدا معیارهای کیفیت شناسایی و سپس با روش‌های عددی کمی میزان کیفیت داده در بستر هادوپ^۱ اندازه‌گیری می‌گردد. در نهایت، خوشه‌بندی تاپدیان^۲ بر اساس کامین^۳ روی مجموعه داده اجرا می‌گردد.

«عبدالوند و بوبه‌رژ» در مطالعه خود به بررسی رویکردی سیستماتیک به چالش کیفیت داده در راهبرد مشتری‌محور در صنعت بانکداری پرداختند. این پژوهش با رویکردی نظام‌مند و با استفاده از روش‌شناسی سیستم نرم^۴ به بررسی این معضل پرداخته است. در این پژوهش بر طبق گام‌های روش‌شناسی به بیان مشکل و استخراج تصویر غنی پرداخته شده است. در ترسیم تصویر غنی با مدیران و کارکنان بانک و نیز دانشگاهیانی که به تحقیق در این حوزه پرداخته‌اند، مصاحبه شده است. این پژوهش استفاده از شش سیگما را برای بهبود کیفیت داده توصیه کرده و بر طبق این روش بهبود مستمر به ارائه مدلی برای بهبود مستمر کیفیت داده مشتریان پرداخته است (۱۳۹۵).

«محبوب و سمیه‌ساز» در پژوهشی به بررسی «کیفیت داده‌ها پیش‌نیاز مدیریت

1. Hadoop

2. Top Down

3. K-Mean

4. Soft Systems Methodology (SSM)

منابع سازمان» پرداختند. محققان در مطالعه خود بحث می‌کنند که سازمان‌ها می‌توانند با مدیریت و برنامه‌ریزی منابع خود به بهره‌وری رسیده و هزینه‌ها را کاهش دهند. منابع اطلاعاتی به‌عنوان یکی از مهم‌ترین منابع سازمانی باید بتواند نیازهای سازمان را در برنامه‌ریزی‌ها و تصمیم‌سازی‌ها برآورده کند. کیفیت داده‌ها مبحثی است که به شیوه‌های بهبود و مدیریت منابع سازمان و منابع اطلاعاتی مرتبط با آن‌ها می‌پردازد. استاندارد ISO8000 با موضوع کیفیت داده‌ها و استاندارد ISO22745 با موضوع یکپارچگی و سیستم‌های اتوماسیون صنعتی در این مقاله مورد بحث واقع شده است (۱۳۹۴).

«مودزا» در پژوهشی به بررسی «مدیریت داده‌های پژوهشی: مروری بر تجربه کتابخانه دانشگاهی امارات» پرداخت. این پژوهش به‌صورت مرور ادبیات تحقیق و همچنین بررسی نظرات کتابداران امارات صورت گرفت. نتایج نشان داد که فعالیت در زمینه مدیریت داده‌های پژوهشی در کتابخانه‌های امارات نسبت به گذشته روبه‌افزایش است و آگاهی از اهمیت مدیریت داده برای همه کتابداران مهم و ضروری است (Mavodza 2022).

«حیکمواتی، سانتوزا و هیدایه» در پژوهشی به بررسی بهبود کیفیت داده و حاکمیت داده با استفاده از مدیریت داده اصلی پرداختند. این پژوهش به روش مرور ادبیات انجام شد. نتایج بررسی نشان می‌دهد که مدیریت داده‌های اصلی می‌تواند بر مشکلات کیفیت داده‌ها از طریق فرایند مدیریت داده‌های اصلی ناشی از داده‌های منشأگرفته از منابع مختلف پراکنده غلبه کند. مدیریت داده‌های اصلی، سازمان‌ها را تشویق می‌کند که مدیریت داده‌ها را با تنظیم نقش‌ها و مسئولیت‌های بازیگران تجاری و کارکنان فناوری اطلاعات که از طریق حاکمیت داده‌ها مستند شده‌اند، بهبود بخشند. ارزیابی موفقیت پیاده‌سازی مدیریت داده‌های اصلی می‌تواند توسط سازمان‌ها برای بهبود کیفیت داده‌ها و حاکمیت داده‌ها با پیروی از چارچوب مطرح شده انجام شود (Hikmawati, Santosa & Hidayah 2021).

«کولتای» در پژوهشی با عنوان «حکمرانی داده، سواد داده و مدیریت کیفیت داده» نتیجه گرفت که حکمرانی داده یک خدمت مبتنی بر استاندارد و فرایندی تکرارپذیر است که برای شفاف‌سازی فرایند داده‌محور و کاهش هزینه طراحی شده است، سودمند است. همچنین چون به قوانین، سیاست‌ها، استانداردها، حقوق تصمیم، جوابگویی و روش‌های اجرایی برمی‌گردد، مفید است (Koltay 2016).

«الشریده» در پژوهش خود که به حکمرانی داده در گزارش‌های سلامت الکترونیک به‌صورت نظام‌مند پرداخت، بیان می‌کند که حکمرانی داده به‌عنوان یک نظم‌دهنده برای

پوشش دادن در هر سازمانی قابل استفاده است. در نهایت، پنج بُعد از کیفیت داده توسط محقق در EHR مشخص گردید که شامل تکمیل، صحت، برازندگی، اعتبار، و جاری بودن است (Al-Sharida 2015).

«دای» در پژوهشی چارچوب کنترل و مدل مدیریت کیفیت داده برای پشتیبانی تصمیم‌گیری در حوزه سلامت را پیشنهاد کرد که یک روش بازرسی کیفیت داده، فرایند و تشخیص اجرایی شدن برای داده‌های سلامت با انواع گوناگون بر پایه مدل پیشنهادی طراحی شده است (Dai 2015).

«کرهونن» در پژوهشی با عنوان «طراحی ساختار حکمرانی داده: یک دیدگاه سازمانی»، روی جنبه‌های پاسخگویی در مورد حکمرانی داده‌ها متمرکز شد. او بیان کرد که داده‌های با کیفیت بالا یک منبع کلیدی با ارزش کسب و کار هستند، اما مسائل مربوط به کیفیت داده‌ها در سازمان‌ها بیشتر به‌طور ناکافی بررسی و مدل حکمرانی داده مناسب برای آن انتخاب می‌شود. وی به دنبال رویکرد علمی طراحی، مسئله تحقیق زیر را مورد بررسی قرار داد: «نقش‌های حاکمیت اطلاعات و حسابدهی لازم در ابتکار عمل حاکمیت داده سازمان در سطح کل سیستم چیست؟» او بر اساس نظریات پیشین و ادبیات موجود، پوشش سازمانی نقش‌های حکومتی رایج را تجزیه و تحلیل کرده و پیشنهاد داد که مدل حکمرانی چابک^۱ چگونه می‌تواند به عنوان مبنایی برای طراحی یک ساختار حکومتی استفاده شود تا اطمینان از نقش‌ها و حسابدهی‌های لازم در تمام شرکت‌ها در سطوح سازمانی مناسب و موقعیت‌های خاص مورد توجه قرار بگیرند (Korhonen et al. 2013).

«ایگان» در پژوهش خود با عنوان «تأثیر کیفیت داده‌ها از طریق حکمرانی داده: مطالعه موردی در صنعت خدمات مالی» که یک مطالعه موردی در مورد شرکت‌های بزرگ خدمات مالی جهت بررسی سیاست‌ها و روش‌های حکمرانی داده است، به دنبال آن است که برخی از اطلاعات را در ایجاد حکمرانی داده، فرایندهایی برای اطمینان از کیفیت داده‌ها، فناوری و افراد درگیر برای کمک به فرایندها و همچنین استفاده از حکمرانی داده در تصمیم‌گیری آشکار کند. وی همچنین به برخی سؤالات اساسی درباره حاکمیت داده، مانند قابلیت ماندگاری رکورد، مالکیت و مسئولیت‌های داده‌ها و نحوه به کارگیری مطلوب حوزه‌های حاکمیت داده جواب می‌دهد (Egan 2011).

1. agile governance model

«ایون، شانکارانارایانان و برگر» یک مدل برای مدیریت کیفیت داده هزینه-اثربخش در یک محیط واقعی مدیریت ارتباط با مشتری ارائه کرد که نشان می‌داد تمامی پارامترهای مدل قابل اندازه‌گیری هستند، اما با سطح کیفی بهینه و سودمندی مطلوب لزوماً امکان‌پذیر منطبق نیستند (Even, Shankaranarayanan & Berger 2010).

و «ونده» در مقاله‌ای با عنوان «مدلی برای حکمرانی داده - سازماندهی پاسخگویی برای مدیریت کیفیت داده»، بیان می‌کند که سازمان‌ها، به مدیریت کیفیت داده‌ای^۱ نیاز دارند که دیدگاه‌های مبتنی بر کسب و کار و فنی را با یکدیگر ترکیب کند تا به چالش‌های استراتژیک و عملیاتی داده‌های شرکت‌های با کیفیت بالا پاسخ دهند و آن چالش‌ها را برطرف نمایند. تاکنون، شرکت‌ها مسئولیت‌های مدیریت کیفیت داده را اغلب به بخش‌های فناوری اطلاعات واگذار کرده‌اند. مدل حکومت داده، وظایف کیفیت داده و نوع تعامل آن‌ها را با فعالیت‌های مدیریت کیفیت داده نشان می‌دهد. شرکت‌های سهامی می‌توانند ساختار مدل حکومت داده‌های خاص شرکت خود را بر اساس این یافته‌ها طراحی نمایند (Wende 2007).

بررسی و تحلیل پیشینه داخلی و خارجی پژوهش نشان از آن دارد که هر کدام از پژوهش‌ها از جنبه‌های مختلف به مدیریت کیفیت داده و یا حکمرانی داده پرداخته‌اند. به عبارتی، تنوع پژوهش‌های انجام‌شده در زمینه مدیریت کیفیت داده زیاد بوده و هیچ‌کدام از پژوهش‌ها به‌طور مستقیم از طریق روش فراترکیب به بررسی این موضوع نپرداخته‌اند. تنها پژوهش Egan (2011) و Wende (2007) به حکمرانی داده از طریق کیفیت داده پرداخته‌اند که این پژوهش‌ها نیز به دلیل روش انجام پژوهش، نوع مطالعه، قلمرو مکانی و زمانی با موضوع پژوهش حاضر متفاوت هستند.

۳. روش

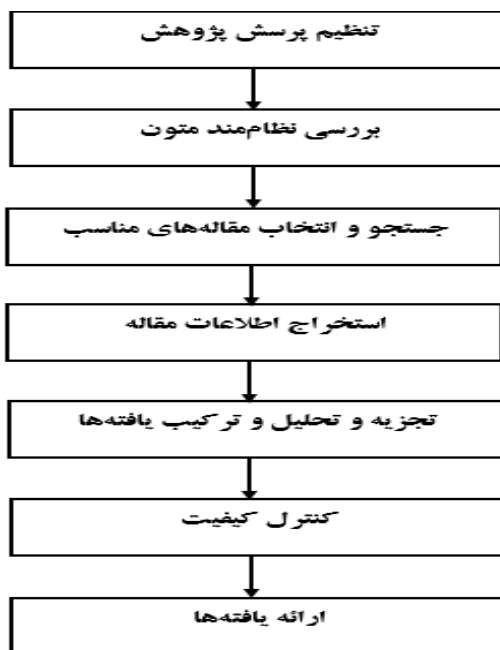
در این پژوهش فراترکیب^۲ به عنوان یک روش مناسب برای دستیابی به یک ترکیب جامع از عوامل تشکیل‌دهنده مدیریت کیفیت داده معرفی می‌شود. به‌طور کلی، ترکیب و تلفیق مطالعات اولیه کیفی و نتایج آن‌ها را فراترکیب می‌نامند. این دسته از مطالعات مروری ساختارمند از طریق استخراج کدها، تعابیر و مفاهیم ارائه‌شده در اطلاعات کیفی با

1. data quality management (DQM)

2. meta synthesis

متدولوژی یکسان و همسان‌سازی و تلفیق آن‌ها، تعابیر و مفاهیم جدید و جامع‌تری را ارائه می‌کنند. به عبارت دیگر، فراترکیب در جست‌وجوی یکی کردن نتایج مطالعات کیفی جهت کشف مفاهیم اساسی آن‌ها و ترجمه این مفاهیم به یک مفهوم نهایی و واحد است. فراترکیب به گونه‌ای ارائه می‌شود که همزمان نتایج مطالعات اولیه در آن قابل ردیابی باشد. نتیجه جدید حاصل از مطالعات اولیه این است که گرچه گاهی از متاسنتر تحت عنوان «متاآنالیز مطالعات کیفی» یاد می‌شود، اما این دو رویکرد دارای تفاوت‌هایی در اهداف خود هستند. هدف از متاآنالیز، تجمیع اطلاعات مطالعات مشابه و تقویت تفسیرهای علت و معلولی و یا مورد رابطه دو متغیر است که با توصیف دقیق‌تر شاخص‌های آماری همراه است، در حالی که هدف از فراترکیب، ارائه توضیح و کمک به درک جامع‌تر پدیده‌هاست.

بنابراین، فراترکیب مستلزم این است که پژوهشگر یک بازنگری دقیق و عمیق انجام داده و یافته‌های پژوهش‌های کیفی مرتبط را ترکیب کند. در این پژوهش از روش هفت-مرحله‌ای فراترکیب (Sandlowski and Barço 2003) استفاده شده است که مراحل آن در شکل ۱، نشان داده شده است.



شکل ۱. مراحل هفت‌گانه فراترکیب (Sandlowski and Barço 2003)

۴. یافته‌های پژوهش

گام یک: تنظیم سؤال‌های پژوهش

برای تنظیم پرسش پژوهش، نخستین گام برای پژوهشگران، تمرکز بر «چه چیزی» پژوهش است. در پژوهش حاضر، شناسایی و گروه‌بندی منافع حاصل از به‌کارگیری مدیریت کیفیت داده بر حکمرانی داده مورد بررسی قرار می‌گیرد. این پارامتر با جواب دادن به پرسش‌های زیر تنظیم می‌شود:

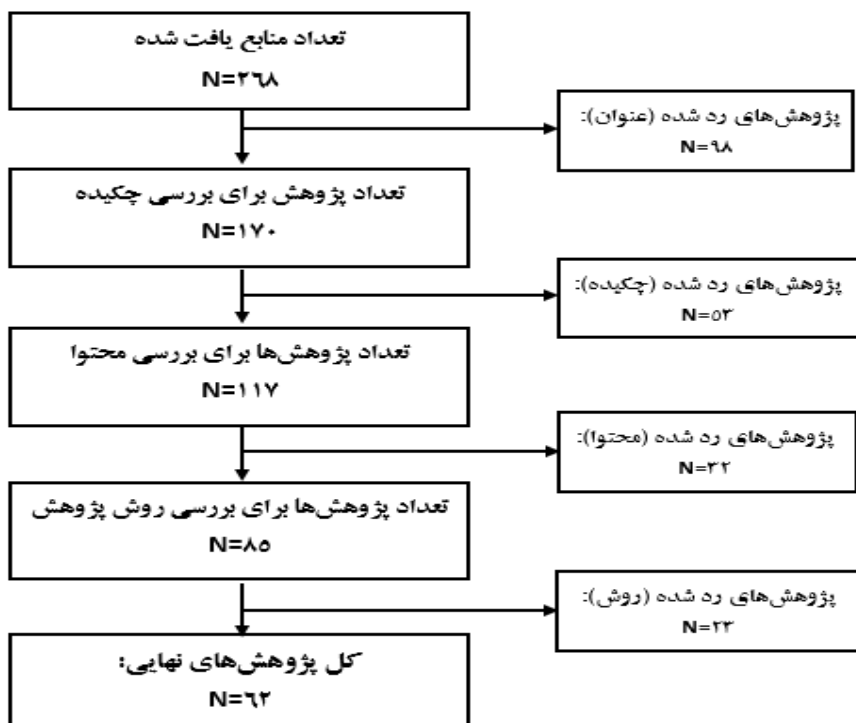
۱. چه چیزی: در این پژوهش پایگاه‌های داده، نشریه‌ها، کنفرانس‌ها و موتورهای جست‌وجوی مختلف مورد بررسی قرار می‌گیرند، و مقاله‌هایی که در مورد حکمرانی داده و مدیریت کیفیت داده باشد، مورد تحلیل قرار می‌گیرند؛
۲. چه وقت: پژوهش‌های مطالعه‌شده از سال‌های ۱۹۹۵ تا ۲۰۲۲ هستند؛
۳. چگونگی: در این پژوهش، روش «تحلیل اسنادی»، یعنی داده‌هایی که به‌صورت ثانویه هستند، مورد استفاده قرار می‌گیرد.

گام دوم: بررسی نظام‌مند متون

در این پژوهش پایگاه‌های داده، نشریه‌ها و موتورهای جست‌وجوی مختلفی بین سال‌های ۱۹۹۵ تا سال ۲۰۲۲ بررسی شده است. واژه‌های کلیدی متنوعی از جمله حکمرانی داده، کیفیت داده، مدیریت کیفیت داده، مدل «داما» و ارزیابی کیفیت داده برای جست‌وجوی مقاله‌های پژوهش مورد استفاده قرار گرفت. در نتیجه جست‌وجو و بررسی پایگاه‌های داده، نشریه‌ها و موتورهای جست‌وجوی مختلف و با استفاده از واژه‌های کلیدی مورد نظر، ۲۶۸ منبع یافت شد.

گام سوم: جست‌وجو و انتخاب پژوهش‌های مناسب

برای جست‌وجو و انتخاب پژوهش‌های مناسب از کلیدواژه‌هایی از جمله داده، مدیریت داده، حکمرانی داده و مدیریت کیفیت داده در پایگاه‌های «ایرانداک»، Science، Direct، Google Scholar، Springer، IEEE و ACM در بین سال‌های (۱۹۹۵-۲۰۲۲) استفاده شد و بخش‌های مختلف پژوهش‌ها از جمله عنوان، چکیده، محتوا، روش و منابع مورد ارزیابی قرار گرفت.



شکل ۲. الگوریتم انتخاب پژوهش‌های نهایی

گام چهارم: استخراج نتایج

پژوهش‌ها بر اساس مرجع مربوط به هر پژوهش، شامل نام و نام‌خانوادگی نویسنده، به همراه سال انتشار مقاله و اجرای هماهنگی بیان‌شده که در هر پژوهش به آن‌ها اشاره شده است، طبقه‌بندی شد.

گام پنجم: تجزیه و تحلیل و تلفیق یافته‌های کیفی

در این پژوهش ابتدا برای تمام عوامل استخراج‌شده از مطالعات پیشین، کدی در نظر گرفته و سپس با در نظر گرفتن مفهوم هر یک از کدها، آن‌ها در یک مفهوم مشابه دسته‌بندی شد. به این ترتیب، مفاهیم پژوهش مشخص شدند. بر اساس تحلیل‌های صورت گرفته به کمک روش تحلیل محتوا روی ۶۲ پژوهش نهایی انتخاب‌شده، در مجموع، تعداد ۱۲ مقوله اصلی و ۴۷ مقوله فرعی برای مدیریت کیفیت داده برای حکمرانی داده کشف و مشخص شدند. در جدول ۱، نتایج حاصل از این قسمت به تفکیک مقوله‌ها نشان داده شده است و در جدول ۲، دسته‌بندی مقوله‌ها صورت گرفته است.

جدول ۱. شناسایی شاخص‌ها یا مقوله‌های فرعی

ردیف	مؤلفه	تعریف	منبع
۱	قابلیت نمایش مقادیر پوچ (Ability to represent null values)	درجه‌ای که یک قالب اجازه می‌دهد تا مقادیر پوچ در یک ویژگی مشخص شود.	Redman & Blanton 1996; Chen 2010; Chen & Tseng 2011; Chung 2006; Goel & Chengalur-Smith 2010; Kim & Han 2011; Lee et al. 2007; Lee & Haider 2011
۲	دسترسی به امنیت (Access security)	میزان دسترسی به مجموعه‌های داده محدود شده است.	Wang 1996; Chen & Tseng 2011; Freitas et al. 2011; Haider & Lee 2012; Lee & Haider 2011; Michnik & Lo 2009; Sung & You 2007
۳	دسترسی (Accessibility)	سهولت بررسی و بازیابی اطلاعات	CDDQ 2019; Eurostat 2015; Wang & Strong 1996; English 1999; ISO 25012 [n. d.]; Chen & Tseng 2011; Chien & Tsaur 2007; Chung 2006; Freitas et al. 2011; Goel, Chengalur-Smith 2010; Haider & Lee 2012; Hazen et al. 2014; Lee & Haider 2011; Lin 2006; Michnik & Lo 2009; Salaün & Flores 2001; Yeganeh, Sadiq & Sharaf 2014
۴	دقت (Accuracy)	میزان نزدیک بودن مقادیر داده به مقادیر واقعی	CDDQ 2019; Eurostat 2015; Wang & Strong 1996; English 1999; DAMA-UK 2013; DAMA 2017; Eurostat 2015; Redman & Blanton 1996; Wagner & et al. 2020; Batini et al. 2009; Bharati & Chaudhury 2004; Bizer, & Cyganiak 2009; Chen 2010; Chen, & Tseng 2011; Freitas et al. 2011; Haider & Lee 2012; Hazen et al. 2014; Jin et al. 2009; Kim & Han 2011; Lee & Haider 2011; Lee, Kim & Kim 2007; Lee & Chung 2009; Li et al. 2005; Lin 2006; Michnik & Lo 2009; Negash, Ryan & Igbaria 2003; Sebastian-Coleman 2012; Yeganeh, Sadiq & Sharaf 2014; Serhani et al. 2016; Taleb & Serhani 2017; Ardagha et al. 2018.
۵	تناسب (Appropriateness)	میزان مناسب بودن قالب برای استفاده	Wang & Strong 1996; Redman & Blanton 1996
۶	قابلیت استفاده (Usability)	میزان مشورت یا بازیابی اطلاعات توسط مصرف‌کنندگان داده یا فرایند	English 1999; ISO 25012; DAMA-UK 2013; Redman & Blanton 1999; Chen & Tseng 2011; Chung 2006; Freitas et al. 2011; Haider & Lee 2012; Jin et al. 2009; Lee et al. 2002; Michnik & Lo 2009;

ردیف	مؤلفه	تعریف	منبع
۷	وضوح (Clarity)	سهولتی که مصرف کنندگان می توانند فراداده را درک کنند.	Eurostat 2014; & Blanton 1996; English 1999
۸	انسجام (Coherence)	میزان ترکیب مجموعه داده ها	Eurostat 2015
۹	قابلیت مقایسه (Comparability)	درجه ای که ارزش داده هایی که دو جمعیت را نشان می دهند، دارای تعریف یکسانی هستند و به یک روش اندازه گیری می شوند.	Eurostat 2014
۱۰	قابلیت مقایسه با گذشت زمان (Comparability over time)	درجه ای که ارزش داده ها در طول زمان تعریف یکسانی دارند و به یک شکل اندازه گیری می شوند.	CDDQ 2019; Eurostat 2015; Wang & Strong 1996; English 1999; DAMA-UK 2013; DAMA 2017; Eurostat 2015; Redman & Blanton 1996; Wagner & et al. 2020
۱۱	کامل بودن (Completeness)	درجه ای که همه ویژگی، سوابق و فایل های مورد نیاز در مجموعه داده وجود دارد.	ISO 25012; Batini et al., 2009; Bharati, Chaudhury, 2004; Bizer, Cyganiak, 2009; Chen, 2010; Chen, Tseng, 2011; Chien, Tsaur, 2007; Chung, 2006; Du, Zhou, 2012; Freitas et al., 2011; Goel, Chengalur-Smith, 2010; Haider, Lee, 2012; Hazen et al., 2014; Jin et al., 2009; Kim, Han, 2011; Kwon, Lee, Shin, 2014; Lee, Chung, 2009; Lee et al., 2002; Lee, Haider, 2011; Michnik, Lo, 2009; Negash et al., 2003; Sebastian-Coleman, 2012; Sung, You, 2007; Yeganeh et al., 2014; Serhani et al., 2016; Taleb & Serhani, 2017.
۱۲	مطابقت با قوانین، مقررات یا استانداردها (Compliance with laws, regulations, or standards)	میزان مطابقت داده ها با قوانین، مقررات یا استانداردها	Wang 1996
۱۳	محرمانه بودن (Confidentiality)	میزان افشای داده ها باید به مصرف کنندگان مجاز داده شود.	DAMA UK 2013; Eurostat 2014; ISO 25012

ردیف	مؤلفه	تعریف	منبع
۱۴	ثبات (Consistency)	میزان رعایت مقادیر داده دو مجموعه از ویژگی‌ها - در یک رکورد، در یک فایل داده، - بین فایل‌های داده، - در یک رکورد در نقاط مختلف زمانی، با یک قاعده مطابقت داشته باشد.	CDDQ 2019; Eurostat 2015; Wang 1996; English 1999; DAMA-UK 2013; DAMA 2017; Eurostat 2015; Redman 1996; Wagner & et al 2020; Batini et al., 2009; Bizer, Cyganiak, 2009; Chen, Tseng, 2011; Chien, Tsaur, 2007; Chung, 2006; Du, Zhou, 2012; Freitas et al., 2011; Haider, Lee, 2012; Hazen et al., 2014; Kwon et al., 2014; Lee, Haider, 2011; Michnik, Lo, 2009; Sebastian-Coleman, 2012; Serhani et al., 2016; Taleb & Serhani, 2017; Ardagna et al., 2018
۱۵	اعتبار (Credibility)	درجه‌ای که ارزش داده‌ها توسط مصرف‌کنندگان به‌عنوان داده‌های واقعی و باورپذیر تلقی شود.	ISO 25012; DAMA 2017; DAMA-UK 2013; English 1999; Wagner & et al. 2020; Bizer & Cyganiak 2009; Chen 2010; Haider & Lee 2012; Hsieh et al. 2010; Lin 2006; Salaün & Flores 2001
۱۶	رایج و شایع (Currency)	میزان به‌روز بودن ارزش داده‌ها	DAMA 2017; CDDQ 2019; Redman & Blanton 1996; ISO 25012 [n. d.]
۱۷	معاذلسازی (Equivalence)	میزان برابر بودن ویژگی‌های ذخیره‌شده در مجموعه داده‌های متعدد از نظر مفهومی	English 1999
۱۸	تقسیم‌بندی صفات یا جداگانه بودن داده (Granularity)	درجه‌ای که یک ویژگی واحد در صفات تقسیم می‌شود.	Redman & Blanton 1996
۱۹	تمامیت (Integrity)	میزان عدم از دست دادن ارزش اطلاعات یا فساد	DAMA 2017; CDDQ 2019; English 1999
۲۰	تفسیرپذیری (Interpretability)	میزان استفاده از داده‌ها به زبان و واحدهای اندازه‌گیری مناسب	Redman & Blanton 1996; Wang & Strong 1996; Chen 2010; Chung 2006; Freitas et al. 2011; Haider & Lee 2012; Lee & Haider 2011; Michnik & Lo 2009
۲۱	تأخیر (Latency)	فاصله زمانی بین زمان ایجاد داده‌ها و زمانی که برای استفاده در دسترس هستند.	DAMA 2017

ردیف	مؤلفه	تعریف	منبع
۲۲	پیوندپذیری (Linkability)	درجه‌ای که می‌توان رکوردهای یک فایل داده را به درستی با رکوردهای یک فایل داده دیگر مرتبط کرد.	Daas 2009; Freitas et al. 2011; Hsieh, Kuo, Yang, Lin, 2010
۲۳	انطباق فراداده (Metadata) (compliance)	میزان مطابقت ارزش داده‌ها با تعریف، مشخصات قالب و دامنه ارزش	Daas 2009; English 1999
۲۴	طبیعی بودن (Naturalness)	میزان همسویی ترکیب مجموعه داده‌ها با اشیای دنیای واقعی که آن‌ها نشان می‌دهند.	Redman & Blanton 1996
۲۵	عینی و هدفمند بودن (Objectivity)	میزان ایجاد مقادیر داده به شیوه‌ای بی‌طرفانه	Wang & Strong 1996; Bizer & Cyganiak 2009; Chen & Tseng 2011; Chung 2006; Freitas et al. 2011; Haider & Lee 2012; Lee et al. 2002; Lee, & Haider 2011; Michnik & Lo 2009
۲۶	قابلیت دستیابی (Obtainability)	میزان دستیابی به داده‌ها	Redman & Blanton 1996
۲۷	قابل قبول بودن (Plausibility)	میزان مطابقت ارزش داده‌ها با دانش دنیای واقعی	Bizer & Cyganiak 2009; Chen & Tseng 2011; Chung 2006; Hazen et al. 2014; Lee & Haider 2011; Li et al. 2005; Michnik & Lo 2009
۲۸	قابلیت حمل (Portability)	میزان نصب، جایگزینی یا انتقال داده‌ها از یک سیستم به سیستم دیگر با حفظ کیفیت موجود	Redman & Blanton 1996; ISO 25012 [n.d.]
۲۹	دقت، درستی (Precision)	میزان دقت ثبت یا طبقه‌بندی ارزش داده‌ها	CDDQ 2019; English 1999; Redman & Blanton 1996; ISO 25012 [n.d.]; Wagner & et al. 2020; Serhani et al. 2016
۳۰	وقت‌شناسی (توجه به جزئیات) (Punctuality)	میزان مناسب بودن فاصله بین زمان واقعی و هدف در دسترس بودن مجموعه داده	Eurostat 2014; Ardagna et al., 2018
۳۱	منطقی بودن (Reasonability)	میزان مطابقت الگوی داده با انتظارات	DAMA 2017
۳۲	قابل بازیابی (Recoverability)	میزان حفظ مجموعه داده‌ها در صورت بروز حادثه	ISO 25012 [n.d.]

ردیف	مؤلفه	تعریف	منبع
۳۳	افزونگی (قابلیت اضافه کردن) (Redundancy)	میزان ذخیره داده‌های منطقی یکسان بیش از یک‌بار	Redman & Blanton 1996; Early 2011
۳۴	یکپارچگی مرجع (Referential integrity)	میزان برابر بودن مقادیر داده کلید اصلی یک فایل داده و مقادیر داده کلید خارجی فایل داده دیگر	Ardagna et al. 2018
۳۵	ارتباط (Relevance)	درجه‌ای که ترکیب مجموعه داده‌ها نیازهای مصرف‌کننده داده را برآورده می‌کند.	Eurostat 2015; Wang & Strong 1996; Redman & Blanton 1996; Bharati & Chaudhury 2004; Bizer & Cyganiak 2009; Chen 2010; Chen & Tseng 2011; Chien & Tsaor 2007; Chung 2006; Goel & Chengalur-Smith 2010; Haider & Lee 2012; Jin et al. 2009; Lederer et al. 2000; Lee & Chung 2009; Lee & Haider 2011; Michnik & Lo 2009; Negash et al. 2003; Salaün & Flores 2001; Song, & Zahedi 2007; Sung & You 2007
۳۶	قابلیت اطمینان (روایی) (Reliability)	نزدیک بودن مقدار داده اولیه به مقدار داده بعدی	Eurostat 2015
۳۷	قابلیت تکرارپذیری (Reproducibility)	میزان بازآفرینی یک مجموعه داده با مقادیر داده یکسان	CDDQ 2019
۳۸	شهرت (اعتبار) (Reputation)	درجه‌ای که داده‌ها از نظر منبع یا محتوای خود مورد اعتماد هستند یا بسیار مورد توجه قرار می‌گیرند.	Wang & Strong 1996; Chen & Tseng 2011; Chung 2006; Freitas et al. 2011; Haider & Lee 2012; Lee & Haider 2011; Michnik & Lo 2009
۳۹	دوره نگهداری (Retention period)	دوره‌ای که مجموعه داده‌ها در دسترس هستند تا زمانی که می‌توانند یا باید حذف شوند.	English 1999

ردیف	مؤلفه	تعریف	منبع
۴۰	به‌موقع بودن (Timeliness)	میزان مناسب بودن فاصله بین زمان ایجاد ارزش واقعی و زمان موجود بودن مجموعه داده مناسب است.	CDDQ 2019; Eurostat 2015; Wang & Strong 1996; English 1999; DAMA-UK 2013; DAMA 2017; Redman & Blanton 1996; Wagner & et al 2020; Batini et al. 2009; Bharati & Chaudhury 2004; Bizer & Cyganiak 2009; Chen 2010; Chen & Tseng 2011; Chien & Tsaur 2007; Chung 2006; Freitas et al. 2011; Haider & Lee 2012; Hazen et al. 2014; Jin et al. 2009; Kim & Han 2011; Lederer et al. 2000; Lee et al. 2007; Lee & Chung 2009; Lee & Haider 2011; Li et al. 2005; Michnik & Lo 2009; Negash et al. 2003; Sebastian-Coleman 2012; Yeganeh et al. 2014; Serhani et al. 2016; Taleb & Serhani 2017
۴۱	قابلیت ردیابی (Traceability)	میزان دسترسی به اصل داده‌ها	Wang & Strong 1996; ISO 25012 [n.d.]
۴۲	منحصربه‌فرد بودن (Uniqueness)	میزان وقوع اجسام (از دنیای واقعی) فقط یک‌بار به‌عنوان یک رکورد در یک فایل داده	DAMA 2017; DAMA-UK 2013; English 1999; Wagner & et al. 2020
۴۳	ارزش (Value)	میزان استفاده از داده‌ها و مزایای آن‌ها	DAMA-UK 2013; Wang & Strong 1996
۴۴	تنوع (Variety)	میزان در دسترس بودن داده‌ها از منابع داده‌های مختلف	Wang & Strong 1996
۴۵	نوسان (Volatility)	میزان تغییر مقادیر داده در طول زمان	DAMA 2017
۴۶	قابل درک بودن (Understandability)	درجه‌ای که با آن اطلاعات به‌سادگی قابل درک است.	Chen, & Tseng 2011; Chung 2006; Haider & Lee 2012; Hsieh et al. 2010; Lee & Haider 2011; Salaün & Flores 2001; Chen 2010; Song & Zahedi 2007; Sung & You 2007)
۴۷	موجز بودن (To be brief)	درجه‌ای که اطلاعات تلخیص شده است.	Bizer & Cyganiak 2009; Chen 2010; Chen & Tseng 2011; Chung 2006; Freitas et al. 2011; Haider & Lee 2012; Michnik & Lo 2009)

گام ششم: کنترل کدهای استخراجی

برای کنترل کدها و مقوله‌های استخراجی از دو کدگذار استفاده شده و شاخص

مورد نظر در این زمینه، شاخص «کاپا» است. برای کنترل مفاهیم استخراجی از مقایسه نظر پژوهشگر با یک خبره استفاده شده است. شاخص «کاپا» بین صفر تا یک نوسان دارد. هرچه مقدار این سنج به عدد یک نزدیک‌تر باشد، نشان می‌دهد که توافق بیشتری بین رتبه‌دهندگان وجود دارد. اما زمانی که مقدار «کاپا» به صفر نزدیک‌تر باشد، در آن صورت توافق کمتر بین دو رتبه‌دهنده وجود دارد (Cohen 1960).

(درصد اتفاق نظری که فقط بر حسب شانس انتظار می‌رود) - (درصد اتفاق نظر مشاهده شده)
کاپا = $\frac{\text{درصد اتفاق نظری که فقط بر حسب شانس انتظار می‌رود} - 100}{\text{درصد}}$

جدول ۲. مقدار اندازه توافق

انحراف استاندارد	مقدار کاپای کوهن	سطح معناداری
۰/۲۵۴	۲/۱۴۹	۰/۰۳۲
۰/۲۹۳		
۴۷	تعداد موارد معتبر	

با توجه به اطلاعات جدول ۲، ملاحظه می‌شود که سطح معناداری شاخص «کاپا» کمتر از سطح خطای ۰/۰۵ قرار دارد. بنابراین، نتیجه گرفته می‌شود که استقلال مفاهیم استخراجی رد می‌شود و به عبارتی، مفاهیم و مقوله‌های استخراجی از پایایی مناسبی برخوردار است.

گام هفتم: ارائه یافته‌ها

بر اساس مطالعه پژوهش‌های پیشین و کدهای استخراج‌شده، مؤلفه‌های مدیریت کیفیت داده برای حکمرانی داده شامل جدول زیر است.

جدول ۳. دسته‌بندی شاخص‌ها در قالب ابعاد مدیریت کیفیت داده برای حکمرانی داده

ابعاد یا مقوله اصلی	مقوله‌های فرعی	تعاریف
ویژگی	کامل بودن	درجه‌ای که همه ویژگی‌های مورد نیاز در مجموعه داده وجود دارد.
	معادل‌سازی	میزان برابر بودن ویژگی‌های ذخیره‌شده در مجموعه داده‌های متعدد از نظر مفهومی
	جداگانه بودن داده	درجه‌ای که یک ویژگی واحد در صفات تقسیم می‌شود.

ابعاد یا مقوله اصلی	مقوله‌های فرعی	تعاریف
داده	دسترسی	سهولت بررسی و بازیابی اطلاعات
	در دسترس بودن	میزان اشتراک یا بازیابی اطلاعات توسط مصرف کنندگان داده یا فرایند
	مطابقت با قوانین و مقررات	میزان مطابقت داده‌ها با قوانین، مقررات یا استانداردها
	محرمانه بودن	میزان افشای داده‌ها باید به مصرف کنندگان مجاز داده محدود شود.
	تفسیرپذیری	میزان استفاده از داده‌ها به زبان و واحدهای اندازه‌گیری مناسب
	تأخیر	فاصله زمانی بین زمان ایجاد داده‌ها و زمانی که برای استفاده در دسترس هستند.
	قابلیت حمل	میزان نصب، جایگزینی یا انتقال داده‌ها از یک سیستم به سیستم دیگر با حفظ کیفیت موجود
	افزونگی	میزان ذخیره داده‌های منطقی یکسان بیش از یک‌بار
	شهرت	درجه‌ای که داده‌ها از نظر منبع یا محتوای خود مورد اعتماد هستند یا بسیار مورد توجه قرار می‌گیرند.
	قابلیت ردیابی	میزان دسترسی به اصل داده‌ها
	قابلیت دستیابی	میزان دستیابی به داده‌ها
	ارزش یا مقدار	میزان استفاده از داده‌ها و مزایای آن‌ها
	تنوع	میزان در دسترس بودن داده‌ها از منابع داده‌های مختلف
فایل داده	کامل بودن	میزان وجود همه فایل‌های داده مورد نیاز
	پیوندپذیری	درجه‌ای که می‌توان رکوردهای یک فایل داده را به‌درستی با رکوردهای یک فایل داده دیگر مرتبط کرد.
	یکپارچگی مرجع	میزان برابر بودن مقادیر داده کلید اصلی یک فایل داده و مقادیر داده کلید خارجی فایل داده دیگر
ارزش داده	دقت	میزان نزدیک بودن مقادیر داده به مقادیر واقعی
	قابلیت مقایسه جمعیت‌ها	درجه‌ای که ارزش داده‌هایی که دو جمعیت را نشان می‌دهند، دارای تعریف یکسانی هستند و به یک روش اندازه‌گیری می‌شوند.
	قابلیت مقایسه با گذشت زمان	درجه‌ای که ارزش داده‌ها در طول زمان تعریف یکسانی دارند و به یک شکل اندازه‌گیری می‌شوند.
	ثبات	مقادیر داده به‌طور مداوم در یک مجموعه داده و بین مجموعه داده‌ها نمایش داده می‌شوند و به‌طور مداوم در مجموعه‌های داده مرتبط هستند.
	قابل قبول بودن	درجه‌ای که ارزش داده‌ها توسط مصرف کنندگان داده‌ها واقعی و باورپذیر تلقی می‌شود.

ابعاد یا مقوله اصلی	مقوله‌های فرعی	تعاریف
	رایج و شایع	میزان به‌روز بودن ارزش داده‌ها
	تمامیت	میزان عدم از دست دادن ارزش اطلاعات
	انطباق فراداده	میزان مطابقت ارزش داده‌ها با تعریف، مشخصات قالب و دامنه ارزش
	قابل قبول بودن	میزان مطابقت ارزش داده‌ها با دانش دنیای واقعی
	دقت، درستی	میزان دقت ثبت یا طبقه‌بندی ارزش داده‌ها
	اعتبار	میزان انطباق ارزش داده‌ها با قوانین
	نوسان	میزان تغییر مقادیر داده در طول زمان
	قابل درک بودن	درجه‌ای که با آن اطلاعات به‌سادگی قابل درک است.
	موجز بودن	درجه‌ای که اطلاعات تلخیص شده است.
ارزش اولیه داده	قابلیت اطمینان	نزدیک بودن مقدار داده اولیه به مقدار داده بعدی
الگوی داده	منطقی بودن	میزان مطابقت الگوی داده با انتظارات
مجموعه داده	دسترسی به امنیت	میزان دسترسی به مجموعه‌های داده محدود شده است.
	قابلیت بازیابی	میزان حفظ مجموعه داده‌ها در صورت بروز حادثه
	قابلیت تکرارپذیری	میزان بازآفرینی یک مجموعه داده با مقادیر داده یکسان
	دوره نگهداری	دوره‌ای که مجموعه داده‌ها در دسترس هستند تا زمانی که می‌توانند یا باید حذف شوند.
دسترسی به مجموعه وقت‌شناسی داده		میزان مناسب بودن فاصله بین زمان واقعی و هدف در دسترس بودن مجموعه داده
	به‌موقع بودن	میزان مناسب بودن فاصله بین زمان ایجاد ارزش واقعی و زمان موجود بودن مجموعه داده مناسب است.
ترکیب مجموعه داده‌ها	انسجام	میزان ترکیب مجموعه داده‌ها
	طبیعی بودن	میزان همسویی ترکیب مجموعه داده‌ها با اشیای دنیای واقعی که آن‌ها نشان می‌دهند.
	ارتباط	درجه‌ای که ترکیب مجموعه داده‌ها نیازهای مصرف‌کننده داده را برآورده می‌کند.
قالب‌بندی	قابلیت نمایش مقادیر پوچ	درجه‌ای که یک قالب اجازه می‌دهد مقادیر پوچ در یک ویژگی مشخص شود.
	تناسب	میزان مناسب بودن قالب برای استفاده
	قابلیت حمل	میزان استفاده از یک قالب در طیف گسترده‌ای از شرایط

ابعاد یا مقوله اصلی	مقوله‌های فرعی	تعاریف
فرا داده	وضوح	سهولتی که مصرف کنندگان داده می‌توانند فراداده یک مجموعه داده را درک کنند.
عینی بودن	کامل و تمامیت منحصر به فرد بودن	میزان توصیف کامل داده‌ها میزان وقوع اجسام (در دنیای واقعی) فقط یک بار به عنوان یک رکورد در یک مجموعه داده

۴. نتیجه‌گیری

هدف از انجام این پژوهش ارائه مدل کیفیت داده برای حکمرانی داده بود. روش پژوهش به صورت فراترکیب به کار گرفته شد. از ۲۶۸ منبع یافت شده، طی مراحل فراترکیب، ۶۲ مقاله با معیارهای مورد پذیرش منطبق بود. در این پژوهش ابتدا برای تمام عوامل استخراج شده از مطالعات پیشین، کدی در نظر گرفته و سپس، با در نظر گرفتن مفهوم هر یک از کدها، آن‌ها در یک مفهوم مشابه دسته‌بندی شد. به این ترتیب، مفاهیم پژوهش مشخص شدند. بر اساس تحلیل‌های صورت گرفته به کمک روش تحلیل محتوا، در مجموع تعداد ۱۲ مقوله اصلی و ۴۷ مقوله فرعی برای مدیریت کیفیت داده برای حکمرانی داده کشف و مشخص شدند.

یکی از مقوله‌های اصلی در زمینه مدیریت کیفیت داده برای حکمرانی داده، «ویژگی داده» است. ویژگی، مشخصه‌ای از نوع موجودیتی است که سازمان مایل است اطلاعات مربوط به آن را در اختیار داشته باشد و شامل کامل بودن (درجه‌ای که همه ویژگی‌های مورد نیاز در مجموعه داده وجود دارد)، معادل‌سازی (میزان برابر بودن ویژگی‌های ذخیره شده در مجموعه داده‌های متعدد از نظر مفهومی) و جداگانه بودن (درجه‌ای که یک ویژگی واحد در صفات مختلف تقسیم می‌شود) داده است. نتایج این بخش از پژوهش را می‌توان با نتایج پژوهش‌های (Hazen & et al. (2014)، (Wagner et al. (2018) و (Ardagna & et al. (2020) همسو و منطبق دانست.

مقوله دیگری که برای مدیریت کیفیت داده جهت حکمرانی داده مورد شناسایی قرار گرفت، «داده» است. داده شامل بازنمایی حقایق، مفاهیم یا دستورالعمل‌ها به صورت رسمی، مناسب برای برقراری ارتباط، تفسیر یا پردازش توسط انسان یا به صورت خودکار است و شامل دسترسی، در دسترس بودن، مطابقت با قوانین و مقررات، محرمانه بودن،

تفسیرپذیری، تأخیر، قابلیت حمل، افزونگی، شهرت، قابلیت ردیابی، قابلیت دستیابی، ارزش یا مقدار، تنوع است. نتایج این بخش از پژوهش را می‌توان با پژوهش‌های Haider & Lee (2012)، (2012)، DAMA (2017)، Ardagna et al. (2018) و Wagner et al. (2020) همسو و منطبق دانست. مقوله دیگری که برای مدیریت کیفیت داده جهت حکمرانی داده مورد شناسایی قرار گرفت، «فایل داده» است. فایل داده به اطلاعات ذخیره‌شده در رایانه به‌عنوان یک واحد با یک نام مشخص را گویند و شامل کامل بودن (میزان وجود همه فایل‌های داده مورد نیاز)، پیوندپذیری (درجه‌ای که می‌توان رکوردهای یک فایل داده را به‌درستی با رکوردهای یک فایل داده دیگر مرتبط کرد) و یکپارچگی مرجع (میزان برابر بودن مقادیر داده کلید اصلی یک فایل داده و مقادیر داده کلید خارجی فایل داده دیگر) است. نتیجه این بخش از پژوهش را می‌توان با شاخص «ذخیره‌سازی داده و عملیات داده» و «انبار داده» (DAMA (2017) همسو و منطبق دانست.

مقوله بعدی که برای مدیریت کیفیت داده جهت حکمرانی داده مورد شناسایی قرار گرفت، «ارزش یا مقدار داده» است. مقدار، یک مورد داده را گویند و نمایانگر ویژگی یک شیء است و شامل دقت، قابلیت مقایسه جمعیت‌ها، قابلیت مقایسه با گذشت زمان، ثبات، قابل قبول بودن، رایج و شایع، تمامیت، انطباق قرارداد، قابل قبول بودن، دقت، درستی، اعتبار، نوسان، قابل درک بودن، و موجز بودن است. در واقع، ترکیب شاخص‌هایی مانند دقت‌پذیری داده، اطمینان داده، صحت داده، و ارزش داده را تعیین می‌کند. نتایج این بخش از پژوهش را می‌توان با پژوهش‌های Haider & Lee (2012) و DAMA (2017) همسو و منطبق دانست.

«ارزش اولیه داده» مقوله دیگری است که برای مدیریت کیفیت داده جهت حکمرانی مورد شناسایی قرار گرفته است. مقدار داده موقتی که با مقدار دقیق‌تری به‌روز می‌شود و به قابل اطمینان بودن داده دلالت دارد. در واقع، هر چقدر مقدار داده اولیه به مقدار داده بعدی نزدیک‌تر باشد، ارزش داده اولیه بیشتر خواهد شد. نتایج این بخش از پژوهش را می‌توان با پژوهش‌های Wang & Strong (1996)، Haider & Lee (2012)، DAMA (2017) و Serhani & et al. (2016) همسو و منطبق دانست.

«الگوی داده» مقوله دیگری است که برای مدیریت کیفیت داده جهت حکمرانی مورد شناسایی قرار گرفته است. مجموعه‌ای از داده‌ها که به روشی قابل تشخیص تکرار می‌شوند و دلالت بر منطقی بودن داده دارد. به عبارتی، منطقی بودن به میزان مطابقت

الگوی داده با انتظارات اشاره دارد.

مقوله دیگری که برای مدیریت کیفیت داده جهت حکمرانی مورد شناسایی قرار گرفته است، «مجموعه داده» است. مجموعه داده، هرگونه جمع‌آوری سازمان‌یافته داده را گویند که شامل دسترسی به امنیت (میزان دسترسی به مجموعه‌های داده محدود شده است)، قابلیت بازیابی (میزان حفظ مجموعه داده‌ها در صورت بروز حادثه)، قابلیت تکرارپذیری (میزان بازآفرینی یک مجموعه داده با مقادیر داده یکسان) و دوره نگهداری (دوره‌ای که مجموعه داده‌ها در دسترس هستند تا زمانی که می‌توانند یا باید حذف شوند) اشاره دارد. نتایج این بخش از پژوهش با شاخص «یکپارچه‌سازی داده‌ها و قابلیت همکاری» DAMA (2017) همسو و منطبق است.

«دسترسی به مجموعه داده»، مقوله دیگری است که برای مدیریت کیفیت داده جهت حکمرانی مورد شناسایی قرار گرفته است. میزان اشتراک یا بازیابی مجموعه داده توسط مصرف‌کنندگان یا فرایندهای داده را گویند و شامل وقت‌شناسی (میزان مناسب بودن فاصله بین زمان واقعی و هدف در دسترس بودن مجموعه داده است) و به موقع بودن (میزان مناسب بودن فاصله بین زمان ایجاد ارزش واقعی و زمان موجود بودن مجموعه داده مناسب است) است.

ترکیب مجموعه داده‌ها: روشی که در آن یک مجموعه داده تشکیل شده است و شامل انسجام، طبیعی بودن و ارتباط است. نتایج این بخش از پژوهش را می‌توان با شاخص «امنیت داده» (2017) DAMA همسو و منطبق دانست.

مقوله دیگری که برای مدیریت کیفیت داده جهت حکمرانی مورد شناسایی قرار گرفته است، «قالب‌بندی یا فرمت» است. قالب‌بندی داده شامل ترکیبی از نوع داده، واحد اندازه‌گیری و مجموعه کاراکتر بوده و شامل قابلیت نمایش مقادیر پوچ (درجه‌ای که یک قالب اجازه می‌دهد تا مقادیر پوچ در یک ویژگی مشخص شود)، تناسب (میزان مناسب بودن قالب برای استفاده) و قابلیت حمل (میزان استفاده از یک قالب در طیف گسترده‌ای از شرایط) است.

مقوله دیگری که برای مدیریت کیفیت داده جهت حکمرانی داده، «فراداده» است. فراداده به داده‌هایی که سایر داده‌ها را تعریف و توصیف می‌کند و شامل وضوح داده (سهولتی که مصرف‌کنندگان داده می‌توانند فراداده یک مجموعه داده را درک کنند)، کامل یا تمامیت (میزان توصیف کامل داده‌ها) است، اشاره دارد. نتایج این بخش از

پژوهش را می‌توان با شاخص «فرا داده» (DAMA (2017) همسو و منطبق دانست. مقوله دیگر و نهایی که برای مدیریت کیفیت داده جهت حکمرانی داده مورد شناسایی قرار گرفت، «عینی بودن داده» است. عینی بودن داده به هر چیزی که قابل درک یا تصور باشد و شامل مقوله منحصر بودن داده است، اشاره دارد و به عبارتی، به میزان وقوع اجسام (در دنیای واقعی) فقط یک بار به عنوان یک رکورد در یک مجموعه داده اشاره دارد. یکی از مدل‌هایی که در زمینه حکمرانی داده وجود دارد، مدل (DAMA (2017) است. «داما» در سال ۲۰۱۷ مدل قبلی خود را بازنگری کرد و شاخص‌های مدل خود را در زمینه حکمرانی داده شامل مدل‌سازی و طراحی داده‌ها، ذخیره‌سازی داده‌ها و عملیات، امنیت داده‌ها، یکپارچه‌سازی داده‌ها و قابلیت همکاری، مدیریت اسناد و محتوا، داده‌های مرجع و اصلی، انبار داده و هوش تجاری، فراداده، کیفیت داده و معماری داده تدوین کرد. شاخص‌های مدل «داما» در ظاهر با شاخص‌های تدوین شده در این پژوهش، تفاوت‌هایی دارد؛ ولی در زیر شاخص‌ها مشابهت‌هایی هم با یکدیگر دارند. بنابراین، در این پژوهش به شاخص‌هایی از جمله خود داده، ارزش داده، قالب‌بندی، عینی بودن داده و ارزش اولیه داده توجه بیشتری شده است و در مدل «داما» به طور دقیق به این شاخص‌ها اشاره نشده است. با توجه به نتایج به دست آمده پیشنهاد می‌شود در پیاده‌سازی مدیریت کیفیت داده جهت حکمرانی داده به شاخص‌هایی از جمله «ویژگی داده، داده، قابل یا قالب داده، ارزش داده، ارزش اولیه داده، الگوی داده، مجموعه داده، دسترسی به داده، ترکیب داده، قالب‌بندی داده، فراداده و عینی بودن داده» توجه شود.

فهرست منابع

- ارشادی، محمدجواد، جلال‌الدین نصیری، و فرهاد شیرانی. ۱۳۹۶. طراحی مدل کیفیت فراداده در سامانه ثبت پایان‌نامه، رساله‌های دانش‌آموختگان داخل و خارج کشور. طرح پژوهشی. وزارت علوم، تحقیقات و فناوری، پژوهشگاه علوم و فناوری اطلاعات ایران.
- خلیلی جعفرآباد، احمد. ۱۳۹۶. بررسی تغییرات حوزه کیفیت داده با استفاده از تحلیل کلمات کلیدی. دو فصلنامه علمی-پژوهشی مدیریت اطلاعات ۲ (۳): ۱۲۱-۱۳۸.
- درزی خلردی، صغری، سید علی اصغر رضوی، و میترا قیاسی. ۱۴۰۰. بلوغ حکمرانی داده در کتابخانه‌های دانشگاهی استان مازندران. علوم و فنون مدیریت اطلاعات ۷ (۳): ۳۵۳-۳۶۸.

- درودی، فریبرز، فریبا جلیلی، و سید علی اکبر فامیل روحانی. ۱۴۰۰. بررسی نگرش کتابداران در پیاده‌سازی و اجرای راهبردهای حاکمیت داده در برنامه‌ریزی سازماندهی منابع اطلاعاتی در کتابخانه‌ها. *پژوهشنامه کتابداری و اطلاع‌رسانی* ۱۱ (۲): ۴۴-۶۴.
- رحیمی، علی‌رضا ۱۳۹۵. بررسی تحولات پژوهش‌های حوزه ارزیابی کیفیت داده‌ها و اطلاعات در نظام‌های اطلاعاتی از سال ۲۰۰۰ تا نیمه نخست سال ۲۰۱۵. *پژوهشنامه پردازش و مدیریت اطلاعات*. <http://Jipm.irandoc.ac.ir> (دسترسی در ۱۳۹۶/۴/۸-۸).
- سلطانی‌نژاد، احمد، و سهیل گودرزی. ۱۳۹۶. فناوری اطلاعات و تحول در مفهوم حکمرانی خوب. *فصلنامه سیاست، مجله دانشکده حقوق و علوم سیاسی* ۴۷ (۱): ۷۹-۹۷.
- عبدالوند، ندا، و آتنا بویه‌رژ. ۱۳۹۵. رویکردی سیستماتیک به چالش کیفیت داده در راهبرد مشتری‌محور در صنعت بانکداری. *فصلنامه علمی- پژوهشی تحقیقات بازاریابی نوین* ۱ (۶): ۱۷۷-۱۹۶.
- غفاری، مسعود، عادل آذر، و اشکان شباک. ۱۳۹۷. نگاشت علی مدیریت محصول آماری با رویکرد کیفیت داده. *پژوهشنامه پردازش و مدیریت اطلاعات* ۳ (۳۳): ۱۰۴۱-۱۰۶۴.
- کارگر، محمدجواد، و میلاد محمدی. ۱۳۹۶. مدلی به‌منظور کیفیت بر روی کلان‌داده. *سومین کنفرانس ملی محاسبات توزیعی و پردازش داده‌های بزرگ*. تهران.
- محبوب، عباس، و حمیدرضا سمیه‌ساز. ۱۳۹۴. کیفیت داده‌ها پیش‌نیاز مدیریت منابع سازمان. *فصلنامه مدیریت استاندارد و کیفیت* ۵ (۱۷): ۵۳-۶۲.

References

- Alhassan, I., D. Sammon, & M. Daly. 2019. Critical success factors for data governance: a theory building approach. *Information Systems Management* 36 (2): 98-110
- Al-Sharida, R. A. 2015. Data governance in electronic health records: A systematic review. Doctoral dissertation. University of Arkansas at Little Rock.
- Ardagna, D., C. Capiello, W. Samá, & M. Vitali. 2018. Context-aware data quality assessment for big data. *Future Generation Computer Systems* 89: 548-562.
- Batini, C., C. Capiello, C. Francalanci, A. & Maurino. 2009. Methodologies for data quality assessment and improvement. *ACM computing surveys (CSUR)* 41 (3): 16.
- Bharati, P., & A. Chaudhury. 2004. An empirical investigation of decision-making satisfaction in web-based decision support systems. *Decision support systems* 37 (2): 187-197.
- Bizer, C., & R. Cyganiak. 2009. Quality-driven information filtering using the WIQA policy framework. *Journal of Web Semantics* 7 (1): 1-10.
- CDDQ. 2019. List of Conformed Dimensions of Data Quality. Retrieved from <https://dimensionsofdataquality.com/alldimensions> (accessed July 12, 2019)
- Chen, C. C., & Y. D. Tseng. 2011. Quality evaluation of product reviews using an information quality framework. *Decision Support Systems* 50 (4): 755-768.
- Chen, C. W. 2010. Impact of quality antecedents on taxpayer satisfaction with online tax-filing systems—An empirical study. *Information & Management* 47 (5-6): 308-315.

- Chengalur-Smith, I. N., D. P. Ballou, & H. L. Pazer. 1999. The impact of data quality information on decision making: an exploratory analysis. *IEEE Transactions on Knowledge and Data Engineering* 11 (6): 853-864.
- Chien, S. W., & S. M. Tsaur. 2007. Investigating the success of ERP systems: Case studies in three Taiwanese high-tech industries. *Computers in industry* 58 (8-9): 783-793.
- Chung, W. 2006. Studying information seeking on the non-English Web: An experiment on a Spanish business Web portal. *International Journal of Human-Computer Studies* 64 (9): 811-829.
- Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement* 20 (1): 37-46.
- Dai, T., H. Hu, Y. Wan, Q. Chen, & Y. Wang. 2015. A data quality management and control framework and model for health decision support. In 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) (pp. 1792-1796). IEEE. Canada.
- DAMA 2017. DAMA-DMBOK. *Data Management Body of Knowledge*. 2nd Edition. Denver, Colorado: Technics Publications LLC.
- DAMA. 2008. The DAMA dictionary of data management. Denver, Colorado: Technics publications.
- DAMA, U. 2013. The Six Primary Dimensions for Data Quality Assessment-Defining Data Quality Dimensions. Bristol: np URL: https://www.whitepapers.em360tech.com/wpcontent/files_mf/1407250286DAMAUKDQDimensionsWhitePaperR37.pdf, (accessed ?)
- Daas, P.J.H., S.J.L.Ossen, & J. Arends-Tóth. 2009. Framework of Quality Assurance for Administrative Data Sources. Paper for the 57th session of the International Statistical Institute, 16- 22 Aug., Durban, South Africa..
- Earley, S. 2011. *The DAMA Dictionary of Data Management*. 2nd ed. NJ: Technics Publications LLC.
- Egan, P. 2011. Effecting Data Quality through Data Governance: a Case Study in the Financial Services Industry, All Regis University Theses. 465. <https://epublications.regis.edu/theses/465>
- English, L. P. 1999. *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits*. Hoboken, NJ: Wiley.
- Eurostat. 2014. Expenditure on education as % of GDP or public expenditure, available at http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=educ_figdp&lang=en (accessed ?)
- Eurostat. 2015. ESS Handbook for Quality Reports. Brussels, Belgium: Eurostat.
- Even, A., G. Shankaranarayanan, & P. D. Berger. 2010. Evaluating a model for cost-effective data quality management in a real-world CRM setting. *Decision Support Systems* 50 (1): 152-163.
- Freitas, A., T. Knap, S. O'Riain, & E. Curry. 2011. W3P: Building an OPM based provenance model for the Web. *Future Generation Computer Systems* 27 (6): 766-774.
- Goel, S., & I. N. Chengalur-Smith. 2010. Metrics for characterizing the form of security policies. *The Journal of Strategic Information Systems* 19 (4): 281-295.
- Haider, A., & S. H. Lee. 2012. Using six sigma for continuous improvement of asset management information quality. CONF-IRM 2012 Proceedings. 59. <https://aisel.aisnet.org/confirm2012/59> (accessed ?)
- Hazen, B. T., C. A. Boone, J. D. Ezell, & L. A. Jones-Farmer. 2014. Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics* 154: 72-80.
- Hikmawati, S., P. I. Santosa, & I. Hidayah. 2021. Improving Data Quality and Data Governance Using Master Data Management: A Review. *IJITEE (International Journal of Information Technology and Electrical Engineering)* 5 (3): 90-95.

- ISO 25012. (n.d.). The Data Quality model defined in the standard ISO/ IEC 25012 .Retrieved from <https://iso25000.com/index.php/en/iso-25000-standards/iso-25012> (accessed May 9, 2023)
- Jin, X. L., C. M. Cheung, M. K. Lee, & H. P. Chen. 2009. How to keep members using the information in a computer-supported social network. *Computers in Human Behavior* 25 (5): 1172-1181.
- Kim, B., & I. Han. 2011. The role of utilitarian and hedonic values and their antecedents in a mobile data service environment. *Expert Systems with Applications* 38 (3): 2311-2318.
- Koltay, T. 2016. Data governance, data literacy and the management of data quality. *IFLA journal* 42 (4): 303-312.
- Korhonen, J. J., I. Melleri, K. Hiekkänen, & M. Helenius. 2014. Designing data governance structure: An organizational perspective. *GSTF Journal on Computing (JoC)* 2 (4): 1-14.
- Lee, H., J. Kim, & J. Kim. 2007. Determinants of success for application service provider: An empirical test in small businesses. *International journal of human-computer studies* 65 (9): 796-815.
- Lee, S. H., & A. Haider. 2011. A Framework for Information Quality Assessment Using Six Sigma Approach. Communications of the IBIMA. <http://www.ibimapublishing.com/journals/CIBIMA/cibima.html> Vol. 2011 (2011), Article ID 927907, 1-11. DOI: 10.5171/2011.927907
- Li, S., S. S. Rao, T. S. Ragu-Nathan, & B. Ragu-Nathan. 2005. Development and validation of a measurement instrument for studying supply chain management practices. *Journal of operations management* 23 (6): 618-641.
- Liaw, S. T., A. Rahimi, P. Ray, J. Taggart, S. Dennis, S. de Lusignan, ... & A. Talaie-Khoei. 2013. Towards an ontology for data quality in integrated chronic disease management: a realist review of the literature. *International journal of medical informatics* 82 (1): 10-24.
- Lin, A. 2006. The acceptance and use of a business-to-business information system. *International Journal of Information Management* 26 (5): 386-400.
- Loshin, D. 2010. *Building a Data Quality Scorecard for Operational Data Governance*. Cary, NC: DataFlux Corporation.
- Mavodza, J. 2022. Research Data Management: A review of UAE academic library experience. *Open Information Science* 6 (1): 16-27.
- Michnik, J., & M. C. Lo. 2009. The assessment of the information quality with the aid of multiple criteria analysis. *European Journal of Operational Research* 195 (3): 850-856.
- Michnik, J., & M. C. Lo. 2009. The assessment of the information quality with the aid of multiple criteria analysis. *European Journal of Operational Research* 195 (3): 850-856.
- Negash, S., T. Ryan, & M. Igbaria. 2003. Quality and effectiveness in web-based customer support systems. *Information & management* 40 (8): 757-768.
- Ofner, M., B. Otto, & H. Österle. 2013. A maturity model for enterprise data quality management. *Enterprise Modelling and Information Systems Architectures-An International Journal*: Vol. 8, Nr. 2.4-21 : (2) 8 .
- Pipino, L. L., Y. W. Lee, & R. Y. Wang. 2002. Data quality assessment. *Communications of the Association for Computing Machinery (ACM)* 45 (4): 211-218.
- Redman, T. C., & A. Blanton. 1996. *Data quality for the information age*. Norwood: Artech House.
- Salaün, Y., & K. Flores. 2001. Information quality: Meeting the needs of the consumer. *International Journal of Information Management* 21 (1): 21-37.
- Sebastian-Coleman, L. 2012. Measuring data quality for ongoing improvement: a data quality assessment framework. Newnes. ???
- Serhani, M. A., H. T. El Kassabi, I. Taleb, & A. Nujum. 2016. An hybrid approach to quality evaluation across big data value chain. In 2016 IEEE International Congress on Big Data (Big Data Congress) (pp. 418-425). IEEE.

- Sung, T. J., & M. You. 2007. A method for establishing an online design audit platform. *Design Studies* 28 (2): 195-211.
- Taleb, I., & M. A. Serhani. 2017. Big Data pre-processing: closing the data quality enforcement loop. In 2017 IEEE International Congress on Big Data (BigData Congress) (pp. 498-501). IEEE. Honolulu, HI, USA.
- Taleb, I., H. T. El Kassabi, M. A. Serhani, R. Dssouli, & C. Bouhaddiou. 2016. Big data quality: A quality dimensions evaluation. In 2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ ATC/ ScalCom/ CBDCom/ IoP/ SmartWorld) (pp. 759-765). IEEE. United Arab Emirates University.
- Wagner, T., N. R. Lottig, M. L. Bartley, E. M. Hanks, E. M. Schliep, E. M., N. B. Wickle, ... & J. Zhou. 2020. Increasing accuracy of lake nutrient predictions in thousands of lakes by leveraging water clarity data. *Limnology and Oceanography Letters* 5 (2): 228-235.
- Wang, R. Y. 1998. A product perspective on total data quality management. *Communications of the ACM* 41 (2): 58-66.
- Wang, R. Y., & D. M. Strong. 1996. Beyond accuracy: What data quality means to data consumers. *Journal of management information systems* 12 (4): 5-33.
- Wende, K. 2007. A model for data governance-Organising accountabilities for data quality management. *ACIS 2007 Proceedings*, 80.
- Were, V., & C. Moturi. 2017. Toward a data governance model for the Kenya health professional regulatory authorities. *The TQM Journal* 29 (4): 579-589.
- Yeganeh, N. K., S. Sadiq, & M. A. Sharaf. 2014. A framework for data quality aware query systems. *Information Systems* 46: 24-44.

سحر بنیادی

دانشجوی دکتری علم اطلاعات و دانش‌شناسی گرایش ذخیره و بازیابی اطلاعات واحد علوم و تحقیقات دانشگاه آزاد اسلامی است. ایشان هم‌اکنون مدیر کتابخانه دانشگاه پیام نور اردبیل است. مدیریت داده، مدیریت کیفیت داده، ذخیره و بازیابی اطلاعات از جمله علایق پژوهشی ایشان است.



نجلا حویری

دارای مدرک دکتری علم اطلاعات و دانش‌شناسی است. ایشان هم‌اکنون استادتمام گروه علم اطلاعات و دانش‌شناسی واحد علوم و تحقیقات دانشگاه آزاد اسلامی است. رفتار اطلاع‌یابی، سازماندهی اطلاعات، مدیریت اطلاعات و پایگاه‌های اطلاعاتی از جمله علایق پژوهشی ایشان است.



سید مهدی طاهری

دارای مدرک دکتری در رشته علم اطلاعات و دانش‌شناسی است. ایشان هم‌اکنون رئیس کتابخانه مرکزی و مرکز اسناد و استادیار دانشگاه علامه طباطبائی است. سازماندهی اطلاعات، مدیریت فراداده و کتابخانه دیجیتال از جمله علایق پژوهشی ایشان است.

**رؤیا پورنقی**

دارای مدرک تحصیلی دکتری در رشته علم اطلاعات و دانش‌شناسی است. ایشان هم‌اکنون استادیار پژوهشگاه علوم و فناوری اطلاعات ایران است. سیستم‌های اطلاعات مکانی، دروازه‌های اطلاعاتی، علم‌سنجی، حق مؤلف، سواد اطلاعاتی و اثر پژوهش از علایق پژوهشی ایشان است.

