

Recognition and Systemic Analysis of Data Quality Methodologies and Proposing a Comprehensive Framework Using the Meta-Synthesis Method

Babak Sohrabi*

PhD in Management Science; Professor; University of Tehran;
Tehran, Iran Email: bsohrabi@ut.ac.ir

Hamid Reza Yazdani

PhD in Human Resources Management; Assistant
Professor; Farabi Campus; University of Tehran; Tehran, Iran;
Email: hryazdani@ut.ac.ir

Mohammad Javad Ershadi

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

Soudeh Dorvash

PhD Candidate in Information Technology Management; University
of Tehran; Tehran, Iran Email: soudeh.dorvash@ut.ac.ir

Received: 18, May 2020 Accepted: 15, Sep. 2020

Abstract: Despite abundant research on data quality, no research has so far been conducted which can provide a comprehensive view of data quality methodologies. In the present study, 3909 articles and related researches in the period before 2020 were selected from Web of Science (WOS) and Scopus citation indexes, from among which 27 articles were finally evaluated in line with the research goals, using meta-synthesis method and inclusion criteria. In this regard, while applying the system view and using the open coding method, the related codes to the three main categories of the systemic approach (input, process and output) were extracted. The similar concepts were categorized in sub-codes and then the sub-codes in main codes. The main inputs included the context and status of organization, data and information resources, and data quality dimensions. Steps of data quality methodologies were also classified into three main stages: state reconstruction, measurement/evaluation, and improvement. In addition, data quality outputs fell into six general categories: list of activities and the related techniques

**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. 36 | No. 3 | pp. 737-766

Spring 2021



* Corresponding Author

for data quality improvement, controlled or redesigned processes, measured or improved flows and databases, data quality status results, revised data quality policies or rules, and costs and benefits. The results of this study can provide an appropriate instrument for identifying the existing data quality methodologies as well as evaluating the strengths and weaknesses of data quality methodologies.

Keywords: Data Quality, Data Quality Methodology, Systemic Approach, Meta-Synthesis

شناخت و تحلیل سیستمی متدولوژی‌های کیفیت داده و ارائه یک چارچوب جامع (با استفاده از روش فراترکیب)

بابک سهرابی

دکتری علوم مدیریت؛ استاد؛

دانشگاه تهران؛ تهران، ایران؛

پدیده‌آور رابط bsohrabi@ut.ac.ir

حمیدرضا یزدانی

دکتری مدیریت منابع انسانی؛ استادیار؛ پردیس فارابی

دانشگاه تهران؛ تهران، ایران hryazdani@ut.ac.ir

محمدجواد ارشادی

دکتری مهندسی صنایع؛ استادیار؛ پژوهشگاه علوم و

فناوری اطلاعات ایران (ایرانداک)؛ تهران، ایران؛

ershadi@irandoc.ac.ir

سوده دوروش

دانشجوی دکتری مدیریت فناوری اطلاعات؛

دانشگاه تهران؛ تهران، ایران؛

soudeh.dorvash@ut.ac.ir



مقاله برای اصلاح به مدت ۳۸ روز نزد پدیدآوران بوده است.

پذیرش: ۱۳۹۹/۰۶/۲۵

دریافت: ۱۳۹۹/۰۲/۲۹

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

شاپا (جایی) ۸۲۲۳-۲۵۰۱

شاپا (الکترونیکی) ۸۲۳۱-۲۵۰۱

نمایه در SCOPUS، LISTA، ISC و

jipm.irandoc.ac.ir

دوره ۳۶ | شماره ۳ | صص ۷۳۷-۷۶۶

بهار ۱۴۰۰



چکیده: به‌رغم وجود تحقیقات فراوان در حوزه کیفیت داده، تاکنون پژوهشی که بتواند دیدگاه جامعی نسبت به متدولوژی‌های کیفیت داده فراهم آورد، انجام نشده است. در این مطالعه ۳۹۰۹ مقاله و پژوهش مرتبط در بازه زمانی قبل از ۲۰۲۰ از نمایه‌های استنادی «وب‌آوساینس» و «اسکوپوس» انتخاب شد که با استفاده از روش فراترکیب و معیارهای ورود در نهایت، ۲۷ مقاله در راستای هدف پژوهش مورد ارزیابی قرار گرفتند. در این راستا، ضمن به‌کارگیری دیدگاه سیستمی و با استفاده از روش کدگذاری باز، کدهای مربوط به سه مقوله اصلی رویکرد سیستمی شامل ورودی، فرایند و خروجی استخراج گردید و مفاهیم مشابه در کدهای فرعی و در ادامه، کدهای فرعی در کدهای اصلی دسته‌بندی شدند. ورودی‌های اصلی شامل زمینه و وضعیت سازمان، داده‌ها و منابع اطلاعاتی و ابعاد کیفیت داده هستند. همچنین، گام‌های متدولوژی‌های کیفیت داده در سه مرحله اصلی بازسازی وضعیت، ارزیابی/اندازه‌گیری و ارتقا طبقه‌بندی شده‌اند. افزون بر این، خروجی‌های کیفیت داده در شش دسته کلی شامل فهرست فعالیت‌ها و

تکنیک‌های مرتبط مشخص شده برای ارتقای کیفیت داده‌ها، فرایندهای کنترل شده یا بازطراحی شده، جریان‌ها و پایگاه‌های داده اندازه‌گیری یا ارتقا داده شده، نتایج ارائه شده از وضعیت کیفیت داده، سیاست‌ها یا قوانین کیفیت داده تصحیح شده، و هزینه‌ها و منفعت‌ها طبقه‌بندی شده‌اند. نتایج حاصل از این پژوهش می‌تواند ابزار مناسبی جهت شناخت متدولوژی‌های کیفیت داده موجود و همچنین ارزیابی نقاط ضعف و قوت متدولوژی‌های کیفیت داده باشد.

کلیدواژه‌ها: کیفیت داده، متدولوژی کیفیت داده، رویکرد سیستمی، فراترکیب

۱. مقدمه

امروزه، داده‌ها به عنوان یک دارایی ارزشمند در سازمان به شمار می‌روند. در سال‌های اخیر حجم داده استفاده شده در سازمان‌ها به شدت افزایش پیدا کرده و این نقش بسیار مهمی در عملیات کسب و کار دارد (Alhassan, Sammon & Daly 2019). سازمان‌ها به دو روش کلی متکی به استفاده از داده‌ها هستند: اول، اجرای فرایندهای استاندارد کسب و کار، و دوم، پشتیبانی از فعالیت‌های عملیاتی. تحلیلگران کسب و کار، داده‌های به دست آمده از نتیجه عملیات روزانه را از طریق گزارش‌ها و موتورهای تحلیلگر به عنوان راهی برای شناسایی فرصت‌های جدید برای بهره‌وری یا رشد استفاده می‌کنند. به عبارت دیگر، از داده‌ها برای اجرا و بهبود راه‌هایی که سازمان‌ها به اهداف کسب و کار خود می‌رسند، استفاده می‌شود (Loshin 2010). از طرف دیگر، چنانچه داده‌ها از کیفیت مناسبی برخوردار نباشند، نتایج و پیامدهای نامطلوبی برای سازمان ایجاد خواهد شد. کیفیت داده غالباً با عنوان «تناسب برای استفاده» تعریف می‌شود (Tayi and Ballou 1998). به عبارت بهتر، لازم است داده‌های جمع‌آوری شده از کیفیت بالایی برخوردار باشند، در غیر این صورت، تجزیه و تحلیل داده‌ها، برنامه‌های کاربردی و فرایندهای کسب و کار غیر قابل اعتماد خواهند بود. بنابراین، داده‌هایی با کیفیت پایین منجر به ریسک‌هایی در دستیابی به موفقیت در اهدافی می‌شود که فعالیت‌های روزانه سازمان را هدایت می‌کنند (Kilkenny and Robinson 2018). تعاریف زیادی در مورد کیفیت داده‌ها وجود دارد، اما داده‌ها معمولاً زمانی با کیفیت بالا در نظر گرفته می‌شوند که «برای استفاده‌های مورد نظر در عملیات، تصمیم‌گیری و برنامه‌ریزی مناسب باشند» (Redman 2013; Fadahunsi et al. 2019). کیفیت اطلاعات، هم در سازمان‌ها و هم در روابط بین سازمانی نقش اساسی دارد. در سازمان‌ها، کیفیت اطلاعات

خطاها را کاهش می‌دهد، باعث افزایش کارایی فرایند و بهبود اثربخشی تصمیم‌گیری می‌شود و در روابط بین سازمانی، یک مؤلفه مهم در قابل اعتماد به‌شمار رفتن هر سازمان محسوب می‌شود (Cappiello, Francalanci and Pernici 2005). اثرات کیفیت پایین داده‌ها در صورت تأثیر بر نارضایتی مشتری یا ایجاد مانع در تصمیم‌گیری‌های مدیریتی مناسب می‌تواند بسیار مهم و تعیین‌کننده باشد؛ به‌طوری که این نکته در ادبیات موضوع کیفیت داده با تأکید بیان شده است (Redman and Blanton 1996; English 1999). یک برنامه جامع مدیریت کیفیت داده تنها با ارزیابی مداوم، نظارت و بهبود داده‌ها قابل دستیابی است (Cappiello, Francalanci and Pernici 2005). متدولوژی‌های کیفیت داده با فراهم آوردن یک فرایند منطقی برای ارزیابی و بهبود کیفیت داده‌های سازمان‌ها از طریق مرحله‌های مشخص و نقاط تصمیم‌گیری، سازمان را در این امر مهم یاری می‌کنند (Batini and Scannapieco 2016).

از طرف دیگر، امروزه با افزایش آگاهی سازمان‌ها نسبت به اهمیت کیفیت داده و همچنین، اقبال روزافزون سازمان‌های ایرانی به مقوله مدیریت کیفیت داده و در نتیجه، لزوم به‌کارگیری متدولوژی‌های کیفیت داده نیاز به ابزاری که بتواند شناختی جامع و فراگیر نسبت به متدولوژی‌های کیفیت داده فراهم کند، بیش از پیش رخ می‌نماید. عدم شناخت کافی نسبت به ورودی‌های متدولوژی‌ها منجر به ارزیابی نادرست از مرزهای یک برنامه مدیریت کیفیت داده و در نتیجه، افزایش احتمال شکست یک پروژه خواهد شد. ممکن است سازمان‌ها در حال استفاده از متدولوژی متداول و مرسوم باشند که احتمالاً با نیازمندی‌های متفاوت و متغیر آن‌ها متناسب نباشد و یا ممکن است شناختی نسبت به سایر متدولوژی‌ها و خروجی‌هایی که می‌توانند فراهم کنند، نداشته باشند (Woodall, Borek and Parlikad 2013). بنابراین، نبود پژوهشی که بتواند دیدگاهی جامع نسبت به متدولوژی‌های موجود فراهم آورد، منجر به عدم رفع نیازمندی‌های سازمان‌ها در مدیریت کیفیت داده شده و در نتیجه، به ایجاد مانعی در دستیابی به اهداف سازمانی و یا اتلاف هزینه و منابع سازمان منتهی شده است.

تاکنون، متدولوژی‌های بسیاری برای توانمندسازی سازمان‌ها در ارزیابی و ارتقای سطح کیفیت داده‌های آن‌ها پیشنهاد شده است. در این پژوهش از رویکرد سیستمی برای تحلیل و شناخت متدولوژی‌های کیفیت داده استفاده شده است. متفکران سیستمی هدف از طرح این نوع تفکر را کمک به درک بهتر رفتار پدیده‌ها و روند تغییر رویدادهایی

می‌دانند که در طبیعت تکوین یافته و می‌یابند. تفکر سیستمی با توجه به مدل ارائه‌شده توسط «پتر سنگه»^۱، یکی از فرمان‌های یادگیری سازمانی است. تفکر سیستمی پرهیز از تفرق‌گرایی سیستمی و حرکت به سمت تشکل‌گرایی سیستمی هدفمند است (محمدخانی، رشادت‌جو و روحانی‌پور ۱۳۹۱). از طرف دیگر، رویکرد سیستمی می‌تواند در ایجاد یک الگوی جامع و منسجم نقش‌آفرینی کند (محمدی و حاجی‌پور ۱۳۹۷). در این راستا، هدف پژوهش حاضر ارائه یک دیدگاه جامع در شناخت متدولوژی‌های کیفیت داده با استفاده از رویکرد سیستمی است.

سؤالات تحقیق

۱. ورودی‌های متدولوژی‌های کیفیت داده کدام‌اند؟
۲. فرایند (فعالیت‌های) پیشنهادشده در متدولوژی‌های کیفیت داده کدام‌اند؟
۳. خروجی‌های متدولوژی‌های کیفیت داده کدام‌اند؟

مبانی نظری

تعاریف متعددی برای کیفیت داده وجود دارد، اما عموماً داده‌ای با کیفیت بالا در نظر گرفته می‌شود که برای مصارف مورد نظر در عملیات، تصمیم‌گیری و برنامه‌ریزی مناسب باشد (Redman 2013). در تعریفی دیگر، کیفیت اطلاعات به‌عنوان «میزانی که اطلاعات و داده‌ها می‌توانند منبع قابل اعتمادی برای یک و یا همه مصارف مورد نیاز باشند»، معرفی گردیده است. این به معنای داشتن مجموعه‌ای از اطلاعات صحیح، در زمان مناسب، در مکان مناسب، برای استفاده افراد مناسب در تصمیم‌گیری، اداره کسب‌وکار، خدمت به مشتریان و دستیابی به اهداف شرکت است (McGilvray 2008).

بر اساس تعریف لغت‌نامه «آکسفورد»، متدولوژی به‌عنوان مجموعه‌ای از روش‌ها و اصول استفاده‌شده برای انجام یک فعالیت خاص در نظر گرفته می‌شود (Oxford 2020). طبق تعریف ارائه‌شده توسط «ایرنی و رز»، متدولوژی عبارت است از تحلیل نظام‌مند و نظری روش‌هایی که در یک زمینه مطالعاتی به کار می‌روند و شامل تحلیل نظری روش‌ها و اصول مرتبط با شاخه‌ای از دانش است. به‌طور معمول، متدولوژی مفاهیمی مانند پارادایم،

1. Peter Senge

مدل نظری، مراحل و فن‌های کمی یا کیفی را دربرمی‌گیرد (Irny and Rose 2005). یک متدولوژی برای ارائه مستقیم راه‌حل‌ها تعریف نشده است؛ بلکه، زیربنای نظری را برای درک این که کدام روش، مجموعه روش‌ها یا بهترین روش‌ها را می‌توان در یک مورد خاص (مثلاً برای محاسبه یک نتیجه خاص) به کار برد، فراهم می‌کند (Shirish 2013). متدولوژی کیفیت داده به عنوان مجموعه‌ای از دستورالعمل‌ها و تکنیک‌ها تعریف می‌شود که از اطلاعات ورودی مربوط به یک واقعیت خاص و مورد علاقه شروع می‌شود. همچنین، یک فرایند منطقی را به منظور ارزیابی و بهبود کیفیت داده‌های سازمان، از طریق مراحل مشخص و نقاط تصمیم‌گیری تعیین می‌کند (Batini and Scannapieco 2016). در اغلب موارد، توالی فعالیت‌های یک متدولوژی کیفیت داده‌ها از سه مرحله تشکیل شده است:

۱. بازسازی وضعیت که با هدف جمع‌آوری اطلاعات زمینه‌ای در مورد فرایندهای سازمان و خدمات، جمع‌آوری داده‌ها و رویه‌های مدیریتی مرتبط، موضوعات مربوط به کیفیت و هزینه‌های مربوط انجام می‌شود.
۲. ارزیابی^۱ / اندازه‌گیری^۲ که کیفیت مجموعه‌های داده را با ابعاد کیفیت مربوط اندازه‌گیری می‌کند. اصطلاح اندازه‌گیری برای پرداختن به مسئله اندازه‌گیری میزان مجموعه‌ای از ابعاد کیفیت داده استفاده می‌شود. اصطلاح ارزیابی وقتی استفاده می‌شود که چنین اندازه‌گیری‌هایی با مقادیر مرجع مقایسه شده و بتوانند کیفیت را مشخص کنند.
۳. ارتقا مربوط به انتخاب گام‌ها، استراتژی‌ها و تکنیک‌ها برای دستیابی به اهداف جدید کیفیت داده است (Batini et al. 2009; Batini and Scannapieco 2016).

بررسی ادبیات موضوع حاکی از آن است که تاکنون متدولوژی‌های کیفیت داده بسیاری پیشنهاد شده‌اند. برخی از متدولوژی‌های کیفیت داده نظیر متدولوژی ارائه شده توسط Lee et al. (2002) صرفاً تا مرحله اندازه‌گیری / ارزیابی را پوشش می‌دهند و در عین حال، برخی دیگر مانند متدولوژی «مدیریت جامع کیفیت داده‌ها»^۳ Wang (1998) مرحله ارتقا را نیز شامل می‌شوند.

در یک تقسیم‌بندی کلی می‌توان متدولوژی‌های کیفیت داده را در دو دسته کلی در

نظر گرفت: متدولوژی‌های عام‌منظوره^۱ و متدولوژی‌های خاص‌منظوره^۲. متدولوژی‌های عام‌منظوره، مانند متدولوژی ارائه‌شده توسط Wang (1998)، طیف گسترده‌ای از مراحل، ابعاد و فعالیت‌ها را دربرمی‌گیرند، در حالی که متدولوژی‌های خاص‌منظوره بر فعالیت‌های خاص نظیر اندازه‌گیری (Vaziri, Mohsenzadeh & Habibi (2019)، تمیزسازی (Azeroual, Saake & Abuosba (2019) و ... و یا بر یک حوزه اطلاعاتی خاص به‌عنوان مثال، ثبت‌نام (Falorsi et al. (2003 و ...، یا در حوزه‌های کاربردی خاص مانند انبار داده (Jeusfeld, Quix & Jarke (1998) و غیره متمرکز هستند.

در تقسیم‌بندی دیگری، متدولوژی‌های کیفیت داده به دو دسته کلی اطلاعات‌محور و فرایند‌محور تقسیم‌بندی می‌شوند. این طبقه‌بندی مربوط به استراتژی کلی است که برای فرایند ارتقای کیفیت داده انتخاب شده است. استراتژی‌های اطلاعات‌محور، مانند متدولوژی ارائه‌شده توسط Jeusfeld, Quix & Jarke (1998) منحصراً مبتنی بر استفاده از منابع اطلاعاتی برای بهبود کیفیت اطلاعات هستند؛ اما در استراتژی‌های فرایند‌محور نظیر متدولوژی پیشنهادشده توسط Loshin (2001) فرایند تولید اطلاعات، تجزیه و تحلیل شده و احتمالاً اصلاح می‌شود تا دلایل اصلی مشکلات کیفیتی شناسایی و حذف شوند (Batini and Scannapieco 2016).

۴. پیشینه پژوهش

مرور ادبیات در حوزه متدولوژی‌های کیفیت داده نشان می‌دهد که تاکنون متدولوژی‌های بسیاری برای توانمندسازی سازمان‌ها در ارزیابی و ارتقای سطح کیفیت داده‌های آن‌ها پیشنهاد شده است که هر کدام با توجه به معیارهایی مانند استراتژی، مرحله‌های مورد پوشش، فعالیت‌ها، ابعاد کیفیت داده و ... با یکدیگر متفاوت هستند. همچنین، در ادبیات کیفیت داده پژوهش‌هایی نیز برای طبقه‌بندی این حوزه ارائه شده است.

در پژوهشی که توسط «وانگ، ستوری و فیرث» انجام شده، تحلیلی بر ادبیات کیفیت داده به‌منظور ایجاد شناخت کلی بر وضعیت تحقیقات در این حوزه انجام شده است. با استفاده از شباهت بین تولید محصول و تولید داده، این مقاله چارچوبی را برای

1. general purpose

2. special purpose

تحلیل تحقیقات کیفیت داده توسعه داده و آن را به عنوان مبنایی برای سازمان‌دهی ادبیات کیفیت داده به کار می‌برد. این چارچوب شامل هفت جزء است: مسئولیت‌های مدیریتی، هزینه‌های عملیاتی و تضمین کیفیت داده، تحقیق و توسعه، تولید، توزیع، مدیریت کارکنان، و عملکرد قانونی. نتایج حاصل از پژوهش نشان می‌دهد که اکثر تلاش‌های تحقیقاتی بر هزینه‌های عملیاتی و تضمین کیفیت داده، تحقیق و توسعه و تولید محصولات داده متمرکز شده است (Wang, Storey and Firth 1995).

در پژوهشی که توسط «بتینی» و همکاران انجام شده، فعالیت‌های شناسایی شده در متدولوژی‌های کیفیت داده به ۱۵ دسته تقسیم شده است که ۵ فعالیت مربوط به مرحله ارزیابی و ۱۰ فعالیت دیگر مربوط به مرحله ارتقای کیفیت داده‌هاست. در ادامه، ۱۳ متدولوژی کیفیت داده مورد بررسی و تحلیل قرار گرفته‌اند. همچنین، مشخص شده است که هر کدام از این ۱۳ متدولوژی شامل کدام یک از این فعالیت‌ها هستند (Batini et al. 2009).

در پژوهشی که توسط «سیدی» و همکاران انجام شده، ابعاد کیفیت داده معرفی شده در تحقیقات پیشین با استفاده از روش مرور نظام‌مند استخراج شده است. همچنین، از داده کاوی و تکنیک‌های آماری برای اندازه‌گیری وابستگی بین ابعاد کیفیت داده استفاده شده است (Sidi et al. 2012).

در پژوهشی که توسط «وودال، بورک و پارلیکاد» انجام شده، هشت متدولوژی کیفیت داده انتخاب شده و فعالیت‌هایی که در هر کدام از این متدولوژی‌ها انجام شده، مشخص شده‌اند. این پژوهش تنها بر مرحله ارزیابی تمرکز نموده و فعالیت‌های موجود در مرحله ارتقای کیفیت داده را نادیده گرفته است (Woodall, Borek and Parlikad 2013).

در مقاله‌ای که توسط «ویسکوف و ونگ» انجام شده، به مرور ادبیات پژوهشی در رابطه با متدولوژی‌های ارزیابی کیفیت داده حوزه سلامت پرداخته شده است. در این پژوهش با استفاده از یک فرایند تکرارشونده، جنبه‌های کیفیت داده اندازه‌گیری شده و روش‌های ارزیابی استفاده شده خلاصه و دسته‌بندی شده‌اند. نتایج حاصل از این پژوهش پنج بُعد کیفیت داده شامل کامل بودن، درست بودن، تطابق، محتمل بودن، و معاصر بودن و هفت دسته کلی روش‌های ارزیابی کیفیت داده شامل مقایسه با استانداردهای طلایی، توافق جزء داده‌ها، توافق منبع داده‌ها، مقایسه توزیع، بررسی اعتبار، بررسی ورود به سیستم، و حضور جزء را مشخص می‌کند (Weiskopf and Weng 2013).

در پژوهشی دیگر نتایج حاصل از مرور نظام‌مند رویکردهای ارزیابی کیفیت داده پیوندی باز^۱ LOD ارائه شده است. در این پژوهش رویکردهای موجود جمع‌آوری شده و در ادامه، تحت یک طرح طبقه‌بندی مشترک مقایسه و طبقه‌بندی شده‌اند. این پژوهش رویکردهای موجود را بر اساس ابعاد کیفیت داده، معیارها، انواع داده و سطح اتوماسیون طبقه‌بندی نموده است (Zaveri et al. 2013).

به‌رغم وجود تحقیقاتی در جهت جمع‌بندی تحقیقات کیفیت داده (Sidi et al. 2013; Wang, Storey and Firth 1995; Zaveri et al. 2013; Weiskopf and Weng 2013)، این پژوهش‌ها غالباً مطالعات اولیه را بر اساس مواردی مانند ابعاد ارزیابی کیفیت داده، روش‌ها و تکنیک‌های مورد استفاده، انواع داده و ... طبقه‌بندی نموده‌اند و شناخت و جمع‌بندی متدولوژی‌ها تا حد زیادی مغفول باقی مانده و یا تنها بر بخش محدودی شامل فعالیت‌ها و یا ورودی‌ها و خروجی‌ها تمرکز کرده‌اند و پژوهشی که بتواند شناخت و دیدگاه جامعی نسبت به متدولوژی‌های کیفیت داده به لحاظ ورودی‌ها، فرایند (فعالیت‌ها) و خروجی‌های قابل دستیابی از متدولوژی‌ها فراهم نماید، انجام نشده است. از این رو، در پژوهش حاضر سعی بر آن است که از طریق مرور نظام‌مند و در نهایت، استفاده از روش فراترکیب، ترکیب جامعی از متدولوژی‌های کیفیت داده ارائه گردد تا تصویر روشنی نسبت به متدولوژی‌های موجود ایجاد شود.

۴. روش پژوهش

روش مورد استفاده در این پژوهش فراترکیب است. فراترکیب روش تحقیقی است که برای تولید ترجمه‌های تفسیری، روایت‌های زمینه‌ای یا نظریه‌هایی با تلفیق و مقایسه یافته‌ها یا استعاره‌های مطالعات کیفی مختلف مورد استفاده قرار می‌گیرد (Beck, 2002; Sandelowski Docherty and Emden 1997). فراترکیب یک ترکیب تفسیری به‌جای یک جمع‌بندی کلی از یافته‌ها ایجاد می‌کند. بر خلاف رویکرد فراتحلیل کمی که به داده‌های کمی از ادبیات موضوع و رویکردهای آماری دقیق متکی است، فراترکیب بر روی مطالعات کیفی متمرکز است که ممکن است لزوماً دربرگیرنده یک پایگاه بزرگ از ادبیات موضوع نشود (Noblit and Hare 1988). فراترکیب رویکردی نظام‌مند را برای محققان

1. Linked Open Data

فراهم می‌کند تا تحقیقات کیفی را ترکیب کرده، مضامین و استعاره‌های اساسی را کشف کنند، دانش فعلی را پیش ببرند و دیدگاهی گسترده و جامع تولید کنند. فراترکیب ابزار ارزشمندی برای تسهیل رویه نظریه‌سازی از طریق ترکیب نظام‌مند است (Siau and Long, 2005).

از آنجا که هر متدولوژی شامل فرایندی با هدف تبدیل ورودی‌ها به خروجی‌های مشخص و در راستای اهداف از پیش تعریف شده است، رویکرد سیستمی، رویکرد مناسبی برای تحلیل و شناخت متدولوژی‌هاست. بنابراین، از رویکرد سیستمی به‌عنوان راهبردی برای ایجاد الگوی منسجم و جامع استفاده شده است.

برای انجام فراترکیب، روش هفت مرحله‌ای «سندلوسکی و باروسو» شامل مراحل زیر مورد استفاده قرار گرفته است (Sandelowski and Barroso 2006):

- ۱) تنظیم سؤال تحقیق، ۲) مرور نظام‌مند پیشینه، ۳) جست‌وجو و انتخاب مقالات مناسب، ۴) استخراج اطلاعات مقالات، ۵) تجزیه و تحلیل و ترکیب یافته‌های کیفی، ۶) کنترل کیفیت، و ۷) ارائه یافته‌ها.

شایان ذکر است با توجه به این که پژوهش حاضر به ارائه چارچوبی جهت شناخت جامع متدولوژی‌های کیفیت داده می‌پردازد، از منظر هدف و ماهیت، بنیادی به‌شمار می‌رود. از طرف دیگر، از آنجا که پژوهش حاضر می‌تواند یاریگر سازمان‌ها در انتخاب یا ترکیب متدولوژی‌های کیفیت داده متناسب با نیازمندی‌های آن‌ها باشد، کاربردی نیز به‌شمار می‌رود. پژوهش حاضر همچنین، از منظر جمع‌آوری داده‌ها کیفی محسوب می‌شود.

مرحله اول: تعیین سؤال پژوهش

برای تعیین سؤال پژوهش طبق جدول ۱، از پارامترهای What (چه چیزی)، Who (جامعه مورد مطالعه) و When (محدودیت زمانی) استفاده شده است. همان‌گونه که اشاره شد، اجزای تشکیل‌دهنده متدولوژی‌های کیفیت داده با در نظر گرفتن دیدگاه سیستمی به سه بخش ورودی، فرایند و خروجی شکسته شده‌اند.

جدول ۱. پارامترها و پاسخ سؤالات

پارامترهای سؤالی	پاسخ سؤال
چه چیزی	شناسایی ورودی‌ها، فرایند (فعالیت‌ها) و خروجی‌های متدولوژی‌ها
جامعه مورد مطالعه	ژورنال‌ها و کنفرانس‌های معتبر نمایه‌شده در نمایه‌های استنادی اسکوپوس و وب‌آوساینس
محدودیت زمانی	بازه زمانی قبل از ۲۰۲۰

مرحله دوم: مرور نظام‌مند ادبیات

در این مرحله جست‌وجوی نظام‌مند کلمات کلیدی منتخب بر روی مقالات انگلیسی منتشرشده در ژورنال‌ها و کنفرانس‌های معتبر و نمایه‌شده در نمایه‌های استنادی «اسکوپوس»^۱ و «وب‌آوساینس»^۲ در بازه زمانی سال‌های قبل از ۲۰۲۰ انجام گرفته است. شایان ذکر است که جهت انتخاب کلمات کلیدی منتخب واژه methodology در لغت‌نامه مورد جست‌وجو قرار گرفته، کلمات مترادف آن مشخص شده‌اند و برای دستیابی به مجموعه کامل تری از منابع (مقالات) مورد استفاده قرار گرفتند. این کلمات عبارت‌اند از: methodology, method, framework, approach, technique, manner, system, way, fashion, strategy, form, how, practice, tactic, mode, channel, design, organization, plan, process. عبارات «data quality» و «information quality»^۳ در قسمت عنوان و کلمه متدولوژی و سایر کلمات مترادف با آن در قسمت «عنوان، چکیده، کلمات کلیدی» مورد جست‌وجو قرار گرفته‌اند.

مرحله سوم: جست‌وجو و انتخاب مقاله‌های مناسب

در این مرحله، محقق در هر گام ارزیابی منابع تعدادی از مقالات را حذف می‌کند که این منابع در فرایند فزاینده‌تر کپیب مورد بررسی قرار نمی‌گیرند. در جست‌وجوی انجام‌شده در پایگاه‌های داده مذکور پس از حذف موارد تکراری جمعاً ۳۹۰۹ مقاله یافت شد که در مرحله اول ارزیابی با مطالعه عنوان، چکیده و کلمات کلیدی تعداد مقالات به ۵۰۵ مقاله تقلیل یافت. به عبارت دیگر، با بررسی عنوان مقالات و محتوای چکیده‌ها حجم زیادی از مقالات نامرتبط با اهداف تحقیق کنار گذاشته شد. در مرحله دوم ارزیابی، محتوای

1. Scopus

2. Web of Science (WoS)

۳. در ادبیات کیفیت داده واژه‌های information quality و data quality هم‌معنا به کار می‌روند.

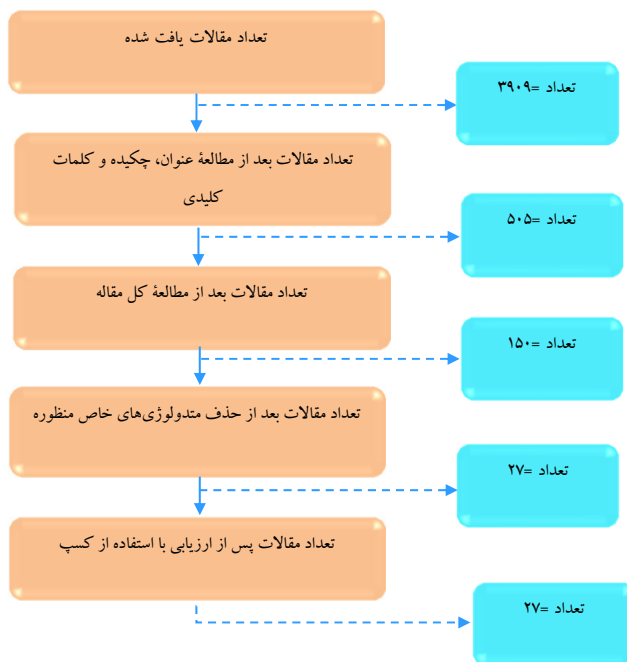
مقالات منتخب بررسی شد و با حذف مقالات نامرتبط، تعداد مقالات به ۱۵۰ مورد کاهش یافت. در این مرحله پس از مطالعه موارد کلیدی مندرج در هر مقاله از جمله سؤالات تحقیق، اهداف تحقیق، چارچوب نظری، نتایج و یافته‌ها و مقایسه آن‌ها با اهداف پژوهش حاضر، تعداد ۳۵۵ مقاله حذف گردید. با توجه به این که متدولوژی‌های عام‌منظوره طیف گسترده‌ای از مراحل، ابعاد و فعالیت‌ها را دربرمی‌گیرد و در مقابل، متدولوژی‌های خاص‌منظوره بر فعالیت‌های خاص، حوزه‌های اطلاعاتی یا بر حوزه‌های کاربرد خاص متمرکز هستند، در مرحله سوم ارزیابی تنها مقالاتی انتخاب شدند که عام‌منظوره بودند و مقالات خاص‌منظوره کنار گذاشته شدند. از آنجا که این پژوهش در صدد شناخت و تبیین ورودی‌ها، فرایند و خروجی‌های متدولوژی‌های کیفیت داده بود، متدولوژی‌های عام‌منظوره که شامل گستره‌ای وسیع از این موارد بودند، انتخاب گردید. به عبارت دیگر، در این مرحله، معیار اساسی جهت گزینش مقالات منتخب، تأکید بر عام‌منظوره بودن متدولوژی کیفیت داده در مقالات بوده است. در نهایت، تعداد ۲۷ مقاله که بیشترین تناسب را با موضوع پژوهش حاضر داشتند، استخراج شدند. همچنین، از برنامه مهارت‌های ارزیابی حیاتی جهت ارزیابی کیفیت پژوهش‌ها استفاده شده است.

برنامه مهارت‌های ارزیابی حیاتی (کسپ)^۱، ابزاری است که معمولاً برای ارزیابی کیفیت مطالعه‌های اولیه پژوهش کیفی استفاده می‌شود. در واقع، «کسپ» شامل ۱۰ سؤال است که به محقق کمک می‌کند تا مفهوم تحقیق کیفی را دریابد. به عبارت دیگر، این ابزار به محقق کمک می‌کند تا دقت، اعتبار و اهمیت مطالعات کیفی تحقیق را مشخص کند. این سؤالات بر موارد زیر تمرکز دارد: (۱) اهداف تحقیق، (۲) منطق روش، (۳) طرح تحقیق، (۴) روش نمونه‌برداری، (۵) جمع‌آوری داده‌ها، (۶) انعکاس‌پذیری، (۷) ملاحظات اخلاقی، (۸) دقت تجزیه و تحلیل داده‌ها، (۹) بیان واضح و روشن یافته‌ها، و (۱۰) ارزش پژوهش. در این مرحله، پژوهشگر به هر کدام از این سؤالات یک امتیاز کمی اختصاص می‌دهد و سپس، یک فرم را ایجاد می‌کند. بنابراین، او می‌تواند امتیازاتی را که به هر مقاله اختصاص می‌دهد جمع کند و به آسانی و به اجمال مجموعه مقالات را بررسی کرده و نتایج ارزیابی را ببیند. بر اساس مقیاس ۵۰ امتیازی «کسپ» روبریک مقالاتی که امتیاز بین ۴۱ تا ۵۰ بگیرند عالی، مقالاتی که امتیاز بین ۳۱ تا ۴۰ بگیرند خیلی خوب، مقالاتی که امتیاز بین ۲۱ تا ۳۰ بگیرند خوب، مقالاتی که امتیاز بین ۱۱ تا

1. Critical Appraisal Skills Programme (CASP)

۲۰ بگیرند متوسط و مقالاتی که بین صفر تا ۱۰ بگیرند ضعیف در نظر گرفته می‌شود. در ارزیابی مقالات با توجه به امتیازات تخصیص‌یافته دو مقاله «عالی»، ۱۵ مقاله «خیلی خوب» و ۱۰ مقاله «خوب» در نظر گرفته شده‌اند. به این ترتیب، هیچ‌یک از مقالات نهایی شده در این مرحله حذف نشده و تمامی ۲۷ مقاله وارد مرحله بعد شدند.

در شکل ۱، فرایند ارزیابی مورد استفاده در پژوهش حاضر نشان داده شده است. علاوه بر این، در جدول ۲، کد مقالات انتخاب‌شده (مشخصات کامل مقالات در جدول پیوست ارائه شده است)، مرحله‌های پوشش داده‌شده در این مقالات، کمی و یا کیفی بودن نوع ارزیابی و استراتژی انتخابی برای ارتقا ارائه شده است.



شکل ۱. خلاصه‌ای از نتایج جست‌وجو و انتخاب مقالات

مرحله چهارم: استخراج اطلاعات از مقاله‌ها

در ادامه، مقالات منتخب و نهایی شده به‌منظور دستیابی به یافته‌های برگزیده و مرتبط مورد مطالعه قرار گرفته‌اند. در این مرحله با در نظر داشتن دیدگاه سیستمی به‌عنوان رویکردی کارآمد در ایجاد الگویی منسجم و جامع، ورودی‌ها، فرایندها (فعالیت‌ها)، خروجی‌ها از متدولوژی‌های ارائه‌شده در مقالات منتخب و نهایی شده استخراج شده‌اند.

مشخصات کامل مقالات انتخابی در جدول ۲، ارائه شده است.

جدول ۲. مقالات انتخابی در روش فراترکیب و ویژگی‌های آنها

کد مقاله	مرحله‌های پوشش داده‌شده در متدولوژی	استراتژی انتخابی برای ارتقا	عینی و یا ذهنی بودن نوع ارزیابی
۱	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور و داده‌محور	عینی
۲	بازسازی وضعیت، ارزیابی	-	عینی و ذهنی
۳	ارزیابی	-	عینی
۴	بازسازی وضعیت، ارزیابی	فرایندمحور	ذهنی
۵	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور	عینی
۶	ارزیابی، ارتقا	-	ذهنی
۷	بازسازی وضعیت، ارزیابی	-	عینی
۸	بازسازی وضعیت، ارزیابی	داده‌محور	عینی
۹	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور و داده‌محور	عینی و ذهنی
۱۰	بازسازی وضعیت، ارزیابی	-	عینی و ذهنی
۱۱	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور	عینی و ذهنی
۱۲	بازسازی وضعیت، ارزیابی	داده‌محور	عینی و ذهنی
۱۳	ارزیابی، ارتقا	فرایندمحور	عینی
۱۴	بازسازی وضعیت، ارزیابی، ارتقا	داده‌محور	عینی و ذهنی
۱۵	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور و داده‌محور	عینی و ذهنی
۱۶	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور و داده‌محور	عینی و ذهنی
۱۷	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور و داده‌محور	عینی و ذهنی
۱۸	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور و داده‌محور	عینی
۱۹	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور	عینی و ذهنی
۲۰	بازسازی وضعیت، ارزیابی	-	عینی و ذهنی
۲۱	بازسازی وضعیت، ارزیابی	-	عینی و ذهنی
۲۲	بازسازی وضعیت، ارزیابی	-	عینی
۲۳	بازسازی وضعیت، ارزیابی	-	عینی و ذهنی
۲۴	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور	عینی و ذهنی

کد مقاله مرحله‌های پوشش داده‌شده در متدولوژی استراتژی انتخابی برای ارتقا عینی و یا ذهنی بودن نوع ارزیابی

۲۵	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور	عینی
۲۶	بازسازی وضعیت، ارزیابی	-	عینی و ذهنی
۲۷	بازسازی وضعیت، ارزیابی، ارتقا	فرایندمحور	عینی و ذهنی

مرحله پنجم: تجزیه و تحلیل و تلفیق یافته‌های کیفی

هدف فراترکیب به‌عنوان روشی پذیرفته‌شده در شفاف‌سازی مفاهیم، الگوها و نتایج حاصل از پالایش حالت‌های موجود دانش و ظهور مدل‌های عملیاتی و تئوری‌های پذیرفته‌شده، ایجاد تفسیری یکپارچه و جدید از یافته‌هاست. در طول تجزیه و تحلیل، محقق موضوعات یا تم‌هایی^۱ را جست‌وجو می‌کند که در میان مطالعات موجود در فراترکیب پدیدار شده‌اند. «ساندلوسکی و باروسو» به این شیوه به‌عنوان «بررسی موضوعی» اشاره می‌کنند، به‌طوری که محقق در ابتدا تم‌ها یا موضوعاتی را شناسایی و مشخص می‌کند. در ادامه، پس از شناسایی و مشخص‌شدن موضوعات، بررسی‌کننده یک طبقه‌بندی موضوعی را شکل داده و موضوعات مشابه را ذیل موضوعی قرار می‌دهد که آن را به بهترین نحو ممکن توصیف می‌کند. تم‌ها اساس و پایه‌ای را برای ایجاد «توضیحات و مدل‌ها، تئوری‌ها و فرضیات کاری» ارائه می‌دهند (Sandelowski and Barroso 2006).

مرحله ششم: حفظ کنترل کیفیت

در روش فراترکیب، محقق رویه‌های زیر را برای حفظ کیفیت در مطالعه خود در نظر می‌گیرد:

- ◇ محقق در سرتاسر تحقیق سعی دارد تا با فراهم‌نمودن توضیحات و توصیفات روشن و واضح برای گزینه‌های موجود در تحقیق گام‌های اتخاذشده را بردارد؛
- ◇ هر دو استراتژی جست‌وجوی الکترونیک و دستی برای یافتن مقالات مربوط توسط محقق به کار گرفته می‌شود؛
- ◇ محقق، روش‌های کنترل کیفیت استفاده‌شده در مطالعات تحقیق کیفی اصلی را به کار می‌برد؛
- ◇ از ابزار «کسپ» برای ارزیابی مطالعات اولیه مورد استفاده در پژوهش استفاده می‌کند.

مرحله هفتم: ارائه یافته‌ها

در این مرحله از روش فراترکیب یافته‌های حاصل از انجام مراحل قبلی ارائه می‌شود. همان‌گونه که در مراحل پیش نیز گفته شد، با استفاده از روش فراترکیب متدولوژی‌های عام‌منظوره کیفیت داده مورد بررسی قرار گرفت و کدهای مرتبط با متدولوژی‌ها در چارچوب ورودی‌ها، فعالیت‌ها و خروجی‌ها استخراج شد. در ادامه، با ترکیب کدهای استخراج‌شده مشابه در یک مفهوم واحد (تم‌های تحقیق)، مقوله‌های سه‌گانه که در جداول ۳، ۴ و ۵ به تفکیک ارائه شده‌اند، تبیین گردید.

نتایج حاصل از فراترکیب نشان می‌دهد که ورودی‌های متدولوژی‌های کیفیت داده در ۹ دسته متمرکز هستند. این ورودی‌ها همان‌گونه که در جدول ۳ نیز ارائه شده‌اند، عبارت‌اند از: منابع (منابع فنی، منابع انسانی و ...)، منابع اطلاعاتی داخلی و جریان‌های اطلاعاتی، داده‌های مرجع، سازمان و زمینه آن، قوانین کسب‌وکار، فرایندها، ابعاد کیفیت داده و بودجه. از بین ورودی‌های مشخص شده «منابع اطلاعاتی داخلی و جریان‌های اطلاعاتی» بیشترین تکرار را به خود اختصاص داده است. پس از آن قوانین کسب‌وکار یک ورودی بسیار مهم برای متدولوژی‌های کسب‌وکار فراهم به‌شمار می‌رود. شایان ذکر است که ورودی «داده‌های مرجع» به‌رغم این که در میان مقالات مورد بررسی تنها در سه مورد به صراحت مطرح شده است، یک ورودی مهم در متدولوژی‌های کیفیت داده به‌شمار می‌رود؛ چرا که برای اندازه‌گیری ابعادی مانند «دقت»^۱ بسیار ضروری است. ورودی‌های متدولوژی‌های کیفیت داده در سه دسته کلی شامل «زمینه و وضعیت سازمان»، «داده‌ها و منابع اطلاعاتی» و «ابعاد کیفیت داده» طبقه‌بندی شده‌اند.

1. accuracy

جدول ۳. ورودی‌ها^۱

کدهای اصلی	کدهای فرعی	منابع
زمینه و وضعیت سازمان	منابع سازمانی	۱۶،۱۹،۲۳
	سازمان و زمینه آن	۱،۲،۱۱،۱۶،۲۳،۲۱،۲۳،۲۶
	قوانین کسب و کار	۱،۳،۵،۹،۱۱،۱۴،۱۶،۱۷،۱۸،۱۹،۲۰،۲۱،۲۲ ۳،۲۴،۲۵
	فرایندهای کسب و کار	۱،۲،۴،۵،۹،۱۰،۱۳،۱۴،۱۶،۱۷،۱۸،۲۱،۲۲ ۳،۲۴
	بودجه تخصیص یافته	۱،۵،۱۱،۱۹،۲۳
داده‌ها و منابع اطلاعاتی	منابع اطلاعاتی داخلی و جریان‌های اطلاعاتی	۱،۲،۳،۴،۵،۷،۹،۱۱،۱۲،۱۳،۱۴،۱۶،۱۷،۱۸، ۱۹،۲۰،۲۱،۲۲،۲۴،۲۵
	منابع اطلاعاتی خارجی	۱،۱۱،۱۶
	داده‌های مرجع ^۱	۱۳،۱۷،۲۰
ابعاد کیفیت داده	ابعاد کیفیت داده	۳،۶،۹،۱۵،۲۰،۲۱،۲۶

افزون بر این، نتایج حاصل از فراترکیب، همان گونه که در جدول ۴ نیز ارائه شده، نشان می‌دهد که در مرحله «بازسازی وضعیت» فعالیت‌های مربوط به «پروفایلینگ داده» که خود در برگیرنده سه فعالیت «جمع‌آوری متادیتا»، «کشف قوانین کیفیت داده» و «تعیین داده‌های مرجع» است، بیشترین تکرار را در متدولوژی‌های کیفیت داده به خود اختصاص داده است. در مرحله «اندازه‌گیری/ارزیابی»، فعالیت «اندازه‌گیری کیفیت داده» که طی آن مقدار معیارهای کیفیت داده از مجموعه‌ای از داده‌های واقعی (عینی) و یا کسب نظرات در مورد وضعیت موجود کیفیت داده اندازه‌گیری می‌شود، بیشترین تکرار را در میان متدولوژی‌های کیفیت داده به خود اختصاص داده است. در میان متدولوژی‌های کیفیت داده مورد بررسی غالب متدولوژی‌ها (۱۷ مورد از ۲۷ مورد) تنها تا مرحله «اندازه‌گیری/ارزیابی» را پوشش داده‌اند و مرحله «ارتقا» کمتر مورد توجه قرار گرفته است. همچنین، فعالیت‌های «کنترل فرایند» و «بازطراحی فرایند» در مقایسه با فعالیت «اصلاح مقادیر داده» از تکرار بالاتری برخوردار بوده‌اند که دلالت بر تمرکز بیشتر ادبیات موضوع بر استراتژی‌های فرایندمحور در مقایسه با استراتژی‌های داده‌محور دارد.

1. reference data

جدول ۴. فعالیت‌ها^۱

منابع	کدهای فرعی	کدهای اصلی
۸،۱۱،۱۶،۲۵	تعیین استراتژی‌ها و اهداف سازمان	بازسازی وضعیت
۱۴،۱۶،۲۳،۲۶	تعیین ذی‌نفعان کلیدی و نیازمندی‌ها و مشکلات آن‌ها در رابطه با کیفیت داده	
۹،۱۹،۲۳	مستندسازی قوانین و استانداردهای کسب‌وکار	
۱،۲،۳،۴،۵،۶،۷،۸،۹،۱۰،۱۱،۱۲،۱۳، ۱۵،۱۶،۱۷،۱۸،۱۹،۲۰،۲۱،۲۲،۲۳،۲۴ ،۲۵،۲۶،۲۷	تعیین بحرانی‌ترین اقلام داده و جمع‌آوری داده‌ها	
۱،۲،۵،۷،۹،۱۱،۱۲،۱۵،۱۹،۲۱،۲۲،۲۳ ،۲۴،۲۵،۲۶	انجام پروفایلینگ داده جمع‌آوری متادیتا	
۲۰،۲۲،۲۵	کشف قوانین کیفیت داده	
۱۷،۱۸،۲۰،۲۳	تعیین داده‌های مرجع	
۱،۲،۳،۴،۵،۶،۷،۸،۹،۱۰،۱۱،۱۲،۱۳، ۱۵،۱۶،۱۷،۱۸،۱۹،۲۰،۲۱،۲۲،۲۳،۲۴ ،۲۵،۲۶،۲۷	تعیین ابعاد کیفیت داده ^۱	
۱،۲،۳،۴،۵،۶،۷،۸،۹،۱۰،۱۱،۱۲،۱۳، ۱۵،۱۶،۱۷،۱۸،۱۹،۲۰،۲۱،۲۲،۲۳،۲۴ ،۲۵،۲۶،۲۷	تعیین معیارهای کیفیت داده ^۲ و روش‌های اندازه‌گیری	
۱،۲،۳،۴،۵،۶،۷،۸،۹،۱۰،۱۱،۱۲،۱۳، ۱۵،۱۶،۱۷،۱۸،۱۹،۲۰،۲۱،۲۲،۲۳،۲۴ ،۲۵،۲۶،۲۷	تعیین نیازمندی‌های کیفیت داده	
۱،۲،۴،۵،۱۵،۱۸،۱۹	تعیین بحرانی‌ترین فرایندها	
۱،۲،۴،۵،۱۱،۱۴،۱۶،۱۸،۲۱،۲۳،۲۴	مدل‌سازی فرایندها	
۵،۷،۱۱	تعیین مکان و زمان اندازه‌گیری	

منابع	کدهای فرعی	کدهای اصلی
۱،۲،۳،۴،۵،۶،۷،۸،۹،۱۰،۱۳،۱۴،۱۵،۱۶،۱۷،۱۸،۱۹،۲۰،۲۱،۲۲،۲۴،۲۵،۲۶	اندازه‌گیری کیفیت داده	اندازه‌گیری / ارزیابی
۵،۱۰،۱۴،۱۵	تأیید و معتبرسازی معیارهای کیفیت داده	
۱،۲،۳،۶،۹،۱۲،۱۵،۱۶،۱۷،۱۸،۲۲،۲۷	تحلیل اندازه‌گیری	
۷،۱۱،۱۲،۱۵،۱۷،۲۰،۲۴،۲۵،۲۷	گزارش‌دهی	
۲،۲۵	ارسال اخطار	
۲،۱۱،۱۴،۱۵،۱۹،۲۷	تعیین هزینه‌های یا ریسک‌های ناشی از کیفیت پایین داده‌ها	
۱،۵،۹،۱۱،۱۳،۱۴،۱۶،۱۹،۲۳،۲۴،۲۵،۲۷	تعیین دلایل ریشه‌ای خطاها	ارتقای کیفیت داده
۱،۵،۱۵،۱۶،۱۷،۱۸،۲۷	انتخاب استراتژی‌ها و تکنیک‌ها	
۱،۱۴،۱۷،۱۸،۱۹	اصلاح مقادیر داده	
۱،۱۳،۱۶،۱۷،۱۸،۱۹،۲۴	کنترل فرایند	
۱،۵،۹،۱۱،۱۳،۱۴،۱۶،۱۸،۱۹،۲۴،۲۷	بازطراحی فرایند	
۱۱،۱۶،۱۷،۱۹،۲۵،۲۷	مدیریت ارتقا	
۹،۱۳،۱۴،۱۶،۱۷،۱۸،۱۹،۲۴،۲۵	نظارت بر ارتقا	

نتایج حاصل از فراترکیب در رابطه با خروجی‌های متدولوژی‌های کیفیت داده، همان‌گونه که در جدول ۵ نیز ارائه شده، نشان می‌دهد که خروجی «جریان‌ها و پایگاه‌های داده» اندازه‌گیری‌شده با بالاترین تکرار در میان خروجی‌ها، در واقع، مهم‌ترین خروجی متدولوژی‌های کیفیت داده در نظر گرفته می‌شود و سایر خروجی‌ها به ترتیب تکرار عبارت‌اند از: «فعالیت‌های ارتقای کیفیت داده و تکنیک‌های مربوط به آن‌ها»، «فرایندهای بازطراحی‌شده یا مهندسی مجددشده»، «گزارش‌های کیفیت داده»، «فرایندهای کنترل‌شده»، «سیاست‌ها یا قوانین کیفیت داده تصحیح‌شده»، «پایگاه‌های داده ارتقا داده‌شده»، «هزینه‌ها و منفعت‌ها» و «اخطارهای کیفیت داده».

جدول ۵. خروجی‌ها

کدهای اصلی	کدهای فرعی	منابع
فهرست فعالیت‌ها و تکنیک‌های مرتبط برای ارتقای کیفیت داده‌ها	فعالیت‌های ارتقای کیفیت داده و تکنیک‌های مرتبط به آن‌ها	۱،۵،۹،۱۴،۱۶،۱۷،۱۸،۱۹،۲۴،۲۵،۲۷
فرایندهای کنترل شده یا بازطراحی شده	فرایند بهینه ارتقا	۱،۱۵،۱۸،۲۴
فرایندهای کنترل شده یا بازطراحی شده	فرایندهای کنترل شده	۱،۴،۱۱،۱۳،۱۷،۱۹،۲۴،۲۷
فرایندهای کنترل شده یا بازطراحی شده	فرایندهای بازطراحی شده یا مهندسی مجدد شده	۱،۵،۹،۱۱،۱۳،۱۶،۱۸،۱۹،۲۴،۲۵،۲۷
جریان‌ها و پایگاه‌های داده اندازه‌گیری یا ارتقا داده شده	جریان‌ها و پایگاه‌های داده اندازه‌گیری شده	۱،۲،۳،۵،۶،۷،۸،۱۰،۱۴،۱۵،۱۶،۱۷،۱۹،۲۱،۲۲،۲۴،۲۵،۲۶
نتایج ارائه شده از وضعیت کیفیت داده	پایگاه‌های داده ارتقا داده شده	۱،۵،۱۴،۱۸
نتایج ارائه شده از وضعیت کیفیت داده	اخطارهای کیفیت داده	۲،۲۵
سیاست‌ها یا قوانین کیفیت داده	گزارش‌ها کیفیت داده	۳،۷،۱۰،۱۲،۱۵،۱۷،۲۰،۲۴،۲۵،۲۷
سیاست‌ها یا قوانین کیفیت داده	سیاست‌ها یا قوانین کیفیت داده تصحیح شده	۱۱،۱۶،۱۷،۱۹،۲۵،۲۷
هزینه‌ها و منفعت‌ها	هزینه‌ها و منفعت‌ها	۱،۲،۱۵،۲۴

نتایج حاصل از انجام فراترکیب ضمن در نظر داشتن رویکرد سیستمی در شکل ۲، ارائه شده است. همان‌طور که در بخش‌های پیشین به آن اشاره شد، هر متدولوژی شامل فرایندی است که با استفاده از ورودی‌های مقتضی منجر به تولید خروجی‌های مطابق با اهداف از پیش تعیین شده می‌شود و به عبارت دیگر، متدولوژی ماهیت سیستمی دارد. بنابراین، کدهای اصلی جداول ۳، ۴ و ۵ شامل ورودی‌ها، فرایندها (فعالیت‌ها) و خروجی‌های مربوط به ۲۷ متدولوژی کیفیت داده به عنوان جمع‌بندی نهایی تدوین شده است.



شکل ۲. نتایج حاصل از فراترکیب

۵. نتیجه‌گیری و ارائه پیشنهادات

داده‌ها یک دارایی اساسی در سازمان‌های امروزی هستند و مشکلات مربوط به کیفیت این داده‌ها می‌تواند مخاطرات متعدد و گاه فاجعه‌آمیزی را به همراه داشته باشد. متدولوژی‌های کیفیت داده می‌توانند یاریگر سازمان‌ها در ارزیابی و ارتقای کیفیت داده‌ها باشند. در این پژوهش سعی شده است تا با استفاده از روش فراترکیب و ضمن به کارگیری رویکرد سیستمی در جهت ایجاد الگویی جامع و منسجم، شناخت جامع و کاملی از متدولوژی‌های کیفیت داده به دست آید. نتایج این پژوهش الگوی جامعی را نسبت به فعالیت‌های اساسی مورد نیاز برای ارزیابی و ارتقای کیفیت داده در قالب‌های سه‌گانه (در چارچوب رویکرد سیستماتیک) شامل ورودی‌ها، فرایند و خروجی‌ها در متدولوژی‌های کیفیت داده فراهم نموده است و فهرستی از کدهای مرتبط با هر مقوله استخراج شده است. ورودی‌های متدولوژی‌های کیفیت داده در سه دسته کلی، شامل زمینه و وضعیت سازمان، داده‌ها و منابع اطلاعاتی و ابعاد کیفیت داده طبقه‌بندی شده‌اند. نتایج حاصل از پژوهش نشان می‌دهد که در بخش ورودی‌ها، تمرکز اکثر پژوهش‌ها بر منابع اطلاعاتی داخلی است و به رغم اهمیت بالای داده‌ها و منابع اطلاعاتی خارجی و نقش تعیین‌کننده آن‌ها در تصمیمات سازمانی، منابع داده خارجی تا حد زیادی مغفول مانده است. همچنین، منابع سازمانی مانند منابع انسانی، تجهیزات، و بودجه به عنوان پیش‌نیازهای ضروری پیاده‌سازی متدولوژی‌های کیفیت داده کمتر در تحقیقات مورد توجه قرار گرفته‌اند. افزون بر این، با

وجود این که داده‌های مرجع به عنوان معیاری برای اندازه‌گیری، ارزیابی و ارتقای کیفیت داده ضروری هستند، در تعداد اندکی از متدولوژی‌ها به صراحت به آن اشاره شده است. نتایج این پژوهش در فرایندهای متدولوژی‌های کیفیت داده حاکی از آن است که فعالیت‌های کیفیت داده به عنوان اجزای تشکیل دهنده فرایند مدیریت کیفیت داده در ۲۶ کد فرعی دسته‌بندی می‌شوند و در نهایت، این ۲۶ کد فرعی در قالب ۳ کد اصلی که در واقع، مراحل اصلی یک متدولوژی کیفیت داده به شمار می‌روند، طبقه‌بندی شده‌اند. این مراحل عبارت‌اند از بازسازی وضعیت، اندازه‌گیری/ارزیابی و ارتقا. نتایج حاصل از پژوهش نشان می‌دهد که فعالیت‌های «تعیین بحرانی‌ترین اقلام داده و جمع‌آوری داده‌ها»، «تعیین ابعاد کیفیت داده»، «تعیین معیارهای کیفیت داده» و «اندازه‌گیری کیفیت داده» در همه متدولوژی‌های کیفیت داده دیده می‌شود و بنابراین، وجود این فعالیت‌ها در هر متدولوژی کیفیت داده الزامی به شمار می‌رود. در این پژوهش سعی شده است فعالیت‌های عنوان شده در متدولوژی‌ها تا حد امکان به اجزای تشکیل دهنده آن‌ها شکسته شوند. برخی از این فعالیت‌ها مانند تعیین ابعاد کیفیت داده و اندازه‌گیری تقریباً در همه متدولوژی‌ها به طور شفاف مطرح شده‌اند، در عین حال، برخی از فعالیت‌ها مانند تعیین و جمع‌آوری داده‌های مرجع به طور شفاف بیان نشده‌اند و باید در زمان اجرای متدولوژی و به تناسب ابعاد کیفیت داده انتخاب شده برای اندازه‌گیری و همچنین، استراتژی انتخابی (داده‌محور یا فرایندمحور بودن) در مورد پیاده‌سازی یا عدم پیاده‌سازی آن تصمیم‌گیری نمود.

مطابق با نتایج پژوهش، خروجی‌های متدولوژی‌های کیفیت داده در شش دسته کلی شامل فهرست فعالیت‌ها و تکنیک‌های مرتبط برای ارتقای کیفیت داده‌ها، فرایندهای کنترل شده یا بازطراحی شده، جریان‌ها و پایگاه‌های داده اندازه‌گیری یا ارتقای داده‌شده، نتایج ارائه شده از وضعیت کیفیت داده، سیاست‌ها یا قوانین کیفیت داده تصحیح شده و هزینه‌ها و منفعت‌ها طبقه‌بندی شده‌اند. خروجی‌های قابل حصول از متدولوژی‌های کیفیت داده تا حد زیادی وابسته به استراتژی انتخابی (داده‌محور یا فرایندمحور) و شمول یا عدم شمول برخی از فعالیت‌ها در مرحله ارتقا است. به عنوان مثال، فرایندهای کنترل شده یا بازطراحی شده تنها در متدولوژی‌های فرایندمحور وجود دارد و یا سیاست‌ها و قوانین کیفیت داده تصحیح شده تنها در متدولوژی‌هایی قابل دستیابی است که در مرحله ارتقا شامل فعالیت مدیریت ارتقا باشند.

همان‌طور که در بخش‌های پیشین نیز اشاره شد، اگرچه تحقیقاتی به‌منظور جمع‌بندی تحقیقات کیفیت داده انجام شده است (Sidi et al. 2012; Wang, Storey and Firth 1995; Zaveri et al. 2013; Weiskopf & Weng 2013)، این پژوهش‌ها، غالباً مطالعات اولیه را بر اساس مواردی مانند ابعاد ارزیابی کیفیت داده، روش‌ها و تکنیک‌های مورد استفاده، انواع داده و ... طبقه‌بندی نموده‌اند و شناخت و جمع‌بندی متدولوژی‌ها تا حد زیادی مغفول مانده است. در پژوهش انجام‌شده توسط «بتینی» و همکاران که به مقایسه ۱۳ متدولوژی کیفیت داده پرداخته است، تنها بر مرحله‌ها و فعالیت‌های تشکیل‌دهنده هر مرحله پرداخته و ورودی‌ها و خروجی‌های قابل دستیابی متدولوژی‌ها مغفول مانده‌اند (Batini et al. 2009). پژوهش حاضر علاوه بر استفاده از روش نظام‌مند در جست‌وجوی مقالات و در نتیجه، شمول تعداد بیشتر پژوهش‌های مورد استفاده و همچنین، استفاده از رویکرد سیستمی و ارائه فهرستی جامع از ورودی‌ها، فعالیت‌ها و خروجی‌های کیفیت داده می‌تواند شناخت جامع‌تری نسبت به متدولوژی‌های کیفیت داده فراهم آورد. همچنین، در مقایسه با پژوهش انجام‌شده توسط «بتینی و اسکناپیکو» که به جمع‌بندی ورودی‌ها و خروجی‌های متدولوژی‌ای کیفیت داده پرداخته، در بخش ورودی‌ها، «استراتژی و اهداف سازمان»، «منابع سازمانی» و «داده‌های مرجع» و در بخش خروجی‌ها «سیاست‌های و قوانین کیفیت داده تصحیح‌شده» اشاره نشده‌اند که در چارچوب ارائه‌شده در این پژوهش مشخص شده‌اند (Batini and Scannapieco 2016). افزون بر این، پژوهش حاضر در مقایسه با پژوهش انجام‌شده توسط (Woodall, Borek and Parlikad 2013)، ضمن شمول تعداد بیشتری از متدولوژی‌ها، ورودی‌ها و خروجی‌های متدولوژی‌ها را نیز پوشش داده و مرحله ارتقا را نیز دربرمی‌گیرد.

این پژوهش که با هدف ارائه چارچوبی جدید و جامع در جمع‌بندی متدولوژی‌های کیفیت داده انجام شده، ضمن در نظر داشتن رویکرد سیستمی، تصویری یکپارچه از ادبیات متدولوژی‌های عام‌منظوره کیفیت داده را فراهم می‌نماید. این پژوهش با ارائه فهرستی شفاف از ورودی‌ها، فعالیت‌ها و خروجی‌ها مبنایی دقیق جهت تشخیص تفاوت‌ها و همپوشانی‌های بین متدولوژی‌های موجود و حتی متدولوژی‌هایی که در آینده توسعه داده خواهند شد، هم در زمینه فرایندها و هم در زمینه ورودی‌ها و خروجی‌ها فراهم نموده است. این پژوهش به پژوهشگران و سازمان‌ها کمک می‌کند تا ضمن شناخت نقاطی که در متدولوژی‌های کیفیت داده مورد توجه و تمرکز واقع شده‌اند و با توجه به نیازمندی‌های

پویا و متغیر سازمان‌ها در بحث مدیریت کیفیت داده (مانند مطرح شدن تکنولوژی‌های جدید همچون داده‌های بزرگ^۱، مدیریت کیفیت داده‌های رسانه‌های اجتماعی و ...)، متدولوژی‌ها را بهبود بخشیده و یا متدولوژی‌هایی متناسب با نیازمندی‌ها توسعه دهند. لازم به ذکر است که نیازمندی‌های سازمان‌ها در مدیریت کیفیت داده عمده‌تاً متفاوت و متغیر هستند. به عبارت دیگر، با استفاده از این متدولوژی می‌توان هر متدولوژی کیفیت داده را با چارچوب ارائه‌شده در این پژوهش مقایسه کرد و مشخص نمود که متدولوژی مورد نظر به لحاظ فرایند (فعالیت‌هایی که شامل می‌شود) و خروجی‌هایی که می‌تواند فراهم نماید تا چه حد می‌تواند پاسخگوی نیازمندی‌های مورد نظر در مدیریت کیفیت داده‌ها باشد و چه مواردی را پوشش نمی‌دهد. در نتیجه، استفاده از این چارچوب ابزار مناسبی را در اختیار پژوهشگران و سازمان‌ها در ارزیابی نقاط ضعف و قوت متدولوژی‌ها قرار می‌دهد.

شایان ذکر است که در پژوهش حاضر، عوامل زمینه‌ای مؤثر بر پیاده‌سازی متدولوژی‌های کیفیت داده (در دیدگاه سیستمی، محیط سیستم) شناسایی نشده‌اند. عوامل زمینه‌ای طیف گسترده‌ای از عوامل مانند فرهنگ سازمانی را دربرمی‌گیرد که در پژوهش‌های آتی به آن پرداخته خواهد شد. همچنین، ارزیابی چارچوب ارائه‌شده در این پژوهش در زمینه‌های خاص نظیر کیفیت داده‌های رسانه‌های اجتماعی و شناسایی نقاط ضعف و قوت آن و تکوین چارچوب پیشنهادی نیز می‌تواند در پژوهش‌های آتی انجام پذیرد. در عین حال، می‌توان از رویکرد استفاده‌شده در این پژوهش برای طراحی مدل بلوغ کیفیت داده در سازمان‌ها نیز استفاده نمود.

فهرست منابع

- محمدخانی، کامران، حمیده رشادت‌جو، و معصومه روحانی‌پور. ۱۳۹۱. بررسی رابطه تفکر سیستمی و میزان خلاقیت مدیران مدارس متوسطه دخترانه آموزش و پرورش شهر تهران و تأثیر آن بر آینده‌پژوهی. *آینده‌پژوهی مدیریت* ۲۳ (ویژه‌نامه شماره ۹۶): ۵۱-۶۳.
- محمدی، مهدی، و بهمن حاجی‌پور. ۱۳۹۷. شناسایی الگوی همکاری رقابت در صنعت خودروسازی ایران بر اساس رویکرد سیستمی: مطالعه داده‌بنیاد. *بررسی‌های بازرگانی* ۱۶ (۸۸-۸۹): ۱-۲۲.

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.
- Aljumaili, M., R. Karim, & P. Tretten. 2016. Metadata-based data quality assessment. *VINE Journal of Information and Knowledge Management Systems* 232-250 : (2) 46 .
- Azeroual, O., G. Saake, & M. Abuosba. 2019. Data quality measures and data cleansing for research information systems. *Journal of Digital Information Management* 16 (1): 12-21.
- Bargh, M. S., J. van Dijk, & S. Choenni. 2015. Dynamic data quality management framework using issue tracking systems. *IADIS International Journal on Computer Science & Information Systems* 10 (2): 32-51.
- Batini, C., and M. Scannapieco. 2016. *Data and information quality: Dimensions. Principles and Techniques*. Cham, Switzerland: Springer.
- Batini, C., C. Cappiello, C. Francalanci, & A. Maurino. 2009. Methodologies for data quality assessment and improvement. *ACM computing surveys (CSUR)* 41 (3): 1-52.
- Batini, C., D. Barone, M. Mastrella, A. Maurino, & C. Ruffini. 2007. A Framework and a Methodology for Data Quality Assessment and Monitoring. In *International Conference on Information Quality (ICIQ)*. Cambridge. USA. 333-346.
- Batini, C., F. Cabitza, C. Cappiello, & C. Francalanci. 2008. A comprehensive data quality methodology for web and structured data. *International Journal of Innovative Computing and Applications* 1 (3): 205-218.
- Beck, C. 2002. Mothering multiples: a meta-synthesis of qualitative research. *Maternal and Child Nursing* 27 (4): 214-221.
- Berti-Équille, L., I. Comyn-Wattiau, M. Cosquer, Z. Kedad, S. Nugier, V. Peralta, & V. Thion-Goasdoué. 2011. Assessment and analysis of information quality: a multidimensional model and case studies. *International Journal of Information Quality* 2 (4): 300-323.
- Bringel, H., A. Caetano, & J. Tribolet. 2004. Business process modeling towards data quality assurance; an organizational engineering approach. In *International Conference on Enterprise Information Systems*. 4. 565-568. SCITEPRESS. Porto, Portugal.
- Bugajski, J., & R. L. Grossman. 2007. An Alert Management Approach To Data Quality: Lessons Learned From The Visa Data Authority Program. In *International Conference on Information Quality (ICIQ)*. Cambridge. USA. 5-18. Cambridge 5-18.
- Caballero, I., & M. Piattini. 2003. CALDEA: a data quality model based on maturity levels. In *Third International Conference on Quality Software*, 2003. Proceedings. (pp. 380-387). IEEE. Texas, USA.
- Caballero, I., A. Caro, C. Calero, & M. Piattini. 2008. IQM3: Information Quality Management Maturity Model. *Journal of Universal Computer Science* 14 (22): 3658-3685.
- Caballero, I., A. Caro, M. Piattini, & C. Calero. 2005. Improving Information Quality Management Using Caldea and Evamecal. In *International Conference on Information Quality (ICIQ)*. Cambridge, USA.
- Caballero, I., E. Verbo, C. Calero, & M. Piattini. 2008. MMPRO: A Methodology Based on ISO/IEC 15939 to Draw Up Data Quality Measurement Processes. In *International Conference on Information Quality (ICIQ)*. Cambridge, USA. 326-340.
- Cappiello, C., C. Francalanci, and B. Pernici. 2005. A self-monitoring system to satisfy data quality requirements. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*. Berlin, Heidelberg: Springer. pp. 1535-1552.
- Cappiello, C., P. Ficiaro, & B. Pernici. 2006. HIQM: A methodology for information quality monitoring, measurement, and improvement. In *International Conference on Conceptual Modeling* (pp. 339-

- 351). Berlin, Heidelberg: Springer.
- Dejaeger, K., B. Hamers, J. Poelmans, & B. Baesens. 2010. A novel approach to the evaluation and improvement of data quality in the financial sector. In International Conference on Information Quality (ICIQ). Cambridge, USA.
- English, L. P. 1999. *Improving data warehouse and business information quality: methods for reducing costs and increasing profits*. (Vol. 1). New York: Wiley.
- Fadahunsi, K. P., J. T. Akinlua, S. O'Connor, P. A. Wark, J. Gallagher, C. Carroll, and J. O'Donoghue. 2019. Protocol for a systematic review and qualitative synthesis of information quality frameworks in eHealth. *BMJ open* 9 (3): e024722.
- Falorsi, P. D., S. Pallara, A. Pavone, A. Alessandroni, E. Massella, & M. Scannapieco. 2003. Improving the quality of toponymic data in the Italian public administration. In Proceedings of the ICDT (Vol. 3).
- Ge, M., & M. Helfert. 2008. Data and information quality assessment in information manufacturing systems. In International Conference on Business Information Systems 380-389. Berlin, Heidelberg: Springer.
- _____, & D. Jannach. 2011. Information quality assessment: Validating measurement dimensions and processes. www.aisel.aisnet.org (accessed June 26, 2019).
- Günther, L. C., E. Colangelo, H. H. Wiendahl, & C. Bauer. 2019. Data quality assessment for improved decision-making: a methodology for small and medium-sized enterprises. *Procedia Manufacturing* 29: 583-591.
- Huh, Y. U., F. R. Keller, T. C. Redman, & A. R. Watkins. 1990. Data quality. *Information and software technology* 32 (8): 559-565.
- Irny, S. I., and A. A. Rose. 2005. Designing a Strategic Information Systems Planning Methodology for Malaysian Institutes of Higher Learning (isp-ipta). Universiti of Teknologi Malaysia.
- Jeusfeld, M. A., C. Quix, & M. Jarke. 1998, November. Design and analysis of quality information for data warehouses. In International Conference on Conceptual Modeling (pp. 349-362). Berlin, Heidelberg: Springer.
- Kilkenny, M. F. and K. M. Robinson. 2018. Data quality: "Garbage in—garbage out". *Health Information Management Journal* 47 (3): 103-105.
- Lee, Y. W., D. M. Strong, B. K. Kahn, & R. Y. Wang. 2002. AIMQ: a methodology for information quality assessment. *Information & management* 40 (2): 133-146.
- Liu, Z., Q. Chen, & L. Cai. 2018. Application of Requirement-oriented Data Quality Evaluation Method. In 2018 19th IEEE/ ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD) (pp. 407-412). IEEE. Busan, Korea.
- Loshin, D. 2001. Enterprise knowledge management: The data quality approach. California: Morgan Kaufmann press.
- _____. 2010. The practitioner's guide to data quality improvement. California: Morgan Kaufmann press.
- Man, Y., L. Wei, H. Gang, & G. Juntao. 2010. A novel data quality controlling and assessing model based on rules. In 2010 Third International Symposium on Electronic Commerce and Security (pp. 29-32). IEEE. Nanchang, China.
- McGilvray, D. 2008. *Executing data quality projects: Ten steps to quality data and trusted information (TM)*. California: Morgan Kaufmann press.
- Noblit, G. W. and R. D. Hare. 1988. *Meta-ethnography: Synthesizing qualitative studies*. (Vol. 11). California: Sage.
- Oxford dictionaries. 2020. Retrieved from oxford dictionaries: <https://www.oxfordlearnersdictionaries.com/definition/americanenglish/methodology> (accessed Feb. 18, 2019).
- Redman, T. C. 2013. Data's credibility problem. *Harvard Business Review* 91 (12): 84-88.

- _____, and A. Blanton. 1996. *Data quality for the information age*. Norwood: Artech House.
- Sandelowski, M., and J. Barroso. 2006. *Handbook for synthesizing qualitative research*. New York: Springer publishing company.
- Sandelowski M., S. Docherty, and C. Emden. 1997 Qualitative Metasynthesis: Issues and Techniques. *Research in Nursing and Health* 20: 365–371.
- Shankaranarayan, G., M. Ziad, & R. Y. Wang. 2003. Managing data quality in dynamic decision environments: An information product approach. *Journal of Database Management (JDM)* 14 (4): 14-32.
- Shirish, T. S. 2013. *Research methodology in education*. Carolina: Lulu publishing company.
- Siau, K., and Y. Long. 2005. Synthesizing e-government stage models—a meta-synthesis based on meta-ethnography approach. *Industrial Management & Data Systems* 105 (4): 443-458.
- Sidi, F., P. H. S. Panahy, L. S. Affendey, M. A. Jabar, H. Ibrahim, and A. Mustapha. 2012. March. Data quality: A survey of data quality dimensions. In 2012 International Conference on Information Retrieval & Knowledge Management (pp. 300-304). IEEE. Kuala Lumpur.
- Silvola, R., J. Harkonen, O. Vilppola, H. Kropsu-Vehkaperä, & H. Haapasalo. 2016. Data quality assessment and improvement. *International Journal of Business Information Systems* 22 (1): 62-81.
- Su, Y., & Z. Jin. 2007. Assuring Information Quality in Knowledge intensive business services. In 2007 International Conference on Wireless Communications, Networking and Mobile Computing (pp. 3243-3246). IEEE. Shanghai, China.
- Tayi, G. K., & D. P. Ballou. 1998. Examining data quality. *Communications of the ACM* 41 (2): 54-57.
- Vaziri, R., M. Mohsenzadeh, & J. Habibi. 2019. Measuring data quality with weighted metrics. *Total Quality Management & Business Excellence* 30 (5-6): 708-720.
- Wang, R. Y. 1998. A product perspective on total data quality management. *Communications of the ACM* 41 (2): 58-65.
- _____, V. C. Storey, and C. P. Firth. 1995. A framework for analysis of data quality research. *IEEE transactions on knowledge and data engineering* 7 (4): 623-640.
- Weiskopf, N. G., and C. Weng. 2013. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *Journal of the American Medical Informatics Association* 20 (1): 144-151.
- Wijnhoven, F., R. Boelens, R. Middel, & K. Louissen. 2007. Total data quality management: A study of bridging rigor and relevance. European conference on information systems (ECIS). Illinois, USA.
- Woodall, P., A. Borek, and A. K. Parlikad. 2013. Data quality assessment: the hybrid approach. *Information & management* 50 (7): 369-382.
- Worthington, J. C., & G. Brilis. 2000. How good are my data?: Information quality assessment methodology. *Quality Assurance* 8 (3-4): 245-260.
- Zaveri, A., A. Rula, A. Maurino, R. Pietrobon, J. Lehmann, S. Auer, and P. Hitzler. 2013. Quality assessment methodologies for linked open data. Submitted to *Semantic Web Journal* 1 (1): 1-5.

جدول پیوست. مقالات نهایی مورد تحلیل در مطالعه

ردیف	سال	نویسندگان	عنوان مقاله	کد مقاله
۱	۱۹۹۰	Huh et al.	Data quality	۱۳
۲	۱۹۹۵	Redman	Improve Data Quality For Competitive Advantage	۱۸
۳	۱۹۹۸	Wang	A Product Perspective on Total Data Quality Management	۵
۴	۲۰۰۰	Worthington and Brilis	How good are my data? Information quality assessment methodology.	۱۷
۵	۲۰۰۲	Lee et al.	AIMQ: a methodology for information quality assessment	۶
۶	۲۰۰۳	Shankaranarayan et al.	Managing data quality in dynamic decision environments: An information product approach	۲۱
۷	۲۰۰۳	Caballero and Piattini	CALDEA: A Data Quality Model Based on Maturity Levels	۱۱
۸	۲۰۰۴	Bringel, Caetano & Tribolet	Business Process modeling towards data quality assurance an organizational engineering approach	۱۰
۹	۲۰۰۵	Caballero et al.	Improving information quality management using CALDEA and EVAMECAL	۱۹
۱۰	۲۰۰۶	Cappiello, Ficiaro & Pernici	HIQM: A methodology for information quality monitoring, measurement, and improvement	۱۶
۱۱	۲۰۰۷	Batini et al.	A framework and a methodology for data quality assessment and monitoring	۲
۱۲	۲۰۰۷	Bugajski and Grossman	An alert management approach to data quality: Lessons learned from the visa data authority program	۲۵
۱۳	۲۰۰۷	Wijnhoven et al.	Total data quality management: A study of bridging rigor and relevance	۲۴
۱۴	۲۰۰۷	Su and Jin	Assuring information quality in knowledge intensive business services	۹
۱۵	۲۰۰۸	Batini et al.	A comprehensive data quality methodology for web and structured data	۱
۱۶	۲۰۰۸	Caballero et al.	IQM3: Information quality management maturity model	۲۷
۱۷	۲۰۰۸	Ge and Helfert	Data and information quality assessment in information manufacturing systems	۱۲
۱۸	۲۰۰۸	Caballero et al.	MMPRO: A methodology based on ISO/IEC 15939 to draw up data quality measurement processes	۲۳

ردیف	سال	نویسندگان	عنوان مقاله	کد مقاله
۱۹	۲۰۱۰	Dejaeger et al.	A Novel Approach to the Evaluation and Improvement of Data Quality in the Financial Sector	۴
۲۰	۲۰۱۱	Ge, Helfert & Jannach	Information quality assessment: Validating measurement dimensions and processes	۲۰
۲۱	۲۰۱۰	Man et al.	A noval data quality controlling and assessing model based on rules	۳
۲۲	۲۰۱۵	Bargh, van Dijk & Choenni	dynamic data quality management framework using issue tracking systems	۱۵
۲۳	۲۰۱۶	Silvola et al.	Data quality assessment and improvement	۱۴
۲۴	۲۰۱۶	Aljumaili, Karim & Tretten	Metadata-based data quality assessment	۲۲
۲۵	۲۰۱۸	Liu, Chen, & Cai	Application of requirement-oriented data quality evaluation method	۷
۲۶	۲۰۱۱	Berti-Équille et al.	Assessment and analysis of information quality: A multidimensional model and case studies	۸
۲۷	۲۰۱۹	Günther et al.	Data quality assessment for improved decision-making: A methodology for small and medium-sized enterprises	۲۶

بابک سهرابی

دارای مدرک تحصیلی دکتری در رشته علوم مدیریت از دانشگاه لنکستر انگلستان است. ایشان هم‌اکنون استاد گروه مدیریت فناوری اطلاعات دانشگاه تهران و سردبیر مجله مدیریت فناوری اطلاعات است. هوش مصنوعی، مدیریت دانش، تجارت الکترونیکی، یادگیری الکترونیکی، سیستم‌های پشتیبانی تصمیم‌گیری و تأثیر سازمانی فناوری اطلاعات، کلان‌داده و داده‌کاوی از جمله علایق پژوهشی وی است.



حمیدرضا یزدانی

متولد سال ۱۳۵۹ و دارای مدرک تحصیلی دکتری در رشته مدیریت منابع انسانی از دانشگاه تهران است. ایشان هم‌اکنون استادیار گروه بازرگانی و کسب‌وکار پردیس فارابی دانشگاه تهران است. سیستم‌های اطلاعات مدیریت، منابع انسانی، و علوم شناختی از جمله علایق پژوهشی وی است.



محمدجواد ارشادی

دانش‌آموخته دکتری تخصصی دانشگاه علم و صنعت در رشته مهندسی صنایع در سال ۱۳۹۳ است و مدارک کارشناسی ارشد و کارشناسی خود را در همین رشته به‌ترتیب از دانشگاه‌های صنعتی شریف و صنعتی اصفهان در سال‌های ۱۳۸۷ و ۱۳۸۴ دریافت داشته است. وی کار علمی خود را از سال ۱۳۹۴ با عضویت در هیئت علمی پژوهشگاه علوم و فناوری اطلاعات ایران آغاز کرد و هم‌اکنون استادیار گروه پژوهشی مدیریت فناوری اطلاعات است. بیشتر فعالیت‌های او در این پژوهشگاه پیرامون کاربرد مدیریت و مهندسی کیفیت در علوم و فناوری اطلاعات بوده است.



زمینه‌های پژوهشی مورد علاقه وی را کنترل کیفیت آماری، مدیریت کیفیت جامع، بازمهندسی فرایندهای کسب‌وکار، بهینه‌سازی، الگوریتم‌های فراابتکاری، تجزیه و تحلیل سیستم‌ها و داده‌کاوی تشکیل می‌دهند.

نام پدیدآور

سوده دوروش، متولد سال ۱۳۶۵ و دانشجوی دکتری در رشته مدیریت فناوری اطلاعات دانشگاه تهران است. هوشمندی کسب‌وکار، داده‌کاوی و کیفیت داده از جمله علایق پژوهشی وی است.

