



Quality of Students Register Statistics: A Case Study University of Kurdistan

Savareh Rashidi¹, Bakhtiar Javaheri^{2*}, Naser Shirbagi³

¹PhD Student in Higher Education Development Planning, Department of Education, Faculty of Humanities and Social Sciences, University of Kurdistan, Sanandaj, Iran.

²Assistant Professor, Department of Economics, Faculty of Humanities and Social Sciences, University of Kurdistan, Sanandaj, Iran.

³Professor, Department of Education, Faculty of Humanities and Social Sciences, University of Kurdistan, Sanandaj, Iran.

ARTICLE INFO

Article Type:

Original Research

Received: 06.15.2021

Revised: 09.21.2021

Accepted: 10.02.2021

Keyword:

Data Quality
Real world System
Information System
Registration Statistics

*Corresponding Author:

Bakhtiar Javaheri

Email: bjavaheri@uok.ac.ir

ABSTRACT

Nowadays, majority of organizations rely to a large extent on information systems in conducting their affairs. Consequently, the quality of information obtained from these systems has a crucial effect on the work process of organizations. Low-quality data results in incurring high costs in organizations leading to lower efficiency and productivity. Improving data quality is a suitable and cost-effective solution to this problem. To improve the information quality, one must consider improving the design and performance of these information systems, and these demands knowing the sense and concept of data quality as a product, knowing methods of measuring it, and taking measures in properly designing and operating the system. Consequently, this study aimed to evaluate the quality of student registration data of University of Kurdistan in general and also in individual work areas of the registration statistics system. For this purpose, a combination of quantitative-qualitative and exploratory research methods were applied. In the qualitative part of the research, employing a library research method, theoretical concepts were presented, the data quality indicators were identified and the most important indicators selected. Accordingly, the "registration statistics quality evaluation checklist" was developed using expert opinion. In the quantitative part of the research, Golestan Educational System was monitored as a field study, and the checklist was completed and analyzed. The results indicated that students' registration statistics quality was 79 percent. The quality of statistics in identifying, registering, transferring and presenting areas were 76, 85, 92, 54, and 63 percent, respectively.



EXTENDED ABSTRACT

Introduction

In general, data quality in organizations and particularly in educational organizations is one of the pillars of decision-making and even educational policy. incorrect data can result in high cost to organizations and companies, causing the loss of customers and making it impossible or at least difficult to implement new strategies in the company or organization. Today, most universities, in addition to their main activities such as teaching and research, also focus on transferring university technologies to industry and commercializing research results, and such a process requires reliable data in higher education. University of Kurdistan was established as one of Iran's comprehensive universities in 1991, which in 2020 had approximately 12,000 students, 380 faculty members, and 260 employees. Golestan is the educational platform used by University of Kurdistan and many universities in the country. Golestan is an information system that is used to handle all matters related to admission, registration, course selection, grading, graduation and other student- related educational issues. Considering the fact that the authors of the present paper have been closely involved with the production of registration data, its quality has always been their concern, and they have sought to scrutinize the efficiency of the Golestan and quality of data it has generated at the University of Kurdistan. There are different frameworks to examine data quality, including practitioners and participants' opinion, the theory of semiotics, the theory of quality of services and products (data-oriented), and ontological principles (plan-oriented). Research literature shows that the concept of data quality depends on the actual use of data and thus the reason that design-oriented framework was used here.

Methodology

Considering the ontological nature of the research, a mixed-method research approach and an exploratory design including the following steps were used. For conceptual clarity and definition, in the qualitative part and in the first step, the current literature (published papers and books) from 1996 to 2021 was reviewed to examine the theoretical concepts such as data quality and its dimensions, various perspectives on data quality, information systems, register data system and its fields and their relationship with data quality. In the second step, data quality indicators were identified and the most important ones selected for this research. In the third step, based on the results of the previous steps, a data collection tool titled "Register Statistics Quality Assessment Checklist" was compiled using the experts' opinions. Due to the expert-oriented nature of this method, to check the validity and reliability of the checklist, we studied the opinions of 20 experts and managers of Iran Statistics Center. The results showed that 90% of the participants confirmed the validity of the checklist. In the quantitative part and the fourth step, the educational system of Golestan was monitored using the developed tool to examine the quality of register data on students.

Results and discussion

The students' statistics produced by the Golestan educational system was evaluated using our checklist and the calculated score for University of Kurdistan was 520, which is approximately 79% of the total score. The findings showed that in the field of identification, the quality of students' register statistics was approximately 76%, which can be considered a good score. The quality of student register statistics in the field of register was approximately 85%, which meant that the Golestan educational system was able to adequately cover a good percentage of the activities in this field. In the field of transfer, the quality of student register statistics was approximately 92%, which can be considered an excellent score and indicates the efficiency of the Golestan educational system to cover a high percentage of transactions required in the field of transfer. For the processing area, the quality of students' register statistics was approximately 54%, which might be considered a poor score, and it conveys the idea that Golestan's educational system failed to perform the necessary processing of events to produce management reports, and finally, in the area of providing the quality of students' register statistics, the score was 63% indicating that Golestan was able to partially provide the required managerial data as form, format and content.

Conclusion

In the present paper, after presenting the definitions related to register data, register statistics, register statistics system and information systems, it was determined that the register statistics system can be considered an information system. Accordingly, data quality was described with a plan-oriented approach. Then, a check list tool consisting of 12 dimensions was developed to examine the quality of student statistics in Golestan. Finally, the quality of student statistics was analyzed according to the areas of register statistics and quality dimensions. The quality of student register statistics in the identification, register, transfer, processing, and presentation fields were appropriate, good, excellent, poor, and medium respectively. It seems that all organizations seek to improve the quality of their data, although it should be noted that in some situations poor data quality is desired and useful. Poor data quality can be beneficial, for example, to prevent an adversary from learning about real-world situations of issues such as national defense and security, or to mislead business and economic competitors. Thus, identifying sources of defects is necessary to produce low quality data. Of course, this should be carried out consciously and in parallel with the main information system. Therefore, it is suggested that the University of Kurdistan and other higher education centers in the country should measure the quality of their own statistics in order to check the quality of other statistics produced by Golestan. In addition, it is recommended that Golestan should be redesigned, and in doing so, using the research findings, try to block unused fields, create control mechanisms at different stages of data generation, consider methods and tools to check the results of data processing and embed graphical reporting facilities.



دانشگاه تربیت مدرس
مجله حرفه‌فنی

کارافن

فصلنامه علمی دانشگاه فنی و حرفه‌ای

ویژه‌نامه ۱۴۰۱، دوره ۱۹، ۲۷۸-۲۵۵

آدرس نشریه: <https://karafan.tvu.ac.ir/>

doi: 10.48301/KSSA.2021.287020.1539



کیفیت آمار ثبتي دانشجویان: مطالعه موردی دانشگاه کردستان

سواره رشیدی^۱، بختیار جواهری^{۲*}، ناصر شیربگی^۳

- ۱- دانشجوی دکتری برنامه‌ریزی توسعه آموزش عالی، گروه علوم تربیتی، دانشکده علوم انسانی و اجتماعی، دانشگاه کردستان، سنندج، ایران.
- ۲- استادیار، گروه علوم اقتصادی، دانشکده علوم انسانی و اجتماعی، دانشگاه کردستان، سنندج، ایران.
- ۳- استاد، گروه علوم تربیتی، دانشکده علوم انسانی و اجتماعی، دانشگاه کردستان، سنندج، ایران.

چکیده

اطلاعات مقاله

امروزه بیشتر سازمان‌ها برای انجام امورشان تا حدود زیادی به سیستم‌های اطلاعاتی وابسته هستند پس کیفیت اطلاعات به‌دست‌آمده از این سیستم‌ها بر روند کاری سازمان اثر تعیین‌کننده‌ای دارد. داده‌های بی‌کیفیت هزینه‌های زیادی بر سازمان‌ها تحمیل می‌کنند که باعث کاهش کارایی و بهره‌وری می‌گردد. بهبود کیفیت داده راهکاری مناسب و با هزینه‌ی بهینه برای رفع این مشکل است. برای بهبود کیفیت اطلاعات باید به فکر بهبود طراحی و عملکرد این سیستم‌های اطلاعاتی بود و این کار مستلزم این است که معنی و مفهوم کیفیت داده به‌عنوان یک محصول شناخته شود، راه‌های اندازه‌گیری آن را دانست و در جهت بهبود کیفیت داده به طراحی مناسب و عملکرد صحیح سیستم اهتمام ورزید. بر این اساس هدف این تحقیق سنجش کیفیت آمار ثبتي دانشجویان دانشگاه کردستان به‌صورت کلی و همچنین به تفکیک حوزه‌های کاری نظام آمار ثبتي است که به این منظور از رویکرد تحقیق آمیخته (کیفی- کمی) و شیوه اکتشافی استفاده شده است. در بخش کیفی مفاهیم نظری از طریق مطالعات کتابخانه‌ای ارائه، شاخص‌های کیفیت داده شناسایی و مهم‌ترین آنها انتخاب‌شده است. لذا «چک‌لیست ارزیابی کیفیت آمار ثبتي» با استفاده از نظرات خبرگان تدوین‌شده است. در بخش کمی سامانه آموزشی گلستان به‌صورت میدانی مورد پایش قرار گرفته، چک‌لیست تکمیل و تحلیل شده است. نتایج نشان می‌دهد کیفیت آمار ثبتي دانشجویان به میزان ۷۹ درصد بوده که در حوزه‌های شناسایی، ثبت، انتقال، پردازش و ارائه به ترتیب ۷۶، ۸۵، ۹۲، ۵۴ و ۶۳ درصد بوده است.

نوع مقاله: مقاله پژوهشی

دریافت مقاله: ۱۴۰۰/۰۳/۲۵

بازنگری مقاله: ۱۴۰۰/۰۶/۳۰

پذیرش مقاله: ۱۴۰۰/۰۷/۱۰

کلید واژگان:

کیفیت داده
سیستم جهان واقعی
سیستم اطلاعاتی
آمار ثبتي

*نویسنده مسئول: بختیار جواهری

پست الکترونیکی:

b.javaheri@uok.ac.ir



©2022 the authors. Published by Technical and Vocational University, Tehran, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution-Noncommercial 4.0 International (CC BY-NC License) (<https://creativecommons.org/licenses/by-nc/4.0/>)

شاپای الکترونیکی: ۲۵۳۸-۴۴۳۰

شاپای چاپی: ۲۳۸۲-۹۷۹۶

مقدمه

کیفیت داده^۱ به‌طور اعم در سازمان‌ها و به‌طور اخص در سازمان‌های آموزشی یکی از ارکان تصمیم‌سازی و حتی سیاست‌گذاری آموزشی است. «آنچه در جهان هستی اتفاق می‌افتد، چه انسان دخالتی در آن داشته باشد و چه نداشته باشد، سبب تولید داده می‌شود بنابراین جهان هستی به‌تناوب در حال تولید داده است» (کیانی و رشیدی، ۲۰۰۹). سازمان‌های آموزشی هم به‌عنوان بخشی از جهان هستی از امر تولید داده مستثنی نیستند. بدیهی است که داده باکیفیت پایین اثرات سوئی بر کارایی یک سازمان خواهد داشت. مطالعات و پژوهش‌ها نشان داده، سازمان‌ها اغلب سیستم‌هایی در اختیار دارند که داده‌های باکیفیت بالا ارائه نمی‌دهند (واند و وانگ، ۱۹۹۶)^۲. به‌عنوان نمونه IBM هزینه داده‌های بی‌کیفیت در سال ۲۰۱۶ را فقط برای ایالات‌متحده آمریکا معادل ۳۰۱ تریلیون دلار برآورد کرده است (ردمن، ۲۰۱۶)^۳. داده‌ها در بیشتر فعالیت‌های سازمان‌ها مورداستفاده قرار می‌گیرند و اساس تصمیم‌گیری‌های عملیاتی و راهبردی می‌باشند. بنابراین داده‌های بی‌کیفیت به‌طور قابل توجهی تأثیر منفی بر کارایی سازمان خواهد داشت، درحالی‌که داده‌های باکیفیت بالا برای موفقیت سازمان مهم و ضروری هستند. داده‌های باکیفیت پایین هزینه‌های زیادی بر سازمان تحمیل می‌کند، برخی از کارشناسان اقتصاد و صنعت مانند گروه گارتنر^۴، پرایس واترهاوس کوپرز^۵ و موسسه دیتا ویرهاوس^۶ مطالعاتی در مورد کیفیت داده انجام داده‌اند (هاوگ و همکاران، ۲۰۱۱)^۷. اسلون^۸ به نقل از ردمن بیان می‌کند که خطاهای موجود در داده‌ها می‌تواند میلیون‌ها دلار برای شرکت هزینه ایجاد کند، باعث از دست دادن مشتریان شود و اجرای راهبردهای جدید شرکت را غیرممکن یا حداقل دشوار سازد (اسلون، ۲۰۰۶)^۸. در شکل ۱ نشان داده می‌شود که بین هزینه‌های یک سازمان و کیفیت داده‌های آن رابطه‌ای وجود دارد به‌طوری‌که هزینه‌های ناشی از داده‌های بی‌کیفیت به‌طوری تصاعدی افزایش می‌یابد ولی هزینه‌های ناشی از افزایش کیفیت به‌طور خطی افزایش می‌یابد و مقدار بهینه کل هزینه‌ها کمترین مقدار آن است (هاوگ و همکاران، ۲۰۱۱).

¹ Data Quality

² Wand & Wang

³ Redman

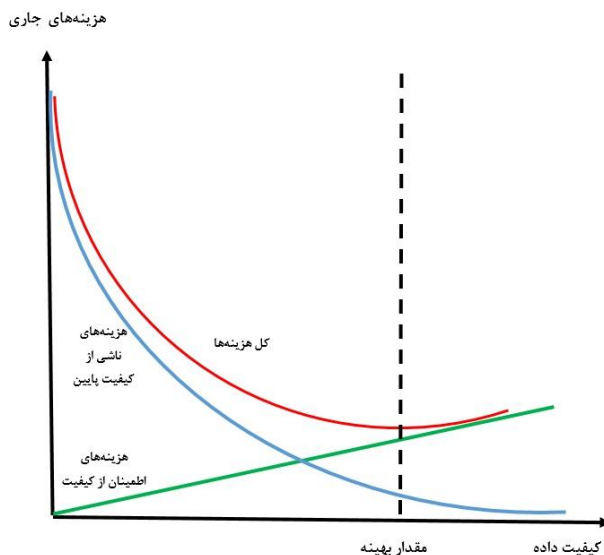
⁴ Gartner Group

⁵ Price Waterhouse Coopers

⁶ Data Warehousing Institute

⁷ Haug

⁸ Sline



شکل ۱. کل هزینه‌های ناشی از کیفیت داده‌های سازمان (هاوگ و همکاران، ۲۰۱۱).

ملاحظه می‌شود که بهبود کیفیت داده‌های سازمان علاوه بر تأثیر مثبت بر روی تصمیم‌گیری و سیاست‌گذاری و ارتقا عملکرد، باعث کاهش کل هزینه‌های سازمان نیز می‌شود. تحقیقات در مورد کیفیت داده‌ها با شناسایی ابعاد کیفیت شروع شد و از آن زمان به بعد پژوهشگران زیادی با دیدگاه‌های مختلف به موضوع کیفیت داده پرداخته‌اند. ابتدا داده به‌عنوان یک محصول مدنظر بوده و سپس از دیدگاه مدیریت کیفیت داده فراگیر، مدل‌سازی ساخت داده و اندازه‌گیری ابعاد کیفیت با استفاده از فراداده‌ها به موضوع کیفیت داده پرداخته شده است. پژوهشگران حوزه‌های مختلفی مانند آمار، علوم کامپیوتر، روانشناسی شناختی، مدیریت و اقتصادسنجی در زمینه کیفیت داده و اطلاعات به مطالعه پرداخته‌اند (شانکارانارایانا و بلیک، ۲۰۱۷)^۱. اسلون به نقل از ردمن پایگاه داده را به یک دریاچه تشبیه می‌کند و می‌گوید برای اطمینان از تمیز بودن دریاچه، ابتدا باید منابع آلوده‌کننده را از بین برد و سپس به تمیز کردن دریاچه پرداخت (اسلون، ۲۰۰۶).

یکی از مشکلات مهم در کشور برای برنامه‌ریزی و تصمیم‌گیری در حوزه‌های مختلف از قبیل اقتصادی، اجتماعی، آموزشی و فرهنگی کمبود داده و آمار باکیفیت است. غفاری به نقل از پولاک بیان می‌کند چنانچه داده‌ها از سطح کیفیت پایینی برخوردار باشند، سازمان‌ها در سه سطح عملیاتی، فنی و راهبردی با مشکل مواجه خواهند شد. داده‌ها در هر سازمان حکم آب را دارند و برای بقای هر سازمانی ضروری هستند. همان‌طور که بررسی کیفیت آب دارای اهمیت است و چنانچه آب موردنظر استانداردهای لازم را نداشته باشد می‌تواند جان انسان‌ها را به خطر اندازد، سازمان نیز در صورت عدم برخورداری از داده و اطلاعات باکیفیت نمی‌تواند به حیات خود ادامه دهد (غفاری و همکاران، ۲۰۱۸). تحقیق در مورد کیفیت داده‌ها به یک موضوع مهم پژوهشی در سیستم‌های اطلاعاتی و یک مجموعه دانش در بسیاری از رشته‌های دیگر تبدیل شده است. نحوه اطمینان از کیفیت داده‌ها، روش تجزیه و تحلیل آنها و استخراج اطلاعات و دانش نهفته در این داده‌ها، به یکی از موضوعات اصلی صنایع و دانشگاه‌ها تبدیل شده است و کیفیت پایین داده‌ها باعث کاهش بهره‌وری

¹ Shankaranarayanan & Blake

سازمانی و تصمیم‌گیری‌های اشتباه می‌شود (کای و ژو، ۲۰۱۵)^۱. از وقتی که کیفیت داده به‌عنوان یک زمینه پژوهشی مورد توجه قرار گرفت، پژوهش‌های زیادی درباره معانی و مفاهیم کیفیت داده و روش‌های اندازه‌گیری آن انجام شده است. این امر منجر به ظهور فلسفه‌ها و دیدگاه‌های مختلفی درباره کیفیت داده و توسعه روش‌های اندازه‌گیری آن شده است. باین حال سؤالی که همواره مطرح است این است که «آیا داده‌ها از کیفیت بالایی برخوردار هستند؟». روش‌های اندازه‌گیری کیفیت تنها در نحوه پاسخگویی به این سؤال متفاوت می‌باشد (تیمرمان و برونسلیر، ۲۰۱۹)^۲.

دانشگاه یکی از مهم‌ترین ارکان اساسی برای تحول و نوآوری در جامعه است و یکی از مهم‌ترین راه‌های تحقق اهداف متعالی چون استقلال، پیشرفت اجتماعی، اقتصادی و فرهنگی از طریق دانشگاه و فرایند آموزش عالی است (رحیمی و آقابابی، ۲۰۲۰). دانشگاه سازمانی است که نقش مهمی در جامعه برای انجام آموزش و پرورش ایفا می‌کند. در دهه‌های اخیر، بیشتر دانشگاه‌ها علاوه بر انجام فعالیت‌های اصلی‌شان مانند آموزش و پژوهش، بر روی انتقال فناوری‌های دانشگاهی به صنعت و تجاری‌سازی نتایج پژوهش‌ها نیز تمرکز کرده‌اند (الماسی فرد و همکاران، ۲۰۲۱) و چنین روندی نیاز به داده‌های معتبر در آموزش عالی دارد. دانشگاه کردستان به‌عنوان یکی از دانشگاه‌های جامع کشور در سال ۱۳۷۰ تأسیس شده است که به‌عنوان دانشگاه مادر در استان کردستان به انجام رسالت آموزش و پژوهش مشغول است. این دانشگاه در سال ۱۳۹۹ حدود ۱۲۰۰۰ دانشجو و ۳۸۰ هیئت‌علمی و ۲۶۰ کارمند داشته است. سیستم آموزشی دانشگاه سامانه گلستان می‌باشد که بسیاری از دانشگاه‌های کشور از خدمات این سامانه بهره می‌گیرند. سامانه آموزشی گلستان یک سیستم اطلاعاتی^۳ است که تمامی امور مربوط به پذیرش، ثبت‌نام، انتخاب واحد درسی، ثبت نمرات دروس، دانش‌آموختگی و سایر موارد مربوط به امور آموزشی دانشجویان به کمک آن صورت می‌گیرد. با توجه به این‌که نویسندگان این مقاله به‌عنوان شاغلین در دانشگاه کردستان از نزدیک با تولید آمار ثبتي درگیر بوده‌اند، همواره کیفیت آمار ثبتي تولیدی حاصل از عملکرد سامانه آموزشی گلستان دغدغه ایشان بوده است و به دنبال بررسی عملکرد سامانه آموزشی گلستان و سنجش میزان کیفیت آمار ثبتي تولیدشده توسط این سامانه بوده‌اند. به همین خاطر دانشگاه کردستان و سامانه آموزشی آن انتخاب گردیده و هدف سنجش کیفیت آمار ثبتي دانشجویان و ارائه راهکاری جهت بهبود کیفیت این آمار بوده است.

مبانی نظری و پیشینه

چارچوب‌های مختلفی در مورد کیفیت داده وجود دارد که عبارت‌اند از: بر پایه نظرات شاغلین و شرکت‌کنندگان (وانگ و استرونگ، ۱۹۹۶)^۴، بر پایه نظریه نشانه‌شناسی (شانکس و دارکه، ۱۹۹۸)^۵، بر پایه نظریه کیفیت خدمات و محصولات (داده گرا) (کان و همکاران، ۱۹۹۷)^۶ و بر پایه اصول هستی‌شناسانه (طرح گرا) (واند و وانگ، ۱۹۹۶). این موضوع روشن است که مفهوم کیفیت داده به کاربرد حقیقی داده بستگی دارد. کیفیت داده به‌طور مستقل از فرایند تولید داده و زمینه‌های استفاده از داده نمی‌تواند بهبود یابد یعنی برای بهبود پایدار کیفیت داده، استفاده از فناوری‌های روز لازم اما کافی نیست. برای اطمینان از بهبود کیفیت داده به‌صورت بلندمدت، علاوه بر راه‌حل‌های داده محور باید به سیستم تولید داده نیز پرداخته شود (آباته و همکاران، ۱۹۹۸)^۷. برخی قوانین کلی در مورد کیفیت داده وجود دارد که از دید سیستمی می‌توان در مورد سیستم‌های اطلاعاتی استنباط کرد: داده‌هایی که استفاده نمی‌شوند نمی‌توانند برای مدت

¹ Cai & Zhu

² Timmerman & Bronselaer

³ Information System

⁴ Wang & Strong

⁵ Shanks & Darke

⁶ Kahn

⁷ Abate

طولانی صحیح باشند. کیفیت داده در یک سیستم اطلاعاتی تابعی از کاربرد داده است و نه جمع‌آوری آن. مشکلات کیفیت داده با افزایش سن سیستم بیشتر می‌شود. قوانین مربوط به کیفیت داده به‌طور یکسان در داده‌ها و فراداده‌ها اعمال شود (اور، ۱۹۹۸)^۱. کیفیت داده تولیدشده توسط یک سیستم اطلاعاتی به طراحی و عملکرد آن سیستم بستگی دارد. اما استفاده صحیح از داده خارج از کنترل طراح سیستم است. بنابراین ارائه تعریفی طرح‌گرا^۲ از کیفیت داده، که استفاده خواسته‌شده از اطلاعات را منعکس کند، لازم به نظر می‌رسد (وانگ و وانگ، ۱۹۹۶).

مبنای نظری: دیدگاه طرح‌گرا

بحث را در ابتدا با تمایز بین دید درونی و دید بیرونی یک سیستم اطلاعاتی آغاز می‌کنیم. دید بیرونی به استفاده و اثر یک سیستم اطلاعاتی توجه دارد. این دید اهداف و توجیهات سیستم و به‌کارگیری آن در سازمان را خطاب قرار می‌دهد. اما دید درونی، ساختار و ملزومات موردنیاز برای عملکرد را مورد خطاب قرار می‌دهد (وانگ و وبر، ۱۹۹۵)^۳. بنابراین تحلیل ما بر روی دید درونی متمرکز می‌باشد و به طراحی سیستم و تولید داده گرایش دارد. مدلی که در اینجا ارائه می‌شود بر چهار فرض استوار است: فرض ارائه^۴، فرض تفسیر^۵، فرض استنباط^۶ و فرض دید درونی^۷ که به‌طور مختصر به هر یک از این‌ها می‌پردازیم.

فرض ارائه: یک سیستم اطلاعاتی ارائه‌ای از یک سیستم جهان واقعی^۸ است همان‌طوری که توسط کاربران درک می‌شود. یک سیستم اطلاعاتی مدلی کوچک و زیرمجموعه‌ای محدود از دنیای واقعی است (کنت، ۲۰۱۲)^۹.

فرض تفسیر: یک سیستم اطلاعاتی برای استفاده کاربرانی ایجاد می‌شود که دید آنها از دنیای واقعی در طراحی سیستم مدنظر قرار گیرد. این امر اساس تعریف ما از کاستی و ناکارایی داده است (وانگ و وانگ، ۱۹۹۶). یک داده ناکارا از عدم تطابق بین دید سیستم جهان واقعی که می‌تواند ناشی از سیستم اطلاعاتی باشد و دید به‌دست‌آمده از مشاهده مستقیم جهان واقعی، به دست می‌آید شکل ۲ را ملاحظه فرمایید. سیستم اطلاعاتی همانند یک آینه عمل می‌کند. هرگاه بین ناظر و شیء مانعی وجود داشته باشد به‌طوری‌که مشاهده شیء به‌صورت مستقیم امکان‌پذیر نباشد، برای مشاهده شیء می‌توان از آینه بهره گرفت (کیانی و رشیدی، ۲۰۰۹). شکل ۳ را ملاحظه فرمایید.

¹ Orr

² Design-Oriented

³ Wand & weber

⁴ Representation Assumption

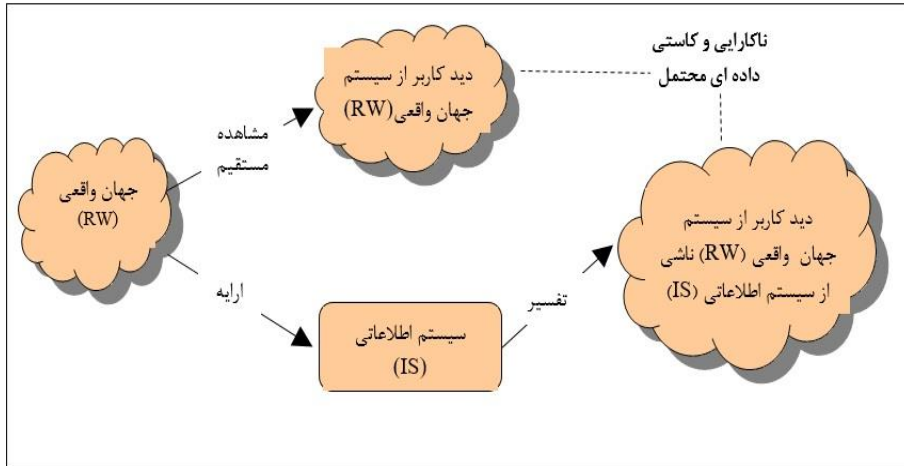
⁵ Interpretation Assumption

⁶ Inference Assumption

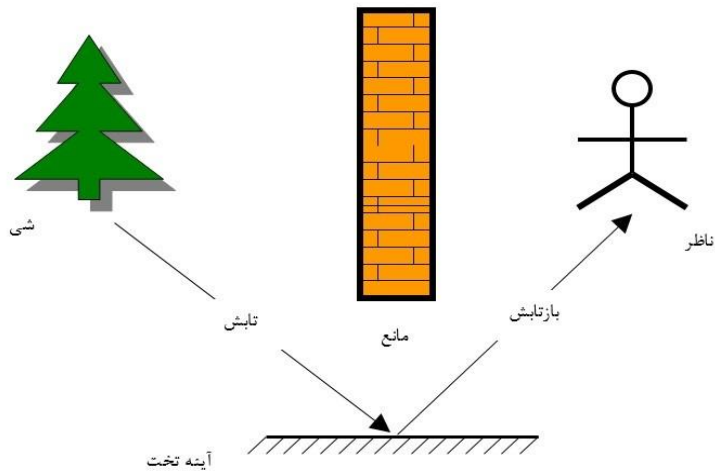
⁷ Internal View Assumption

⁸ Real World System

⁹ Kent



شکل ۲. ناکارایی و کاستی داده‌ای محتمل در مدل کیفیت داده (واند و وانگ، ۱۹۹۶).



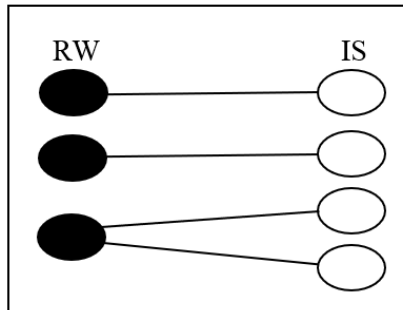
شکل ۳. تشابه آینه و سیستم اطلاعاتی (کیانی و رشیدی، ۲۰۰۹).

فرض استنباط: سیستم اطلاعاتی می‌تواند یک ارائه ملموس ایجاد کند که کاربر به کمک آن قادر خواهد بود دیدی از سیستم جهان واقعی که در سیستم اطلاعاتی وجود دارد، استنباط کند (واند و وانگ، ۱۹۹۶). در مثال آینه، ناظر با دانستن نوع آینه تصویری را که می‌بیند تفسیر می‌کند. یعنی اگر آینه تخت باشد به این نتیجه می‌رسد که شیء به همان اندازه است که می‌بیند و در همان فاصله است و اگر آینه کوژ باشد استنباطش این است که شیء نزدیک‌تر و بزرگ‌تر از آن چیزی است که می‌بیند و در مورد آینه کاو استنباط دیگری وجود دارد (کیانی و رشیدی، ۲۰۰۹).

فرض دید درونی: بر طبق این فرض ابعاد کیفیت داده به نیازهای کاربر ارتباط دارد. کار ما در آغاز شناسایی ضوابطی است که بر طبق آن سیستم جهان واقعی به‌طور مناسبی به‌وسیله یک سیستم اطلاعاتی به نمایش گذاشته

می‌شود. بر این اساس، به شناسایی کاستی‌هایی می‌پردازیم که ممکن است در فرایند طراحی سیستم و تولید داده بروز دهد. این کاستی‌ها برای تعریف ذاتی ابعاد کیفیت داده مورد استفاده قرار می‌گیرند.

ارائه مناسب^۱: فرض کنید RW بیانگر فضای حالت سیستم جهان واقعی باشد و IS یک ارائه سیستم اطلاعاتی از جهان واقعی است شکل ۴. اول، هر حالت مجاز سیستم جهان واقعی باید به حداقل یک حالت مجاز از سیستم اطلاعاتی نگاشت شود (حالت جهان واقعی می‌تواند به چندین حالت سیستم اطلاعاتی نگاشت شود). دوم، به‌طور اصولی، این امر امکان‌پذیر می‌باشد که یک حالت سیستم اطلاعاتی به حالت جهان واقعی «درست» برگشت داده شود.



شکل ۴. ارائه مناسب (واند و وانگ، ۱۹۹۶).

یک سیستم جهان واقعی در صورتی به‌خوبی ارائه می‌شود که:

- ۱- یک نگاشت جامع برای ارائه وجود داشته باشد.
 - ۲- هیچ دو حالتی در RW به یک حالت یکسان در IS نگاشت نشود.
- در کل می‌توان دو نوع کاستی در نظر گرفت که منشأ افت کیفیت یا تولید داده بی‌کیفیت می‌گردد که عبارت‌اند از: کاستی‌های طراحی و کاستی‌های عملیات.

چارچوب مفهومی

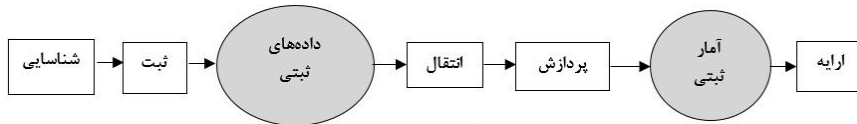
داده‌های ثبتی^۲، داده‌هایی هستند که در حین انجام فعالیت‌های عینی، واقعی و جاری یک سازمان و با استفاده از برگه‌ها، اسناد و مدارک و فایل‌های رایانه‌ای جاری در داخل سازمان ثبت می‌شوند. آمار ثبتی^۳، آمارهایی هستند که از پردازش (طبقه‌بندی، مرتب کردن، محاسبات، تلخیص و ...) داده‌های ثبتی حاصل می‌شوند. نظام آمار ثبتی، نظامی است که داده‌های ثبتی مناسبی در درون خود ثبت کرده و این داده‌های ثبت‌شده در حین عملیات جاری و روزمره یک سازمان را به آمار ثبتی موردنیاز کاربران تبدیل کرده و آنها را به نحو مناسب در اختیار آنان قرار می‌دهد (نورمحمدی و نواب پور، ۲۰۰۵). باید توجه داشت که منظور از داده ثبتی مناسب داده‌ای است که از تعریف مشخص، روشن و بدون ابهام برخوردار باشد، امکان ثبت داده‌ها در پایین‌ترین نقطه فراهم باشد، ثبت داده‌ها به‌هنگام وقوع عمل امکان‌پذیر باشد، فراگرد منطقی وجود داشته باشد و افراد مسئول و پاسخگو در سرتاسر فراگرد وجود داشته باشند.

¹ Proper Representation

² Register Data

³ Register Statistics

حوزه‌های کاری نظام آمار ثبتی



شکل ۵. حوزه‌های کاری نظام آمار ثبتی (نورمحمدی و نواب پور، ۲۰۰۵).

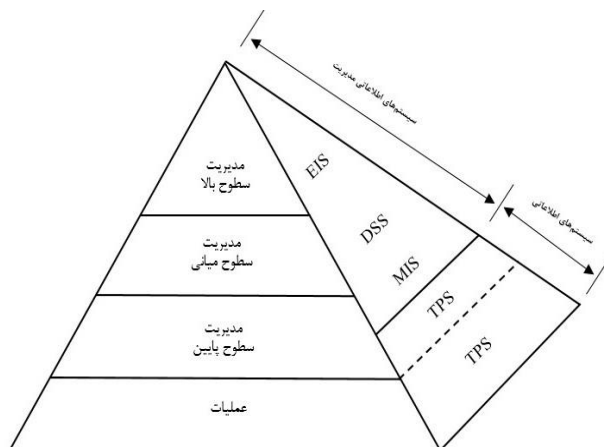
با توجه به ماهیت آمارهای ثبتی و ۵ حوزه مطرح شده در فوق و ضرورت بررسی کیفیت داده‌ها در این حوزه‌ها تبیین موضوعات مرتبط با آمارهای ثبتی اهمیت بیشتری پیدا می‌کند، لذا تعاریف مربوطه در ادامه آمده است. یک سیستم، مجموعه‌ای منظمی از عناصر به هم وابسته است که برای رسیدن به اهدافی مشترک باهم در تعامل مفید و مؤثر هستند. یک سیستم اطلاعاتی، سیستمی است که داده‌های خام را به‌عنوان درون داده می‌پذیرد و پس از پردازش، برون داده‌هایی به‌صورت اطلاعات عملیاتی و گزارش‌های مدیریتی تولید می‌کند. پس نظام آمار ثبتی، یک سیستم اطلاعاتی است که داده‌ها را درون سیستم ثبت کرده و پس از پردازش، برون داده‌هایی به‌صورت آمار ثبتی ارائه می‌کند. سیستم‌های اطلاعاتی انواع مختلفی دارند که عبارت‌اند از: سیستم پردازش تراکنش (TPS)^۱، سیستم اطلاعاتی مدیریت (MIS)^۲، سیستم اطلاعاتی پشتیبان تصمیم‌گیری (DSS)^۳ و سیستم اطلاعاتی مدیریت سطوح بالا (EIS)^۴. سازمان سیستمی است که از سازمان‌دهی و ترکیب منابع مختلفی از قبیل نیروی انسانی، سرمایه، مواد، تجهیزات، و سایر منابع به‌صورت یک مؤسسه پایدار و مؤثر اعم از تجاری و غیرتجاری به وجود می‌آید (ذاکری، ۲۰۰۷). همان‌طور که در شکل (۶) نشان داده شده است می‌توان سیستم‌های اطلاعاتی مختلف را به‌صورت یک هرم در نظر گرفت که سطوح عملیاتی در قاعده هرم و سطوح مدیریتی در رأس هرم قرار گرفته‌اند و این بیانگر این موضوع است که داده‌ها، اطلاعات و گزارش‌ها از طرف قاعده به سمت رأس حرکت کرده و مبنای تصمیم‌گیری و برنامه‌ریزی خواهد بود. پس اگر اطلاعات و گزارش‌های قاعده (TPS) دچار خدشه گردد این خدشه‌ها در سطوح بالاتر مضاعف گردیده و هنگامی که به رأس هرم می‌رسد می‌تواند از واقعیت خیلی به دور باشد و منجر به تصمیم‌گیری‌ها و برنامه‌ریزی‌های نادرست گردد.

¹ Transaction Processing System

² Management Information System

³ Decision Support System

⁴ Executive Information System



شکل ۶. انواع سیستم اطلاعاتی و سطوح مختلف مدیریت و عملیات (ذاکری، ۲۰۰۷).

می‌توان نظام آمار ثبتي در یک سازمان را به‌عنوان یک سیستم اطلاعاتی در نظر گرفت که بخش قابل توجهی از آن در سطوح عملیاتی (قاعده هرم) قرار می‌گیرد پس کیفیت داده‌های ثبتي تولیدشده در این نظام تأثیر بسزایی در کیفیت گزارش‌های مدیریتی سطوح بالا (رأس هرم) دارد. اینجاست که کیفیت داده‌های ثبتي مطرح می‌گردد و بایستی بر اساس سیستم‌های اطلاعاتی، تعریف کیفیت داده و ابعاد آن ارائه شود (کیانی و رشیدی، ۲۰۰۹).

مؤسسه پژوهش و برنامه‌ریزی آموزش عالی به‌عنوان متولی آمار آموزش عالی ملزم به گردآوری آمار و اطلاعات دانشجویان، دانش‌آموختگان، اساتید و کارکنان دانشگاه‌ها و مؤسسات آموزش عالی کشور و ارائه آن به سازمان برنامه و بودجه کشور است. این آمارها یکی از معیارهای تعیین سهم هر دانشگاه از بودجه عمومی کشور است. با توجه به این‌که بسیاری از دانشگاه‌ها و مؤسسات آموزش عالی کشور از سامانه گلستان استفاده می‌کنند اگر داده‌های تولیدشده در این سامانه از کیفیت لازم برخوردار نباشند کیفیت آمار و اطلاعات تجمیع شده از دانشگاه‌های سراسر کشور به‌صورت مضاعف کاهش می‌یابد و این موضوع ممکن است بر توزیع بودجه بین دانشگاه‌ها تأثیر سو بگذارد و عدالت آموزشی را تحت‌الشعاع قرار دهد. مسئله این است که کیفیت پایین داده‌ها می‌تواند باعث برنامه‌ریزی نادرست گردد به همین دلیل ارزیابی کیفیت آمارهای بخش آموزش عالی در جهت بهبود کیفیت آن برای برنامه‌ریزی و بودجه‌ریزی امری لازم و ضروری می‌باشد.

کیفیت داده

خیلی از افراد در موقعیت‌های مختلف عبارت «کیفیت» را به کار می‌برند اما همگی برداشت یکسانی از این عبارت ندارند. به‌عنوان مثال جوران و گریان^۱ کیفیت را مقبولیت در استفاده از کالا و خدمات دانسته‌اند یا طبق تعریف ایزو ۹۰۰۰ کیفیت عبارت است از: توانایی مجموعه‌ای کاملی از ویژگی‌های ذاتی محقق شده یک محصول، سیستم یا فرایند برای برآورده کردن نیازهای موجود (کیانی و رشیدی، ۲۰۰۹). می‌توان کیفیت را، به‌طور کلی، به‌عنوان «کلیت ویژگی‌های کالایی که توانایی آن در تأمین نیازهای اعلام‌شده یا ضمنی را تحت تأثیر قرار می‌دهد»، همچنین «تناسب برای استفاده»، «انطباق با موارد موردنیاز» یا «رضایت کاربر» تعریف کرد (باتینی و اسکناپیکو، ۲۰۱۶)^۲. تنوع معنی‌داری در مفهوم‌سازی کیفیت اطلاعات وجود دارد، باین‌حال، تعاریف ساختار کیفیت اطلاعات بسیار متفاوت است. این سازه به‌طور مداوم

^۱ Juran & Gryan

^۲ Batini & Scannapieco

به‌عنوان ماهیتی چندبعدی، هم در زمینه هوش تجاری و زمینه تحلیلی و هم در دامنه وسیع‌تر سیستم‌های اطلاعاتی مفهوم‌سازی شده است (تورس و سیدورووا، ۲۰۱۹)^۱. تجربه نشان داده است که در بحث کیفیت داده یک‌راه حل یگانه برای همه مشکلات نمی‌توان به دست آورد زیرا ارزیابی کیفیت داده تلاشی مداوم است که مستلزم شناخت ذهنی و عینی مفهوم کیفیت و ابعاد آن است. می‌توان با روش‌هایی، ارزیابی ذهنی و عینی کیفیت داده‌ها را باهم ترکیب کرد و از این راه به رویکردی رسید که به‌طور مؤثر در ارزیابی کیفیت داده استفاده شود (پپینو و همکاران، ۲۰۰۲)^۲. کیفیت اطلاعات یک نگرانی اصلی در علم است. هدف اصلی این است که مقدار زیادی داده باکیفیت بسیار بالا تولید شود. در واقع تمام ملاحظات دیگر فرعی هستند. به همین خاطر کیفیت داده یکی از بالاترین اولویت‌ها است (لوکیانکو و همکاران، ۲۰۲۰)^۳. کیفیت پایین داده می‌تواند تأثیر نامطلوبی بر سطوح عملیاتی، فنی و راهبردی سازمان‌ها داشته باشد این موارد شامل هزینه زیاد (۸-۱۲ درصد درآمد)، تصمیم‌گیری ضعیف و افزایش مشکلات در تدوین استراتژی است (واهیودی و همکاران، ۲۰۱۸)^۴.

پیشینه پژوهش

از منظر پژوهشی، کیفیت اطلاعات در زمینه‌های مختلف، از جمله آمار، مدیریت و علوم رایانه مورد توجه قرار گرفته است. آمارشناسان اولین گروهی بودند که در اواخر دهه ۱۹۶۰ با ارائه یک نظریه ریاضی برای در نظر گرفتن موارد تکراری در مجموعه داده‌های آماری برخی از مشکلات مربوط به کیفیت اطلاعات را بررسی کردند. این موضوع در آغاز دهه ۱۹۸۰ روی چگونگی کنترل سیستم‌های تولید اطلاعات به‌منظور کشف و از بین بردن مشکلات کیفیت اطلاعات توسط محققان مدیریت پیگیری شد. در آغاز دهه ۱۹۹۰ دانشمندان علوم رایانه‌ای در مورد مسئله تعریف، اندازه‌گیری و بهبود کیفیت اطلاعات الکترونیکی ذخیره‌شده در پایگاه‌های اطلاعاتی، انبارهای داده و سیستم‌های اطلاعاتی اقدام به بررسی کردند (باتینی و اسکاناپیکو، ۲۰۱۶).

در پژوهش «کیانی و رشیدی» ضمن ارائه تعریفی از کیفیت داده، شاخص‌های موردنظر برای سنجش کیفیت داده‌های ثبتي، مورد مطالعه و بررسی قرار گرفته و با مطالعه اسناد و مدارک موجود در این زمینه به تعریف ۱۱ بعد کیفیت داده پرداخته‌اند (کیانی و رشیدی، ۲۰۰۹).

در پژوهشی که «غفاری، آذر و شباک» انجام داده‌اند با بازشناسی مفهوم مدیریت کیفیت و ابعاد آن پرداخته‌اند. در ادامه ضمن شناسایی عناصر مرتبط با بعد مدیریت کیفیت محصولات آماری به تبیین شدت هر کدام از این عناصر پرداخته‌اند (غفاری و همکاران، ۲۰۱۸).

در پژوهشی که «جان پی اسلون» انجام داده است پرسشنامه‌ای برای اندازه‌گیری کیفیت داده‌ها ارائه داده و نشان می‌دهد که رابطه بین کیفیت داده‌ها و عملکرد سازمانی به‌طور سیستماتیک قابل‌اندازه‌گیری است و این رابطه در بیشتر موارد مثبت است (اسلون، ۲۰۰۶).

در پژوهشی که «واهیودی، کوک و جانسن» انجام داده‌اند با مطالعه موردی یک شرکت مخابراتی بزرگ، یک مدل برای بهبود کیفیت داده‌ها ارائه کرده‌اند. آنها چهار الگوی مختلف برای مقابله با کیفیت متغیر داده‌ها ارائه داده‌اند که می‌توان از آن به‌عنوان یک مدل مرجع برای بهبود کیفیت داده‌های بزرگ استفاده کرد (واهیودی و همکاران، ۲۰۱۸).

¹ Torres & Sidorova

² Pipino

³ Lukyanenko

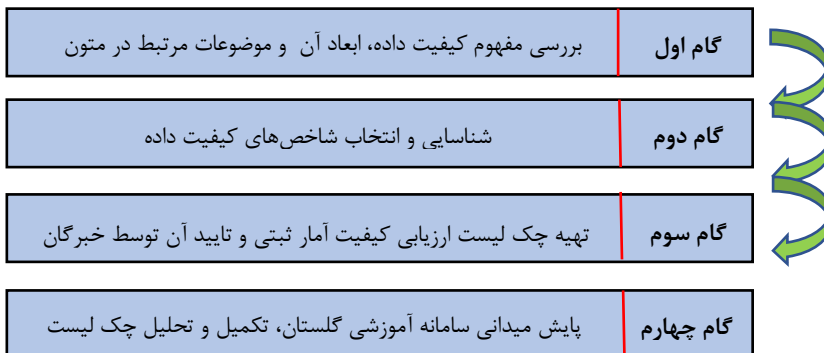
⁴ Wahyudi

با توجه به موارد مطرح‌شده در چارچوب مفهومی، این پژوهش به دنبال پاسخ به پرسش‌های زیر است:

- ۱- کیفیت آمار ثبتي دانشجویان در حوزه شناسایی به چه میزان است؟
- ۲- کیفیت آمار ثبتي دانشجویان در حوزه ثبت به چه میزان است؟
- ۳- کیفیت آمار ثبتي دانشجویان در حوزه انتقال به چه میزان است؟
- ۴- کیفیت آمار ثبتي دانشجویان در حوزه پردازش به چه میزان است؟
- ۵- کیفیت آمار ثبتي دانشجویان در حوزه ارائه به چه میزان است؟

روش شناسی

در این پژوهش، با توجه به ماهیت کار از رویکرد تحقیق آمیخته (کیفی- کمی) و شیوه اکتشافی استفاده شده است. در بخش کیفی در گام اول مفاهیم نظری یعنی کیفیت داده و ابعاد آن، بررسی انواع نگاه‌ها به موضوع کیفیت داده، سیستم‌های اطلاعاتی، نظام آمار ثبتي و حوزه‌های آن و ارتباط آنها با کیفیت داده از طریق مطالعات کتابخانه‌ای گسترده از مقالات و منابع منتشرشده طی سال‌های ۱۹۹۶ تا ۲۰۲۱ جمع‌آوری شده و بر اساس این مطالعات تعاریف مربوطه ارائه شده است. در گام دوم شاخص‌های کیفیت داده شناسایی و مهم‌ترین آنها برای این پژوهش انتخاب شده است. در گام سوم بر اساس نتایج گام‌های قبلی، ابزار گردآوری اطلاعات تحت عنوان «چک‌لیست ارزیابی کیفیت آمار ثبتي» با استفاده از نظرات خبرگان تدوین شده و به دلیل خبره محور بودن این روش، برای بررسی روایی و پایایی چک‌لیست از نظرات ۲۰ نفر از متخصصان، کارشناسان و مدیران مرکز آمار ایران بهره گرفته شده و پس از بررسی نظرات و تأیید ۹۰ درصدی ایشان، چک‌لیست ارزیابی کیفیت آمار ثبتي نهایی شده است. در بخش کمی و گام چهارم سامانه آموزشی گلستان به صورت میدانی مورد پایش قرار گرفته و برای اطلاعات مربوط به دانشجویان تولیدشده توسط این سامانه، چک‌لیست تکمیل و تحلیل شده است. شکل ۷ مراحل اجرای پژوهش را بیان می‌کند.



شکل ۷. مراحل انجام پژوهش.

روش اندازه‌گیری کیفیت آمار ثبتي

همان‌گونه که قبلاً گفته شد در بحث آمارهای ثبتي ۵ حوزه کاری وجود دارند که عبارت‌اند از: شناسایی، ثبت، انتقال، پردازش و ارائه. کیفیت داده‌های تولیدشده در نظام آمارهای ثبتي به عملکرد این پنج حوزه برمی‌گردد و چنانچه سیستم اطلاعاتی نتواند حداقل یکی از حوزه‌های پنج‌گانه فوق را به خوبی تحت کنترل قرار دهد داده‌های تولیدی آن سیستم کیفیتشان مخدوش خواهد شد و با افزایش تعداد حوزه‌هایی که به خوبی تحت کنترل سیستم اطلاعاتی نیستند میزان خدشه به کیفیت داده‌های تولیدی سیستم نیز افزایش خواهد یافت. در مدلی که در اینجا

ارائه شده است ۱۲ بعد کیفیت اندازه‌گیری شده است. باید گفته شود که این ۱۲ بعد با ۵ حوزه مرتبط است به طوری که برخی ابعاد در دو یا چند حوزه مشترک است. لازم به ذکر است که اهمیت ابعاد کیفیت به یک اندازه نیست به طوری که برخی ابعاد از برخی دیگر مهم‌تر می‌باشند. چک‌لیست طراحی شده شامل ۲۹ سؤال است که سؤالات مختلفی در حوزه‌های ۵ گانه آمار ثبتی و ابعاد ۱۲ گانه کیفیت داده را دربرمی‌گیرد. سؤالات ۱ تا ۹ به حوزه شناسایی، ۱۰ تا ۱۷ به حوزه ثبت، ۱۸ تا ۲۲ به حوزه انتقال، ۲۳ تا ۲۵ به حوزه پردازش و ۲۶ تا ۲۹ به حوزه ارائه مرتبط است. برای نمره دهی به سؤالات از طیف لیکرت ۷ گزینه‌ای استفاده شده که به بدترین حالت ممکن نمره ۱ و به بهترین حالت ممکن نمره ۷ داده شده است. ابعاد، تعاریف، وزن و سؤالات مرتبط با کیفیت داده در چک‌لیست ارزیابی کیفیت آمار ثبتی در جدول ۱ آمده است (کیانی و رشیدی، ۲۰۰۹).

جدول ۱. ابعاد، تعاریف، وزن و سؤالات مرتبط با ابعاد کیفیت داده در چک‌لیست ارزیابی کیفیت آمار ثبتی (کیانی و رشیدی، ۲۰۰۹).

شماره پرسش‌ها مرتبط با بعد کیفیت در چک‌لیست ارزیابی آمار ثبتی	وزن	تعریف	بعد کیفیت
۱۰-۱۱-۱۲-۱۳-۱۴-۱۸-۲۰	۲	سیستم اطلاعاتی تمامی حالات ممکن جهان واقعی را در برداشته باشد.	درستی
۱۰-۱۳-۱۶-۱۷-۲۲-۲۴-۲۵-۲۷	۱.۹	یکسانی بین انتظارات و قابلیت‌ها و هم‌نواپی داده‌ها با نیاز کاربران و واقعیات رخ داده در جهان واقعی	قابلیت اعتماد
۱۱-۱۲-۱۳-۱۹-۲۰-۲۱	۱.۸	فاصله زمانی بین تغییرات حالت جهان واقعی و اعمال اصلاحات مربوطه در سیستم اطلاعاتی	به‌هنگام بودن
۱-۲-۳-۴-۵-۶-۷-۸-۹	۱.۷	ثبت تمامی مقادیر مورد نیاز یک متغیر مشخص از جهان واقعی در سیستم اطلاعاتی	کامل بودن
۱۲-۱۸-۲۱	۱.۶	سرعت به‌روزرسانی سیستم اطلاعاتی پس از تغییرات رخ داده در جهان واقعی	جاری بودن
۱۳	۱.۵	عدم نگاهت از یک حالت جهان واقعی به چند حالت سیستم اطلاعاتی	سازگاری
۱۴-۱۶-۱۷-۱۹-۲۰-۲۲-۲۴	۱.۴	عدم نگاهت اشتباه از جهان واقعی به سیستم اطلاعاتی	دقت
۱۵-۲۶-۲۸-۲۹	۱.۳	سیستم اطلاعاتی بتواند تغییرات رخ داده در جهان واقعی را به شکل و قالب مورد نظر کاربر نگاهت کند و آمارهای تولیدی را به صورت کاربر پسند ارائه کند.	شکل و قالب
۱۵-۲۳-۲۶-۲۸	۱.۲	آمارهای مربوط به یک موجودیت برای کاربر کاملاً مفید باشد و بتواند نیازهای وی را تمام و کمال برطرف نماید.	کارایی
۱۵-۲۸	۱.۱	آمارهای مربوط به یک موجودیت برای کاربر کافی باشد و نیازهای اطلاعاتی وی را-بدون اطلاعات اضافی گمراه‌کننده- تأمین کند.	کفایت
۲۳-۲۹	۱.۰	تولید و پردازش اطلاعات در سیستم اطلاعاتی به صورتی است که کاربر به راحتی آن‌ها استفاده کند و نیاز به پردازش اضافی توسط کاربر نباشد.	قابل استفاده بودن
۱-۲-۳-۴-۵-۶-۷-۸-۹	۰.۹	تولید و پردازش اطلاعات در سیستم اطلاعاتی به صورتی است که کاربر به راحتی اطلاعات را با اطلاعات مشابه در سطوح مختلف جغرافیایی و سازمانی و فواصل زمانی مقایسه کند.	مقایسه پذیری

سوال	وزن سوال	نمره یکرت کسب شده	امتیاز مکتسبه	حداکثر امتیاز ممکن	امتیاز کسب شده به تفکیک حوزه	حداکثر امتیاز کسب شده به تفکیک حوزه	درصد امتیاز کسب شده به تفکیک حوزه	امتیاز کسب شده به تفکیک بعد کیفیت	تفکیک بعد کیفیت	درصد امتیاز کسب شده به تفکیک بعد کیفیت
									بعد کفایت	
۲۱	۳.۴	۷	۲۳.۸	۲۳.۸					۵۷.۱	۵۰.۴
۲۲	۳.۳	۷	۲۳.۱	۲۳.۱	۱۲۰.۵	۱۳۰.۹	۹۲.۱	۲۸.۸		
۲۳	۲.۲	۴	۸.۸	۱۵.۴					بعد قابل استفاده بودن	
۲۴	۳.۳	۳	۹.۹	۲۳.۱	حوزه پردازش			۲۲.۶	۳۱.۵	۷۱.۷
۲۵	۱.۹	۵	۹.۵	۱۳.۳	۲۸.۲	۵۱.۸	۵۴.۴		بعد مقایسه پذیری	
۲۶	۲.۵	۳	۷.۵	۱۷.۵				۱۲۴.۸	۱۶۳.۸	۷۶.۲
۲۷	۱.۹	۷	۱۳.۳	۱۳.۳	حوزه ارائه					
۲۸	۳.۶	۳	۱۰.۸	۲۵.۲						
۲۹	۲.۳	۶	۱۳.۸	۱۶.۱	۴۵.۴	۷۲.۱	۶۳.۰			
			۵۱۹.۴	۶۵۵.۹	۵۱۹.۴					
					نمره کل	۷۹.۲				

طبق جدول ۲ در پاسخ به اولین سؤال پژوهش، یافته‌ها نشان می‌دهد در حوزه شناسایی میزان کیفیت آمار ثبتی دانشجویان تقریباً ۷۶ درصد است که می‌توان نمره مناسبی قلمداد کرد یعنی سیستم آموزشی گلستان توانسته است درصد مناسبی از نیازهای حوزه شناسایی را پوشش دهد و برای افزایش میزان کیفیت لازم است تغییراتی در سیستم آموزشی گلستان داده شود تا فیلدهای بلااستفاده مسدود شود. در پاسخ به سؤال دوم میزان کیفیت آمار ثبتی دانشجویان در حوزه ثبت تقریباً ۸۵ درصد است یعنی این که سیستم آموزشی گلستان توانسته است درصد خوبی از فعالیت‌های حوزه ثبت را به‌طور مناسب تحت پوشش خود درآورد، برای افزایش این میزان لازم است مکانیسم‌هایی خودکار و هوشمند تعریف و طراحی گردد تا خطاهای این حوزه به حداقل ممکن کاهش یابد. در حوزه انتقال (سؤال سوم) میزان کیفیت آمار ثبتی دانشجویان تقریباً ۹۲ درصد است که نمره‌ای عالی محسوب می‌شود و بیانگر توانایی سیستم آموزشی گلستان در پوشش درصد بالایی از تراکنش‌های موردنیاز حوزه انتقال می‌باشد. برای پاسخ به سؤال چهارم پژوهش که مربوط به حوزه پردازش است میزان کیفیت آمار ثبتی دانشجویان تقریباً ۵۴ درصد است که نمره ضعیفی به حساب می‌آید و این مفهوم را می‌رساند که سیستم آموزشی گلستان به‌خوبی نتوانسته است پردازش‌های لازم را روی داده‌ها جهت تولید گزارش‌های مدیریتی انجام دهد که بهبود کیفیت این حوزه مستلزم طراحی و تعبیه ابزارهای کنترلی کارآمد برای بررسی نتایج پردازش‌های مختلف است. نهایتاً در پاسخ به آخرین سؤال پژوهش، یافته‌ها نشان می‌دهد که در حوزه ارائه میزان کیفیت آمار ثبتی دانشجویان ۶۳ درصد است و به این مفهوم است که سیستم آموزشی گلستان توانسته است تا حدی نیازهای مدیریتی و کارشناسی را به‌صورت گزارش‌های موردنیازشان از نظر شکل، قالب و محتوی تأمین کند. شرح نقاط ضعف و قوت آمار دانشجویان با در نظر گرفتن تمامی تفکیک‌ها و فراداده‌های مربوطه به تفکیک حوزه‌های ۵ گانه آمار ثبتی در جدول ۳ آمده است.

جدول ۳. نقاط ضعف و قوت کیفیت آمار دانشجویان به تفکیک حوزه‌های آمار ثبتی. منبع: محاسبات.

حوزه	نقاط قوت	نقاط ضعف
شناسایی	پوشش کامل عنوان قلم	
	وجود استانداردها و قراردادهای مشترک در مورد زمان	
	و مکان برای ثبت و انتشار	وجود برخی حالات بلااستفاده در سامانه گلستان
	وجود واحد اندازه‌گیری استاندارد	
ثبت	کامل بودن استانداردها و قراردادهای مشترک در مورد تفکیک‌ها برای ثبت و انتشار	
	ثبت در پایین‌ترین نقطه	عدم وجود مکانیسم‌های کافی برای کنترل و تأیید صحت ثبت
	منطقی بودن فراگرد تولید آمار از پایین‌ترین نقطه تا بالاترین سطح	در مراحل مختلف
	انجام عمل ثبت توسط افراد مسئول و پاسخگو	نامناسب بودن طراحی فرم‌های عملیاتی رایانه‌ای در سامانه گلستان برای تأمین کلیه اطلاعات موردنیاز دانشگاه
انتقال	وجود الزامات قانونی برای ثبت داده‌ها	
	انجام وظایف در ایستگاه‌های انتقال داده	
	انتقال داده‌های ثبتی تولیدشده از مسیرهای مناسب به سطوح بالاتر	
	متناسب بودن تکنولوژی انتقال با آمار و اهمیت فرایند تولید آن	
پردازش	انتقال داده‌های ثبتی توسط افراد مسئول و پاسخگو	
	انجام پردازش‌های لازم روی داده‌های ثبتی	عدم وجود روش‌ها و ابزارهای کنترلی کارآمد برای بررسی نتایج پردازش داده‌ها
	مناسب و کارا بودن شکل و محتوای گزارش‌های خروجی	
	تعیین نسبتاً مناسب سطوح دسترسی افراد به آمار و گزارش‌های خروجی	عدم ارائه گزارش‌های گرافیکی
ارائه	مناسب بودن و کافی بودن تفکیک‌های ارائه‌شده در گزارش‌های خروجی	

نتیجه‌گیری و پیشنهادها

در این مقاله پس از ارائه تعاریف مربوط به داده ثبتی، آمار ثبتی، نظام آمار ثبتی و سیستم‌های اطلاعاتی مشخص شد که نظام آمار ثبتی را می‌توان به‌عنوان سیستم اطلاعاتی در نظر گرفت و بر اساس آن کیفیت داده با رویکرد طرح‌گرا تشریح و سپس ۱۲ بعد کیفیت داده ارائه و با چک‌لیست تدوین‌شده، کیفیت آمار دانشجویان در سیستم گلستان دانشگاه کردستان مورد ارزیابی قرار گرفت و درنهایت کیفیت آمار دانشجویان به تفکیک حوزه‌های آمار ثبتی و ابعاد کیفیت تجزیه و تحلیل شد. کیفیت آمار ثبتی دانشجویان در حوزه شناسایی مناسب، در حوزه ثبت خوب، در حوزه انتقال عالی، در حوزه پردازش ضعیف و در حوزه ارائه متوسط است که با مطالعات (غفاری و همکاران، ۲۰۱۸)، (اسلون، ۲۰۰۶) و

(واهیودی و همکاران، ۲۰۱۸) همسو است. به نظر می‌رسد که تمام سازمان‌ها به دنبال بهبود کیفیت داده‌های خود می‌باشند، البته لازم به ذکر است، در برخی از موقعیت‌ها کیفیت داده ضعیف مورد نظر و مفید می‌باشد. به‌عنوان مثال برای ممانعت از پي بردن دشمن به حالات جهان واقعی در مورد دفاع ملی و مسائل امنیتی یا گمراه کردن رقبای تجاری و اقتصادی، کیفیت داده ضعیف می‌تواند سودمند باشد. پس شناسایی منابع ایجاد نقص برای تولید داده باکیفیت پایین ضروری است. البته این امر بایستی آگاهانه و در موازات سیستم اطلاعاتی اصلی باشد. بنابراین پیشنهاد می‌گردد که دانشگاه کردستان برای بررسی کیفیت دیگر آمارهای تولیدی سامانه آموزشی گلستان و سایر مراکز آموزش عالی کشور نسبت به سنجش کیفیت آمارهای خود اقدام نمایند. همچنین پیشنهاد می‌گردد که در بازطراحی سامانه آموزشی گلستان نقاط ضعف شناسایی شده در یافته‌های پژوهش مدنظر طراحان سیستم قرار گیرد و به‌صورت جزئی می‌توان پیشنهادهای فنی مانند مسدود کردن فیلدهای بلااستفاده در سامانه گلستان، ایجاد مکانیسم‌های کنترلی در مراحل مختلف تولید داده، در نظر گرفتن روش‌ها و ابزارهایی برای بررسی نتایج پردازش داده‌ها و تعبیه امکانات گزارش دهی به‌صورت گرافیکی و نموداری ارائه نمود.

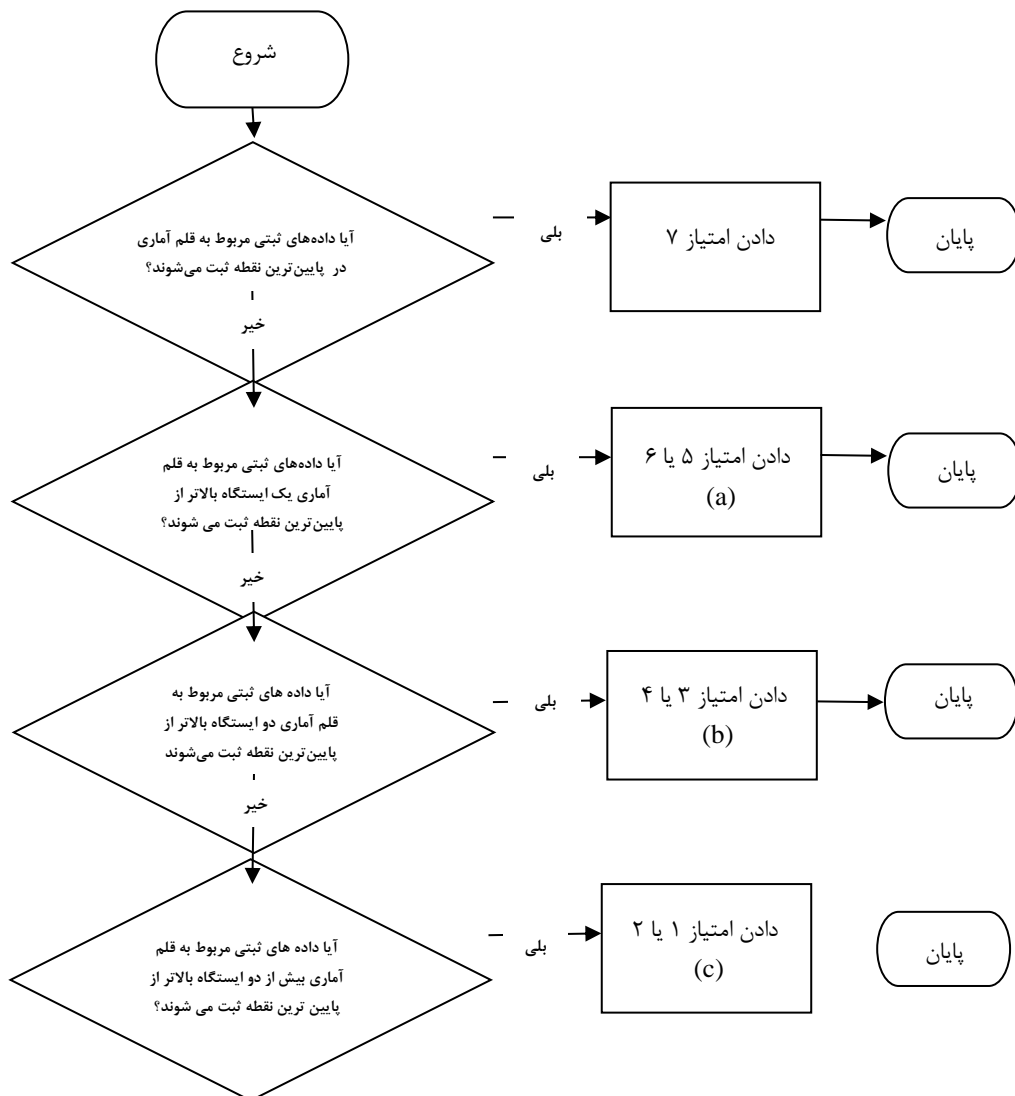
با توجه ماهیت داده‌های ثبتي و شاخص‌های کیفیت داده، چک‌لیست با بررسی دقیق اسناد و مدارک سیستم‌ها، مشاهده و بازرسی میدانی سیستم آموزشی گلستان و گفتگو با کاربران این سیستم توسط دو نفر کارشناس تکمیل که ممکن است باعث ایجاد تورش گردیده باشد.

ضمایم

چک‌لیست ارزیابی کیفیت آمار ثبتی

ردیف	سؤالات چک‌لیست ارزیابی کیفیت آمار ثبتی	۷ بلی	۶	۵	۴	۳	۲	۱ خیر
۱	آیا عنوان قلم تمامی رخدادهای مربوطه را پوشش می‌دهد؟ (کامل بودن - مقایسه پذیری)							
۲	تعاریف و مفاهیم در چه وضعیتی است؟ (کامل بودن - مقایسه پذیری)							
۳	آیا استانداردها و قراردادهای مشترک در مورد زمان (از قبیل سری زمانی، بازه زمانی و مقطع زمانی) برای ثبت قلم آماری مذکور رعایت می‌شود؟ (کامل بودن - مقایسه پذیری)							
۴	آیا استانداردها و قراردادهای مشترک در مورد زمان (از قبیل سری زمانی، بازه زمانی و مقطع زمانی) برای انتشار قلم آماری مذکور رعایت می‌شود؟ (کامل بودن - مقایسه پذیری)							
۵	آیا استانداردها و قراردادهای مشترک در مورد مشخصات مکانی (از قبیل سطوح جغرافیایی، یکپارچگی داده‌ها، ساختار و قابلیت استفاده در سیستم‌های اطلاعات جغرافیایی GIS و...) برای ثبت قلم آماری مذکور رعایت می‌شود؟ (کامل بودن - مقایسه پذیری)							
۶	آیا استانداردها و قراردادهای مشترک در مورد مشخصات مکانی (از قبیل سطوح جغرافیایی، یکپارچگی داده‌ها، ساختار و قابلیت استفاده در سیستم‌های اطلاعات جغرافیایی GIS و...) برای انتشار قلم آماری مذکور رعایت می‌شود؟ (کامل بودن - مقایسه پذیری)							
۷	آیا استانداردها و قراردادهای مشترک در مورد واحد اندازه‌گیری برای قلم آماری مذکور رعایت می‌شود؟ (کامل بودن - مقایسه پذیری)							
۸	آیا استانداردها و قراردادهای مشترک در مورد تفکیک‌ها (از قبیل سن، جنسیت و...) و طبقه‌بندی‌های بین‌المللی (از قبیل ISIC, ISCED و...) برای ثبت قلم آماری مذکور رعایت می‌شود؟ (کامل بودن - مقایسه پذیری)							
۹	آیا استانداردها و قراردادهای مشترک در مورد تفکیک‌ها (از قبیل سن، جنسیت و...) و طبقه‌بندی‌های بین‌المللی (از قبیل ISIC, ISCED و...) برای انتشار قلم آماری مذکور رعایت می‌شود؟ (کامل بودن - مقایسه پذیری)							
۱۰	آیا داده‌های ثبتی مربوط به قلم آماری ارائه‌شده در پایین‌ترین نقطه ثبت می‌شوند؟ (درستی - قابلیت اعتماد)							
۱۱	آیا داده‌های ثبتی مربوط به قلم آماری به هنگام وقوع عملیات ثبت می‌شوند؟ (درستی - بهنگام بودن)							
۱۲	آیا ورود اطلاعات به سیستم هم‌زمان با انجام عملیات ثبت صورت می‌گیرد؟ (جاری بودن - بهنگام بودن - درستی)							
۱۳	آیا فراگرد تولید قلم آماری ارائه‌شده از پایین‌ترین نقطه تا بالاترین سطح منطقی به نظر می‌رسد؟ (درستی - قابلیت اعتماد - بهنگام بودن - سازگاری)							
۱۴	آیا عمل ثبت در فرآیند تولید توسط افراد مسئول و پاسخگو انجام می‌شود؟ (درستی - دقت)							

ردیف	سؤالات چک‌لیست ارزیابی کیفیت آمار ثبتی	۷ بلی	۶	۵	۴	۳	۲	۱ خیر
۱۵	آیا فرم‌های عملیاتی که در فرآیند تولید اقلام آماری به کار می‌روند برای تأمین اطلاعات موردنیاز طراحی مناسبی دارند؟ (شکل و قالب - کارایی - کفایت)							
۱۶	آیا برای کنترل و تأیید صحت ثبت در مراحل مختلف، مکانیسمی وجود دارد؟ (قابلیت اعتماد - دقت)							
۱۷	آیا الزامات قانونی برای ثبت داده‌های مربوط به اقلام آماری وجود دارد؟ (دقت - قابلیت اعتماد)							
۱۸	آیا عمل انتقال در فرآیند تولید توسط افراد مسئول و پاسخگو انجام می‌شود؟ (درستی - جاری بودن)							
۱۹	آیا داده‌های ثبتی تولیدشده از مسیرهای مناسب به سطوح بالاتر منتقل می‌شوند؟ (به‌هنگام بودن - دقت)							
۲۰	آیا در مسیر انتقال، هر ایستگاه وظایف خود (از قبیل کنترل، تجمیع درست یا اقدام به‌موقع برای انتقال اطلاعات و ...) را به‌خوبی انجام می‌دهد؟ (به‌هنگام بودن - درستی - دقت)							
۲۱	آیا تکنولوژی انتقال متناسب با قلم آماری و اهمیت و فرآیند تولید آن می‌باشد؟ (جاری بودن - به‌هنگام بودن)							
۲۲	آیا امنیت داده‌ها و اطلاعات هنگام انتقال رعایت می‌شود؟ (قابلیت اعتماد - دقت)							
۲۳	آیا پردازش‌های لازم روی داده‌های ثبتی صورت می‌گیرد؟ (کارایی - قابل استفاده بودن)							
۲۴	آیا روش‌ها و ابزارهای کنترلی کارآمدی برای بررسی نتایج پردازش داده‌های ثبتی وجود دارد؟ (قابلیت اعتماد - دقت)							
۲۵	آیا در فرآیند ذخیره اطلاعات، امکان تهیه فایل پشتیبان با دوره زمانی مناسب وجود دارد؟ (قابلیت اعتماد)							
۲۶	آیا شکل و محتوای گزارش‌های خروجی مناسب و کارا می‌باشند؟ (شکل و قالب - کارایی)							
۲۷	آیا سطوح دسترسی افراد به آمار و گزارش‌های خروجی تعریف شده است؟ (قابلیت اعتماد)							
۲۸	آیا تفکیک‌های ارائه‌شده در گزارش‌های خروجی کافی و مناسب می‌باشند؟ (کفایت - کارایی - شکل و قالب)							
۲۹	آیا ابزارهای (سخت‌افزار - نرم‌افزار) مناسبی برای بازیابی و ارائه اطلاعات وجود دارد؟ (شکل و قالب - قابل استفاده بودن)							



- (a) بسته به تعداد کل ایستگاه و طول کل فرآیند دارد اگر تعداد ایستگاه‌ها زیاد باشد امتیاز ۶ و اگر کم باشد امتیاز ۵ می‌دهیم.
- (b) بسته به تعداد کل ایستگاه و طول کل فرآیند دارد اگر تعداد ایستگاه‌ها زیاد باشد امتیاز ۴ و اگر کم باشد امتیاز ۳ می‌دهیم.
- (c) بسته به تعداد کل ایستگاه و طول کل فرآیند دارد اگر تعداد ایستگاه‌ها زیاد باشد امتیاز ۲ و اگر کم باشد امتیاز ۱ می‌دهیم.

فلوچارت تکمیل سؤال ۱۰ (به‌عنوان نمونه ضمیمه‌شده است).

References

- Abate, M. L., Diegert, K. V., & Allen, H. W. (1998). A Hierarchical Approach to Improving Data Quality. *Data Quality*, 4(1), 365-369. <https://www.semanticscholar.org/paper/A-Hierarchical-Approach-to-Improving-Data-Quality-Abate-Diegert/4548003c9c89d50b1a3045bfa9970097b7c3b6c9>
- Almasifard, M., Bakhsham, M., Karimi, H., & Behvar, S. (2021). Designing a pathology model of academic entrepreneurship development and knowledge commercialization with a structural interpretive approach. *Karafan Quarterly Scientific Journal*, 17(5), 87-110. <https://doi.org/10.48301/kssa.2021.128449>
- Batini, C., & Scannapieco, M. (2016). *Data and Information Quality: Dimensions, Principles and Techniques*. Springer Cham. <https://doi.org/10.1007/978-3-319-24106-7>
- Cai, L., & Zhu, Y. (2015). The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. *Data Science Journal*, 14(2), 1-10. <https://doi.org/10.5334/dsj-2015-002>
- Ghafari, M., Azar, A., & Shabbak, A. (2018). Causal Mapping in Managing Statistical Outputs with Data Quality Approach. *Iranian Journal Of Information Processing & Management* 33(3), 1053-1076. <https://www.sid.ir/paper/376968/en>
- Haug, A., Zachariassen, F., & Van Liempd, D. (2011). The costs of poor data quality. *Journal of Industrial Engineering and Management* 4(2), 168-193. <https://www.econstor.eu/handle/10419/188448>
- Kahn, B. K., Strong, D. M., & Wang, R. Y. (1997, October 24-26). *A Model for Delivering Quality Information as Product and Service*. International Conference on Information Quality., Cambridge, Massachusetts, USA. https://www.researchgate.net/publication/220918643_A_Model_for_Delivering_Quality_Information_as_Product_and_Service
- Kent, W. (2012). *Data and Reality: A Timeless Perspective on Perceiving and Managing Information* (3 ed.). Technics publications. <https://books.google.com/books?id=4hvBgAAQBAJ&printsec=frontcover#v=onepage&q&f=false>
- Kiani, K., & Rashidi, S. (2009). *Determining the quality indicators of the register data to evaluate the quality of mentioned data*. Report of the research plan approved by the Statistical Research and Training Cente.
- Lukyanenko, R., Wiggins, A., & Rosser, H. K. (2020). Citizen science: An information quality research frontier. *Information Systems Frontiers*, 22(4), 961-983. <https://doi.org/10.1007/s10796-019-09915-z>
- Nourmohammadi, M., & Navvabpour, H. (2005). Introduction to register statistics systems. *Statistical Center of Iran*.
- Orr, K. (1998). Data quality and systems theory. *Communications of the Association for Computing Machinery*, 41(2), 66-71. <https://dl.acm.org/doi/pdf/10.1145/269012.269023>
- Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the Association for Computing Machinery*, 45(4), 211-218. <https://doi.org/10.1145/505248.506010>
- Rahimi, H., & Aghababaei, R. (2020). The effect of authentic leadership on human resource creativity: Mediating role of internal motivation (Case study: staff in University of Kashan). *Karafan Quarterly Scientific Journal*, 16(2), 121-140. https://karafan.tvu.ac.ir/article_105317.html?lang=en

- Redman, T. C. (2016, September 22). *Bad Data Costs the U.S. \$3 Trillion Per Year*. Harvard Business School Publishing. <https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year>
- Shankaranarayanan, G., & Blake, R. (2017). From content to context: The evolution and growth of data quality research. *Journal of Data and Information Quality*, 8(2), 1-28. <https://doi.org/10.1145/2996198>
- Shanks, G. G., & Darke, P. (1998, October 25). *Understanding Data Quality and Data Warehousing: A Semiotic Approach*. Information Quality, Cambridge, Massachusetts, USA. <http://18.4.38.158/ICIQ/Documents/IQ%20Conference%201998/Papers/UnderstandDQinDataWarehouseSemioticAppr.pdf>
- Slone, J. P. (2006). *Information quality strategy: An empirical investigation of the relationship between information quality improvements and organizational outcomes* [Doctoral, Capella University]. Minneapolis, Minnesota. <https://web.mit.edu/smadnick/www/IQ%20Dissertations/Skip.Slone.Dissertation.pdf>
- Timmerman, Y., & Bronselaer, A. (2019). Measuring data quality in information systems research. *Decision Support Systems*, 126, 113138. <https://doi.org/10.1016/j.dss.2019.113138>
- Torres, R., & Sidorova, A. (2019). Reconceptualizing information quality as effective use in the context of business intelligence and analytics. *International Journal of Information Management*, 49(3), 316-329. <https://doi.org/10.1016/j.ijinfomgt.2019.05.028>
- Wahyudi, A., Kuk, G., & Janssen, M. (2018). A Process Pattern Model for Tackling and Improving Big Data Quality. *Information Systems Frontiers*, 20(3), 457-469. <https://doi.org/10.1007/s10796-017-9822-7>
- Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the Association for Computing Machinery*, 39(11), 86-95. <https://doi.org/10.1145/240455.240479>
- Wand, Y., & Weber, R. (1995). On the deep structure of information systems. *Information Systems Journal*, 5(3), 203-223. <https://doi.org/10.1111/j.1365-2575.1995.tb00108.x>
- Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), 5-33. <https://doi.org/10.1080/07421222.1996.11518099>
- Zakeri, B. (2007). *Structured methods of analysis and design of information systems* (10 ed.). Industrial Management Organization. <https://www.gisoom.com/book/1425354/>