

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



وزارت علوم، تحقیقات و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران

طرح پژوهشی

**طراحی مدل کیفیت فراداده در سامانه ثبت پایان نامه /
رساله‌های دانش‌آموختگان داخل و خارج کشور**

مجری

محمدجواد ارشادی

همکاران

جلال‌الدین نصیری

فرهاد شیرانی

بهمن ۱۳۹۶

حقوق: پژوهشگاه علوم و فناوری اطلاعات ایران

طرح پژوهشی «طرحریزی و استقرار نظام تضمین کیفیت در سامانه‌های گردآوری و ثبت، سازمان‌دهی، و اشاعه اطلاعات پایان‌نامه‌ها/رساله‌های دانش‌آموختگان داخل کشور» پیرو قرارداد شماره ۳۳/۱۵۹۵۹ مورخ ۱۳۹۵/۱۲/۱۱ میان پژوهشگاه علوم و فناوری اطلاعات ایران (کارفرما) و آقای محمدجواد ارشادی (به عنوان مجری) اجرا شده است. گزارش حاضر، از مستندات این طرح است. این گزارش و تمامی حقوق مادی آن بر اساس «قانون حمایت حقوق مؤلفان و مصنفان و هنرمندان، مصوب سال ۱۳۴۸ و اصلاحیه‌های بعدی آن و همچنین آیین‌نامه‌های اجرایی این قانون متعلق به پژوهشگاه علوم و فناوری اطلاعات ایران است و هر گونه استفاده از تمامی یا پاره‌ای از آن، شامل: نقل قول، تکثیر، انتشار، کاربرد نتایج، تکمیل و مانند آنها به صورت چاپی، الکترونیکی یا وسایل دیگر فقط با اجازه کتبی پژوهشگاه امکان‌پذیر است. نقل قول در حد هزار واژه در انتشارات علمی مانند کتاب و مقاله با درج اطلاعات کامل کتاب‌شناختی، نیازی به مجوز پژوهشگاه ندارد.

صحت مندرجات گزارش برعهده مجری طرح پژوهشی است.

۳۳/۱۵۱۳۸

۱۳۹۶/۱۱/۱۲

شماره:

تاریخ: ندارد

پیوست:

به نام خدا

وزارت علوم، تحقیقات و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران

ایرانداک



پژوهش، مدیریت دانش، آموزش

گواهی

انجام طرح پژوهشی:

◇ عنوان: طراحی مدل کیفیت فراداده در سامانه ثبت پایان نامه / رساله های دانش آموختگان

داخل و خارج کشور

◇ مجری: دکتر محمدجواد ارشادی

◇ همکاران: دکتر جلال الدین نصیری، دکتر فرهاد شیرانی

◇ شماره قرارداد: ۳۳/۱۵۹۵۹

◇ تاریخ قرارداد: ۱۳۹۵/۱۲/۱۱

◇ تاریخ شروع: ۱۳۹۵/۱۲/۱۱

◇ تاریخ پایان: ۱۳۹۶/۱۱/۱

◇ ناظر: دکتر آرمان ساجدی نژاد

در پژوهشگاه علوم و فناوری اطلاعات ایران گواهی می شود. گزارش پایانی این طرح، بر پایه داوری در نشست ۲۷۵ (تاریخ ۱۳۹۶/۱۱/۱۰) شورای پژوهشی پژوهشگاه به تصویب رسیده است. لازم می دانم مراتب قدردانی خود را به مجری، همکاران و ناظر این طرح تقدیم دارم.

سپاس از روی پژوهشی

پرویز شهریاری

سرپرست معاونت پژوهش و آموزش

طراحی مدل کیفیت فراداده در سامانه ثبت پایان نامه / رساله‌های

دانش‌آموختگان داخل و خارج کشور

مدیر طرح پژوهشی: محمدجواد ارشادی، هیئت علمی پژوهشگاه علوم و فناوری اطلاعات ایران

نشانی: تهران، خیابان انقلاب، چهارراه فلسطین، شماره ۱۰۹۰، طبقه ۴، اتاق ۴۰۲

صندوق پستی: ۱۳۱۵۷۷۳۳۱۴ تلفن: ۰۲۱-۶۶۴۹۴۹۸۰، دورنگار: ۰۲۱-۶۶۴۹۴۹۸۰

وبگاه: <http://www.Irandoc.ac.ir>

رایانامه: Ershadi@irandoc.ac.ir

همکار اصلی: دکتر جلال‌الدین نصیری، دکتر فرهاد شیرانی

چکیده

سامانه ثبت پایان نامه/رساله های دانش آموختگان کل کشور بعنوان یک سامانه های حیاتی در ثبت و اشاعه مدارک علمی بوده و نقشی انکارناپذیر در توسعه پژوهش ها دارد. از این رو سطح کیفیت مطلوب سامانه تاثیر ویژه ای در ارتقاء کیفیت پژوهش های محققین خواهد داشت. نام، کد ملی، مقطع تحصیلی، رشته دانشگاهی مشخصات استاد راهنما، مشخصات پیشنهاددهنده و پایان نامه از مهم ترین اقلام اطلاعاتی می باشند. بمنظور بهبود سطح کیفیت در این سامانه تعیین وضعیت مطلوب هریک از اقلام اطلاعاتی سامانه ثبت اقدامی بنیادین محسوب می گردد. در این راستا مدل های کیفیت داده از متون علمی مختلف استخراج و جمع بندی گردیده و وضعیت مطلوب کیفیت فراداده های مربوط به پیشنهاددهنده/پایان نامه/رساله ها ترسیم گردید. بررسی وضعیت موجود فرایند ثبت و اشاعه مدارک علمی، مشکلات کیفی مشاهده شده، سیاست ها و راهبردهای اشاعه از اقدامات اصلی خواهد بود. نقش هریک از افراد درگیر در سامانه ثبت (نمایندگان دانشگاهها، پرسنل واحد فراهم آوری) در قالب چک لیست های کنترل کیفی فراداده ها تبیین گردید. چارچوبی بمنظور درجه بندی این عدم انطباق ها (اصلی و جزئی) و همچنین اقدامات لازم در صورت بروز تعیین شد. این اقدامات اصلاح فراداده، حذف آن، پذیرش و ارسال فراداده به نمایه ساز و ... را شامل می شود. کلیه مستندات تدوین شده در قالب طرح کیفیت فراداده سامانه ثبت ارائه شد.

کلیدواژه ها: فراداده؛ کنترل کیفیت؛ طرح کیفیت؛ سامانه های اطلاعاتی تحقیقاتی؛ سامانه ثبت پایان نامه/رساله

قدردانی

اکنون که به یاری خدا این طرح به پایان رسیده است، لازم است مراتب تشکر و سپاس صمیمانه خود را از جناب آقای دکتر عمار جلالی منش معاونت گرامی اطلاعات علم و فناوری ایران پژوهشگاه علوم و فناوری اطلاعات ایران، و کلیه همکاران آن معاونت، و همچنین سرکار خانم شمسی رئیس گرامی اداره سامانه‌های اطلاعاتی، که در اجرای این طرح پژوهشی نهایت همکاری و همراهی را با اینجانب داشتند ابراز نمایم.

فهرست مطالب

عنوان	شماره صفحه
چکیده	أ
قدردانی	ب
فهرست مطالب	ت
فصل اول – مقدمه	۱۰
۱-۱- معرفی پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)	۱۱
۱-۲- معرفی مرکز اطلاعات و مدارک علمی ایران	۱۲
۱-۳- معرفی سامانه گنج	۱۴
۱-۴- تبیین موضوع پژوهش	۱۵
فصل دوم – معرفی مبانی کیفیت داده و اطلاعات	۱۸
۲-۱- مقدماتی در مورد کیفیت داده ها	۱۹
۲-۲- مقایسه ی دیدگاهها در تجزیه و تحلیل روشها	۲۱
۲-۲-۱- مراحل و گامهای مشترک	۲۲
۲-۲-۲- استراتژیها و تکنیکها	۲۵
۲-۲-۳- ابعاد	۲۸
۲-۲-۴- هزینه ها	۳۵
۲-۲-۵- انواع داده ها	۳۶
۲-۲-۶- ارتباط کیفیت داده با سیستمهای اطلاعاتی	۳۹
۲-۳- مقایسه ی متدلوژی های مختلف کیفیت داده ها	۴۲
۲-۳-۱- روشها، گامها و مراحل	۴۴
۲-۳-۲- فاز ارزیابی	۴۵
۲-۳-۳- گام بهبود	۴۷

۴-۲	روش‌ها، استراتژی‌ها و تکنیک‌ها.....	۴۸
۴-۲-۱	تکنیک‌های داده محور.....	۴۹
۴-۲-۲	تکنیک‌های فرآیند محور.....	۵۲
۵-۲	روش‌ها و ابعاد.....	۵۳
۶-۲	روش‌ها و هزینه‌ها.....	۶۶
۶-۲-۱	ادسته‌بندی هزینه‌ها.....	۶۶
۶-۲-۲	معیارهای کمی سازی هزینه‌ها.....	۶۸
۷-۲	روش‌ها و انواع داده‌ها.....	۷۰
۸-۲	روش‌ها و انواع سیستم‌های اطلاعاتی.....	۷۳
۹-۲	مقایسه‌ی مجمل روش‌ها.....	۷۸
۱۰-۲	از کیفیت داده به سمت کیفیت اطلاعات.....	۸۷
۱۱-۲	کیفیت داده‌ها و کیفیت فرآیند.....	۸۸
۱۲-۲	کیفیت داده‌ها و انواع جدید سیستم‌های اطلاعاتی.....	۸۹
۱۳-۲	خلاصه‌ای از روش‌ها.....	۹۲
۱۳-۲-۱	روش TDQM (مدیریت جامع کیفیت داده).....	۹۳
۱۳-۲-۲	روش DWQ (کیفیت انبار داده).....	۹۷
۱۳-۲-۳	روش TIQM (مدیریت جامع کیفیت اطلاعات).....	۱۰۰
۱۳-۲-۴	روش AIMQ (روشی برای ارزیابی کیفیت اطلاعات).....	۱۰۲
۱۳-۲-۵	روش CIHI (موسسه کانادایی اطلاعات سلامت).....	۱۰۳
۱۳-۲-۶	روش DQA (ارزیابی کیفیت داده‌ها).....	۱۰۷
۱۳-۲-۷	روش IQM (اندازه‌گیری کیفیت اطلاعات).....	۱۰۹
۱۳-۲-۸	روش ISTAT (اداره ملی سرشماری ایتالیا).....	۱۱۱
۱۳-۲-۹	روش AMEQ (فعالیت بر اساس سنجش و تغییر کیفیت اطلاعات محصول).....	۱۱۴
۱۳-۲-۱۰	روش COLDQ (تأثیر کیفیت پائین داده بر هزینه‌ها).....	۱۱۷
۱۳-۲-۱۱	روش DaQuinCIS (کیفیت داده در سیستم‌های اطلاعاتی تعاونی).....	۱۲۰

۱۲۴	۲-۱۳-۱۲- روش QAFD (ارزیابی کیفیت داده‌های مالی)
۱۲۷	۲-۱۳-۱۳- روش CDQ (کیفیت داده‌های کامل)
۱۳۰	۲-۱۴- خلاصه‌های از روشهای ذکر شده
۱۳۱	فصل سوم - شناسایی وضعیت موجود
۱۳۲	۳-۱ معرفی مدل فرایندی سامانه‌های گردآوری و ثبت، سازماندهی و اشاعه
۱۳۵	۳-۲ شناسایی وضعیت سامانه ثبت
۱۳۸	۳-۳ شناسایی وضعیت داده‌های ارائه شده از سامانه CRM
۱۴۰	۳-۴ شناسایی نحوه کنترل فراداده‌ها در سامانه ثبت به کمک مصاحبه با خبرگان
۱۴۵	۳-۵ کارگروه بهبود و نقش آن در شناسایی مشکلات سامانه گنج
۱۴۹	۳-۶ شناسایی و دسته‌بندی مشکلات سامانه گنج
۱۵۵	فصل چهارم - تعیین وضعیت مطلوب اقلام اطلاعاتی سامانه ثبت
۱۵۶	۴-۱ مقدمه
۱۵۷	۴-۲ معرفی فیلدهای سامانه ثبت
۱۶۰	۴-۳ معرفی مدل کیفیت فراداده انتخابی
۱۷۰	۴-۴ انگاشت اولیه ابعاد کیفی فراداده روی اقلام اطلاعاتی سامانه ثبت
۱۸۶	فصل پنجم - تعیین وضعیت مطلوب اقلام اطلاعاتی سامانه ثبت
۱۸۷	۵-۱ تدوین اولیه طرح کیفیت سامانه ثبت بر مبنای مصاحبه با پرسنل کلیدی
۲۰۷	۵-۲ ارائه طرح نهایی کیفیت فراداده
۲۳۹	۵-۳ فرایند کنترل کیفیت در سامانه ثبت
۲۵۹	۵-۴ ارائه پیشنهادات در حوزه طراحی و ویرایش‌های آتی سامانه ثبت
۱۸۴	فهرست منابع
۲۶۶	منابع لاتین
۲۶۶	منابع فارسی

فهرست جداول

شماره صفحه	عنوان
۳۰	جدول ۲-۱ تعاریف مربوط به دقت در مراجع مختلف
۳۰	جدول ۲-۲ تعاریف مختلف ارائه شده برای جامعیت
۳۳	جدول ۲-۳ تعاریف موجود برای ابعاد مرتبط به زمان
۴۲	جدول ۲-۴ مدل‌های مختلف کیفیت داده‌ها
۴۳	جدول ۲-۵ مقایسه مدل‌های مختلف کیفیت داده‌ها مبتنی بر روش‌های ارزیابی
۴۴	جدول ۲-۶ مقایسه مدل‌های مختلف کیفیت داده‌ها مبتنی بر گام‌های بهبود
۴۵	جدول ۲-۷ مقایسه مدل‌های مختلف کیفیت داده‌ها مبتنی بر گام‌های بهبود
۴۶	جدول ۲-۸ روش‌ها و انواع استراتژی‌ها
۵۶	جدول ۲-۹ روش‌ها و ابعاد کیفی
۵۹	جدول ۲-۱۰ ابعاد و معیارها
۶۲	جدول ۲-۱۱ روش‌ها و معیارهای کیفی
۶۵	جدول ۲-۱۲ مقایسه میان دسته‌بندی English و Loshin
۷۰	جدول ۲-۱۳ روش‌ها و انواع داده‌ها
۷۲	جدول ۲-۱۴ روش‌ها و سیستم‌های اطلاعاتی
۱۳۰	جدول ۲-۱۵ جدولی از جمع‌بندی روش‌های مطالعه شده
۱۳۹	جدول ۳-۳ آمار وضعیت درخواستها در سامانه CRM
۱۴۷	جدول ۳-۱ نتایج بررسی برخی از مشکلات موجود در سامانه گنج
۱۴۹	جدول ۳-۲ فهرست مشکلات موجود در سامانه گنج
۱۴۴	جدول ۴-۱: شناسنامه فیلدهای فراداده‌های ثبت پایان‌نامه‌ها/رساله‌ها
۱۵۸	جدول ۴-۲: طرح کیفی فراداده‌ها در سامانه ثبت (کیفیت ذاتی)
۱۶۲	جدول ۴-۳: طرح کیفی فراداده‌ها در سامانه ثبت (کیفیت زمینه‌ای)
۱۶۵	جدول ۴-۴: طرح کیفی فراداده‌ها در سامانه ثبت (کیفیت بازنمایی)
۱۶۷	جدول ۴-۵: طرح کیفی فراداده‌ها در سامانه ثبت (قابلیت دسترسی)
۱۹۵	جدول ۵-۱: طرح اولیه کیفیت فراداده‌ها در سامانه ثبت
۲۰۴	جدول ۵-۲: طرح اولیه کیفیت فایل Pdf داده‌ها در سامانه ثبت
۲۰۹	جدول ۵-۳: طرح نهایی کیفیت فراداده‌ها در سامانه ثبت
۲۱۸	جدول ۵-۴: طرح نهایی کیفیت فراداده‌ها در سامانه ثبت (فایل Pdf)

جدول ۵-۵: طرح نهایی کیفیت فراداده‌ها در سامانه ثبت (فایل Word).....	۲۲۲
جدول ۵-۶: طرح نهایی کیفیت فراداده‌ها در سامانه ثبت دانشجویان خارجی.....	۲۲۷
جدول ۵-۶: طرح نهایی کیفیت فراداده‌ها در سامانه ثبت دانشجویان خارجی.....	۲۲۷
جدول ۵-۷: طرح نهایی کیفیت فراداده‌ها در سامانه ثبت دانشجویان خارجی (فایل Pdf).....	۲۳۶

فهرست شکل ها

عنوان	شماره صفحه
شکل ۱-۱ میانگین سهم بازدید کنندگان سامانه های مختلف در پژوهشگاه علوم و فناوری اطلاعات ...	۱۶
شکل ۱-۲ مقادیر NULL و جامعیت داده ها	۳۳
شکل ۲-۲ نمایش های مختلفی از اطلاعاتی یکسان در دنیای واقعی	۳۸
شکل ۳-۲ طرح کلی روش ISTAT	۷۵
شکل ۴-۲ معماری چارچوب ارزیابی کیفیت در سیستم های اطلاعات تعاونی	۷۷
شکل ۵-۲ یک دسته بندی از روش ها	۷۹
شکل ۶-۲ فازهای مدیریت کیفیت داده	۹۵
شکل ۷-۲ فازهای روش شناسی کیفیت انبار داده	۹۹
شکل ۸-۲ فازهای مدیریت کیفیت اطلاعات کلی	۱۰۱
شکل ۹-۲ مدل PSP/IQ	۱۰۳
شکل ۱۰-۲ فازهای ارزیابی کیفیت اطلاعات	۱۰۳
شکل ۱۱-۲ فازهای روش موسسه کانادایی برای روش اطلاعات سلامت	۱۰۶
شکل ۱۲-۲ فازهای ارزیابی کیفیت داده	۱۰۸
شکل ۱۳-۲ فازهای اندازه گیری کیفیت اطلاعات	۱۱۰
شکل ۱۴-۲ فازهای روش ISTAT	۱۱۳
شکل ۱۵-۲ فازهای اندازه گیری مبتنی بر فعالیت و ارزیابی کیفیت اطلاعات محصول	۱۱۶
شکل ۱۶-۲ مراحل ارزیابی هزینه مؤثر برای کیفیت کم داده ها	۱۱۹
شکل ۱۷-۲ فازهای ارزیابی کیفیت در سیستم های اطلاعات تعاونی	۱۲۲
شکل ۱۸-۲ زیرساخت های ارزیابی کیفیت در سیستم های اطلاعات تعاونی	۱۲۲
شکل ۱۹-۲ فازهای ارزیابی کیفیت داده های مالی	۱۲۶
شکل ۲۰-۲ فازهای روش جامع برای مدیریت کیفیت داده	۱۲۹
شکل ۱-۳ ورودی و خروجی سامانه های گردآوری و ثبت، سازمان دهی، و اشاعه اطلاعات پارس	۱۳۲
شکل ۲-۳ فرایند ثبت پایان نامه / رساله در سامانه ثبت	۱۳۴
شکل ۳-۳ تحلیل مشکلات کیفی سامانه گنج	۱۵۳
شکل ۴-۱ چهار مرحله اصلی در مدیریت کیفیت جامع داده	۱۶۲
شکل ۲-۴ معرفی چرخه دمیگ	۱۶۴
شکل ۳-۴ حوزه های اصلی مرتبط در زمینه کیفیت داده	۱۶۷

شکل ۴-۴ ابعاد کیفیت فراداده‌ها در سامانه ثبت	۱۶۹
شکل ۵-۱. نمونه چک‌لیست اول در کنترل فراداده‌ها در سامانه ثبت	۱۹۰
شکل ۵-۲. نمونه چک‌لیست دوم در کنترل فراداده‌ها در سامانه ثبت	۱۹۱
شکل ۵-۳. نمونه چک‌لیست سوم در کنترل فراداده‌ها در سامانه ثبت	۱۹۳
شکل ۵-۴. صفحه ورود به سامانه ثبت به منظور کنترل مدرک ارسال شده توسط دانشجویان	۲۳۹
شکل ۵-۵. مشاهده فهرست درخواست‌های در حال انتظار در سامانه ثبت	۲۴۰
شکل ۵-۶. آغاز فرایند بررسی مدرک ارسال شده توسط دانشجو	۲۴۱
شکل ۵-۷. صفحه اطلاعات شخصی و تحصیلی دانشجو	۲۴۳
شکل ۵-۸. صفحه اطلاعات پایان‌نامه / رساله	۲۴۴
شکل ۵-۹. صفحه اطلاعات چکیده‌های پایان‌نامه / رساله	۲۴۵
شکل ۵-۱۰. صفحه اطلاعات فهرست و کلیدواژه‌های پایان‌نامه / رساله	۲۴۵
شکل ۵-۱۱. صفحه اعلام نظر نهایی در خصوص پایان‌نامه / رساله	۲۴۶
شکل ۵-۱۲. چک‌لیست نهایی کنترل فراداده‌ها در سامانه ثبت	۲۶۱

فصل اول – مقدمه

۱-۱- معرفی پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)

به دلیل ضرورت استفاده از پژوهش های انجام شده و معتبر برای علاقه مندان و دانشجویان و همچنین ضرورت استفاده از اطلاعات معتبر پژوهشگاه علوم و فناوری ایران با داشتن ۴۵ سال سابقه به عنوان سازمان متولی و نهاد اداره کننده سامانه ملی مدیریت اطلاعات تحقیقاتی تعیین شده است. پژوهشگاه وظایف خود را با همکاری و همراهی نزدیک ۱۶۰ استاد و کارمند به انجام می رساند. دستاوردهای مدیریت دانش و اطلاعات علمی و فنی کشور در سامانه ها و پایگاه های اطلاعات ارائه می شوند. در این میان، سامانه گنج با صدها هزار رکورد از بزرگ ترین، دیرین ترین و پربازدیدترین پایگاه های اطلاعات علمی و فنی کشور است. اطلاعات پایان نامه ها و رساله ها و پیشنهادیه آن ها که در ایرانداک به ثبت می رسند، بی همتاست. تولید ارزش افزوده و اطلاعات پشتیبان و همچنین تحلیل اطلاعات، بخشی دیگر از مدیریت دانش و سامانه های پیشینه پژوهش، همانندجو، دروازه اطلاعات علمی ایران، اصطلاح نامه، راهنمای پژوهش و پژوهشگران و نمای مقاله های ایرانیان در جهان نمونه هایی از آن ها هستند.

ایرانداک از سال ۱۳۶۱ دوره های آموزشی کوتاه مدت بسیاری را ارائه می کند و از سال ۱۳۹۰ نیز در سه رشته مدیریت فناوری اطلاعات، علم اطلاعات و دانش شناسی، و مهندسی فناوری اطلاعات دانشجوی دکترا پذیرفته است که رساله های خود را در زمینه های گوناگون پژوهشی انجام می دهند.

از سال ۱۳۴۸ همکاری های پژوهشی و اطلاع رسانی با اجرای طرح امانت میان کتابخانه هادر ایرانداک آغاز گردید. این طرح هم اکنون به نام امین خوانده می شود. از سال ۱۳۷۴ نیز طرحی با نام غدیر (عضویت فراگیر کتابخانه ها) در این زمینه به انجام می رسد. فراهم آوری منابع علمی خارجی برای دانشگاه ها و پژوهشگاه ها از سال ۱۳۸۷ در شورای تأمین منابع علمی در ایرانداک دنبال می شود. ایرانداک با نهادهای جامع اطلاع رسانی مانند «ایفلا» و «ترمنت» نیز همکاری دارد. سامانه عرضه و تقاضای پژوهش هم پیوندی است که ایرانداک برای پشتیبانی از همکاری های پژوهشی ساخته است.

سیاست گذاری علم و فناوری در کشور به اطلاعاتی نیاز دارد که فراهم آوری بخشی از آن را این پژوهشگاه در سامانه ملی اطلاعات تحقیقاتی (سمات ملی) و بخشی را نیز در داشبورد مدیران انجام می دهد.

دبیرخانه کمیسیون مدیریت اطلاعات علم و فناوری هم که در قلمرو وزارت عتف، سیاست گذاری می کند در این پژوهشگاه است. پژوهشگاه علوم و فناوری اطلاعات ایران نماینده ایران در شبکه اطلاع رسانی آسیا و اقیانوسیه: «اپین» نیز هست که از سوی برنامه اطلاعات برای همه یونسکو هدایت می شود.

تولید ارزش افزوده و اطلاعات پشتیبان و همچنین تحلیل اطلاعات، بخشی دیگر از مدیریت دانش و سامانه های پیشینه پژوهش، همانندجو، دروازه اطلاعات علمی ایران، اصطلاح نامه، راهنمای پژوهش و پژوهشگران، و نمای مقاله های ایرانیان در جهان نمونه هایی از آنها هستند.

۱-۲- معرفی مرکز اطلاعات و مدارک علمی ایران

ماموریت های اصلی مرکز اطلاعات و مدارک علمی ایران فراهم آوری، سازمان دهی، ذخیره، حفظ، بازیابی، تحلیل و اشاعه اطلاعات و مدارک علمی و فناوری کشور و اطلاعات پشتیبان آنها در سطح ملی و بین المللی، مدیریت تامین و دسترسی پژوهشگران کشور به مدارک اطلاعات علم و فناوری خارج از کشور و پشتیبانی و مدیریت سامانه ملی مدیریت اطلاعات علم و فناوری، می باشند.

این مرکز دارای سه بخش می باشد:

- گروه ارتباطات و فراهم آوری
- مدیریت سازماندهی و تحلیل اطلاعات
- مدیریت حفظ و اشاعه اطلاعات

• گروه ارتباطات و فراهم آوری

این واحد مرحله ورودی و آغازین ورود منابع اطلاعاتی به مرکز اطلاعات می باشد ماموریت های اصلی گروه ارتباطات و فراهم آوری را «فراهم آوری اطلاعات و مدارک علم و فناوری» و «ثبت اطلاعات علم و فناوری کشور» تشکیل می دهد.

گروه ارتباطات و فراهم آوری اطلاعات از بخشهای زیر تشکیل یافته است:

- تهیه مدارک
- ورود اطلاعات
- کنترل اطلاعات
- دیجیتال سازی (اسکن)
- چاپخانه
- مخزن

• مدیریت سازماندهی و تحلیل اطلاعات

مأموریت مدیریت سازماندهی و تحلیل اطلاعات شامل: تحقیق و بررسی روشهای سازماندهی و طبقه‌بندی مدارک و اطلاعات، سازماندهی مدارک به منظور پردازش، تحلیل و بازیابی اطلاعات، انتخاب یکی از روشهای نمایه‌سازی و کنترل واژگان با اصطلاحنامه‌ها می باشد.

اهم وظایف:

- بررسی انواع مدارک و نمایه‌سازی آنها که از بخش تهیه مدارک ارسال می گردد.
- استفاده از اصطلاح‌نامه‌ها جهت انتخاب اصطلاحات مرجع و همچنین یکدست نمودن کلیدواژه‌ها.
- تلاش در جهت استانداردسازی نمایه‌سازی و استفاده از دستورالعمل‌های نمایه‌سازی

خدمات مدیریت ورود سازماندهی و تحلیل اطلاعات :

- گردآوری، سازماندهی، ذخیره، حفظ، بازیابی، تحلیل و اشاعه مدارک و اطلاعات علمی و فنی کشور از جمله پایان‌نامه‌های کارشناسی ارشد و رساله‌های دکترای تخصصی دانش‌آموختگان داخل و دانش‌آموختگان ایرانی خارج از کشور، مقاله‌های سمینارها، مشخصات پژوهشگران و متخصصان، طرحهای پژوهشی، اطلاعات مربوط به اختراعات، نرم افزارهای علمی و همچنین اطلاعات پشتیبان آنها از جمله اطلاعات سازمانها و افراد.

- مدیریت دسترسی به مدارک و اطلاعات علمی و فنی بین المللی مورد نیاز پژوهشگران.
- مدیریت دسترسی به مدارک و اطلاعات علمی و فنی گردآوری شده در پژوهشگاه در سطح ملی و جهانی.
- انتشار نمایه نامه ها، چکیده نامه ها، آمار نامه ها، فرهنگ نامه ها، و اصطلاح نامه ها.

• مدیریت حفظ و اشاعه اطلاعات

- سه مأموریت اصلی برای گروه حفظ و اشاعه اطلاعات تعریف شده است:
- نگهداری و حفظ اطلاعات و مدارک علم و فناوری ورودی به پژوهشگاه و سازمان دهی شده در پژوهشگاه؛
 - بازیابی و اشاعه اطلاعات و مدارک سازمان دهی شده؛
 - ایجاد دسترسی کاربران به اصل مدارک؛
 - اشاعه اطلاعات و تحویل نسخه های تکثیر شده مدارک در قالبها و از روشهای ممکن متناسب با زیرساختهای اطلاعاتی و ارتباطی کشور و نیاز کاربران.
- این واحد شامل سه بخش سالن جستجوی حضوری، سفارش مدرک غیر حضوری، گزارش پیشینه می باشد.

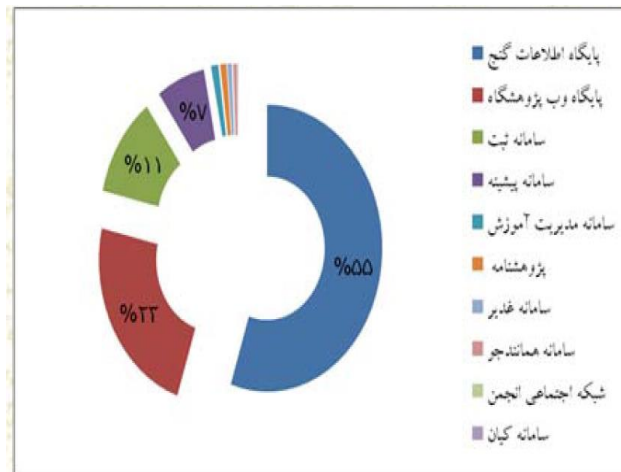
۱-۳- معرفی سامانه گنج

نرم افزار CDS/ISIS بعنوان قدیمی ترین نرم افزار در پژوهشگاه علوم و فناوری اطلاعات ایران از سال ۱۳۶۹ بستر اصلی تولید پایگاه اطلاعات علمی بوده است. این نرم افزار از سال ۱۳۷۲، عرضه شده و روزآمدی پایگاه های اطلاعات متنوعی در این محیط آغاز شده است. در سالهای اخیر با توجه به گسترش برنامه های رایانه ای در بستر وب، این پایگاه اطلاعات با نام سامانه گنج و در محیط MySQL ارائه شده است. سامانه گنج، گنجینه ای با ارزش از اطلاعات علمی و فراداده ای اساتید، دانشجویان و محققان ایرانی است. این پایگاه

علی‌رغم کارایی و اثر بخشی در ثبت و اشاعه اطلاعات علمی به کاربران، با معایب و اشکالاتی مواجه است و نیاز به بهینه سازی دارد تا اهداف راهبردی پژوهشگاه بهتر محقق شود. بخشی از نابسامانی‌ها در این سامانه ناشی از خطاهای انسانی است که به هنگام نمایه‌سازی و ورود اطلاعات در بخش سازماندهی اطلاعات صورت گرفته است. بخشی دیگر از ایرادها منشاء رایانه‌ای و سیستمی دارد که به مرور زمان و با تغییر نرم‌افزارها و سخت‌افزارها و کاراکترها و در هنگام تبدیل‌های مختلف و ورود ماشینی اطلاعات در پایگاه به وجود آمده است (علیدوستی ۱۳۷۵). فرایند فراهم آوری، سازماندهی و اشاعه اطلاعات علمی که منجر به تولید سامانه گنج می‌شود، فرایندی جاری، پیوسته و در حال انجام است. (علیدوستی و عصاره ۱۳۸۴)

۱-۴- تبیین موضوع پژوهش

امروزه سامانه گنج نقشی کلیدی در اشاعه علم و فناوری در کشور و ارائه خدمات به پژوهشگران در راستای انجام پژوهش‌های اثربخش و کارا برعهده دارد. این سامانه برطبق گزارش سامانه شدآمد بالغ بر ۵۰ درصد از بازدیدهای سامانه‌های ایرانداک را به خود اختصاص داده است. (شکل شماره ۱-۱ گزارش ارائه شده از میانگین بازدیدها در سامانه شدآمد را نمایش می‌دهد) با این توصیف کیفیت عملکرد این سامانه یکی از عوامل تاثیرگذار در انجام مناسب ماموریت‌های کلیدی آن محسوب می‌گردد. از سوی دیگر بخش عمده مشکلات کیفی مشاهده شده در حوزه اشاعه اطلاعات از سوی این سامانه بصورت ریشه‌ای به خروجی سامانه دیگری با عنوان سامانه ثبت برمی‌گردد که در مرحله آغازین کلان فرایند ثبت، سازماندهی و اشاعه اطلاعات قرار دارد.



شکل ۱-۱. میانگین سهم بازدید کنندگان سامانه‌های مختلف در پژوهشگاه علوم و فناوری اطلاعات

سامانه ثبت بمنظور گردآوری فراداده‌های پایان‌نامه/ رساله‌های دانشجویان داخل/ خارج کشور به گونه‌ای طراحی شده‌است که متشکل از فیلدها یا اقلام اطلاعاتی مختلفی از جمله نام، نام خانوادگی، شماره دانشجویی، کد ملی، مقطع تحصیلی، رشته دانشگاهی و .. است که به هر رکورد (سابقه) فایل مربوط به پایان‌نامه/ رساله آن دانشجو الصاق گردیده است. مجموعه داده‌های مورد اشاره قرار گرفته یک فراداده در سامانه ثبت خوانده می‌شود. تعیین وضعیت (یا سطح کیفیت) مطلوب در هر فیلد یک فراداده از جنبه‌های کیفی مختلف مانند دقت، صحت، نامتناقض بودن، کامل بودن و ... یکی از حیاتی‌ترین اقدامات در حوزه کیفیت سامانه ثبت به حساب می‌آید. به علاوه در صورتی که در فراداده‌های تکمیل شده توسط دانشجویان تحصیلات تکمیلی نقص یا به عبارتی مشکل کیفی مشاهده گردد عدم انطباق رخ داده بایستی به شکل مناسب درجه‌بندی شده و تصمیم‌گیری مناسب در آن خصوص صورت گیرد. لذا تعیین سطح کیفیت هر اقلام داده‌ای دسته‌بندی انواع عدم انطباق در اقلام داده‌ای یکی دیگر از اقدامات اصلی به نظر می‌آید. مجموعه این

دو اقدام کلیدی یعنی تعیین سطح کیفیت اقلام داده‌ای و درجه بندی نقص‌های مشاهده شده در اقلام داده‌ای و اقدامات واکنشی و اصلاحات بعدی مدل کیفیت فراداده در سامانه ثبت را تشکیل می‌دهند.

در پژوهش حاضر فراداده‌های گردآوری شده در سامانه ثبت حاوی اطلاعات پایان‌نامه‌ها/رساله‌های دانش‌آموختگان به طور ویژه مورد توجه قرار گرفته و مدل کیفیت این فراداده‌ها طراحی می‌گردد. در این راستا با کمک گرفتن از خبرگان حوزه ثبت و اشاعه مدارک علمی در ایرانداک و نیز با طراحی و بکارگیری پرسشنامه به تعیین ابعاد مختلف کیفیت در سامانه ثبت پرداخته خواهد شد. استفاده و بومی‌سازی مدل‌های امروزی در کیفیت داده مانند (TDQM^۱)، (DQA^۲)، (DQM^۳) و ... در تدوین مدل کیفیت یکی از مهم‌ترین دستاوردهای این پژوهش خواهد بود. همچنین چک‌لیست‌های کنترل کیفی و شرح وظایف هریک از همکاران در واحد فراهم‌آوری در کنترل اقلام اطلاعاتی فراداده‌های مربوط به پایان‌نامه/رساله‌ها یکی دیگر از نتایج اجرای طرح حاضر خواهد بود.

^۱ Total Quality Data Management

^۲ Data Quality Assessment

^۳ Data warehouse Quality Methodology

فصل دوم – معرفی مبانی کیفیت داده و اطلاعات

این پژوهش درباره‌ی تعدادی از روش‌ها برای ارزیابی و بهبود کیفیت داده‌هاست. به علت تنوع و پیچیدگی این روش‌ها، این تحقیق بر روی روش‌هایی تمرکز کرده است که در انتخاب، هماهنگ‌سازی و کاربرد روش‌های ارزیابی و بهبود کیفیت داده‌ها به ما کمک می‌کنند. هدف این بخش تشریح سیستماتیک و مقایسه‌ی این روش‌هاست. روش‌ها در ابعاد متفاوتی همچون فازها و مراحل اجرایی، استراتژی‌ها و فنون، ابعاد کیفیتی داده‌ها و نوع سیستم آدرس‌دهی اطلاعات بررسی می‌شوند. همچنین در انتهای این بخش توضیح خلاصه‌ای از هر روش ذکر شده است.

۱-۲ مقدمه‌ای در مورد کیفیت داده‌ها

به علت پیشرفت داده‌های الکترونیکی، کیفیت داده‌ها نقش مهمی در همه کسب و کارها و برنامه‌های دولتی ایفا می‌کند. کیفیت داده‌ها به عنوان یک مسئله عملکردی و ضروری در ارتباط با فرآیندهای عملیاتی (Data Warehousing Institute 2006)، تصمیم‌گیری‌ها (Chengalur-Smith et al. 1999) و همکاری‌های درون سازمانی (Scannapieco & Batini ۲۰۰۲) شناخته شده است. چندین طرح ابتکاری در بخش‌های دولتی و خصوصی راه‌اندازی شده است که در آن کیفیت داده‌ها نقش برجسته‌ای دارد، از جمله این طرح‌ها می‌توان به تصویب قانون کیفیت داده‌ها توسط دولت ایالات متحده در سال ۲۰۰۲ (Office of Management and Budget 2006) و تصویب چارچوب کیفیت داده‌ها توسط دولت Wales در سال ۲۰۰۴ را نام برد که هدف آن‌ها بهبود کیفیت اطلاعات برای شیوه‌های پزشکی عمومی است (DQI^۱ 2004).

با حرکت سیستم‌های اطلاعاتی از ساختار سلسله مراتبی / یکپارچه به ساختار مبتنی بر شبکه، مجموعه منابع داده‌ی بالقوه مورد استفاده در سازمان‌ها، از نظر اندازه و دامنه به طور چشمگیری افزایش می‌یابند. مسئله کیفیت داده‌ها نیز به خاطر اثر گیری از این تکامل، پیچیده‌تر و بحث‌برانگیزتر شده است. در سیستم‌های اطلاعات شبکه‌ای، فرآیندها در تبادل اطلاعات پیچیده‌ای درگیر شده و اغلب بر اساس ورودی به دست آمده از منابع خارجی عمل می‌کنند، که این ورودی‌ها نیز اغلب قابل مقایسه نیستند. در نتیجه اگر کیفیت هر فرآیند و

^۱ Data Quality Information

ورودی اطلاعات کنترل شده نباشد، کیفیت کلی داده‌هایی که در سراسر سیستم اطلاعات جریان دارند می‌تواند به سرعت در طول زمان کاهش یابند. از طرف دیگر، سیستم‌های اطلاعات شبکه‌ای، فرصت‌های جدیدی را برای مدیریت کیفیت داده‌ها ارائه می‌دهند که می‌توان دسترسی به طیف وسیع‌تری از منابع داده‌ها، توانایی انتخاب و مقایسه اطلاعات از منابع مختلف به منظور تشخیص و تصحیح خطاها و در نتیجه، بهبود کلی کیفیت داده‌ها را از نتایج آن دانست.

این متون طیف گسترده‌ای از تکنیک‌های ارزیابی و بهبود کیفیت داده‌ها را برای اموری همچون پیوند مدرک و سابقه، قوانین کسب و کار و مقیاس‌های هماهنگی را ارائه می‌دهد. با گذشت زمان، این تکنیک‌ها برای مقابله با افزایش پیچیدگی کیفیت داده‌ها در سیستم‌های اطلاعات شبکه‌ای تکامل یافتند. با توجه به تنوع و پیچیدگی این تکنیک‌ها، تحقیقات اخیر بر روی تعریف‌های روش‌شناختی که به انتخاب، سفارشی‌سازی و اعمال تکنیک‌های ارزیابی و بهبود کیفیت داده‌ها کمک می‌کنند، متمرکز شده‌اند. این پژوهش روش کیفیت داده را به عنوان مجموعه‌ای از دستورالعمل‌ها و تکنیک‌هایی تعریف می‌کند که با توصیف اطلاعات ورودی در یک زمینه کاربردی شروع شده و فرآیند منطقی را برای ارزیابی و بهبود کیفیت داده‌ها بیان می‌کند. هدف این بخش، ارائه‌ی یک توصیف منظم و مقایسه‌ای در مورد روش‌های کیفیت داده‌های موجود است.

این بخش به شکل زیر تدوین شده است. در ابتدا در مورد کیفیت داده‌ها توضیحاتی بیان می‌شود. در بخش بخش بعدی موضوعات اصلی مشترک در کیفیت داده برای همه روش‌ها بیان می‌شود. دیدگاه‌های مورد استفاده در این بخش برای تحلیل‌های مقایسه‌ای به صورت زیر هستند:

۱. مراحل و گام‌های روش‌شناختی

۲. استراتژی‌ها و تکنیک‌ها

۳. ابعاد کیفیت داده

۴. انواع داده و سرانجام

۵. انواع سیستم‌های اطلاعاتی.

در زیر بخش ۳، روش‌شناسی‌های موجود در راستای ابعاد معرفی شده در زیر بخش ۲ مقایسه می‌شوند. مقایسه کردن با ارزیابی‌های کیفی و بررسی نظرات در مورد جدول‌های هم‌نظیر (که در یک گروه کلی از روش‌هایی با رویکردهای مشابه برجسته می‌شود) انجام می‌شود.

در زیربخش ۴ تحقیقات در حال اجرا و آینده‌ی این حوزه در زمینه‌ی روش‌های کیفیت داده‌ها بررسی می‌شود. در نهایت، این بخش با شرح مختصری از هر یک از روش‌ها در پیوست به پایان می‌رسد که موارد زیر را توصیف می‌کند:

۱. مراحل هر یک از روش‌ها، وابستگی متقابل و تصمیم‌گیری‌های حیاتی هر روش بررسی می‌شود.
۲. توصیف کلی، که تمرکز هر روش و مشارکت اصلی آن برای فرآیند ارزیابی و بهبود کیفیت داده‌ها را برجسته می‌کند.
۳. نظرات دقیق که در مورد کاربرد هر یک از این روش‌ها بررسی می‌شود.

۲-۲ مقایسه‌ی دیدگاه‌ها در تجزیه و تحلیل روش‌ها

دیدگاه‌های مختلفی وجود دارد که می‌توانند برای تجزیه و تحلیل و مقایسه روش‌های کیفیت داده‌ها (DQ) مورد استفاده قرار گیرند، برخی از آن‌ها عبارت‌اند از:

۱. مراحل و گام‌هایی که روش را تشکیل می‌دهند؛
۲. استراتژی و تکنیک‌هایی که در روش برای ارزیابی و بهبود سطح کیفیت داده اتخاذ شده‌اند.
۳. ابعاد و معیارهایی که در این روش‌ها برای ارزیابی سطح کیفیت داده‌ها انتخاب شده‌اند؛
۴. انواع هزینه‌هایی که با مسائل مربوط به کیفیت داده‌ها در ارتباط هستند از جمله:
۱,۴. هزینه‌های مربوط به کیفیت داده ضعیف، که هزینه‌های فرآیند ناشی از خطای داده‌ها و هزینه‌های فرصت به علت درآمد از دست رفته و زیان دیده است؛ این هزینه‌ها به هزینه‌های غیرمستقیم نیز اشاره می‌کنند؛

- ۲,۴. هزینه‌های فعالیت‌های ارزیابی و بهبود که به‌عنوان هزینه‌های مستقیم نیز نام‌برده می‌شوند؛
۵. انواع داده‌هایی که در روش‌ها در نظر گرفته شده‌اند.
۶. انواع سیستم‌های اطلاعاتی برای استفاده، اصلاح و مدیریت داده‌ها که در روش در نظر گرفته شده است.
۷. سازمان‌های درگیر در فرآیندها که به ایجاد یا به‌روزرسانی داده‌ها می‌پردازند و با ساختار و اصول خاص خود در نظر گرفته شده‌اند.
۸. فرآیندهایی که داده‌ها را باهدف تولید خدمات موردنیاز کاربران ایجاد و بروز رسانی می‌کنند، در نظر گرفته شده‌اند.
۹. خدماتی است که توسط فرآیندها تولید شده نیز در نظر گرفته شده‌اند.
- این روش‌ها از نظر نحوه‌ی کاربرد در ابعاد فوق متفاوت هستند. در ادامه این بخش سه دیدگاه آخر یعنی دیدگاه سازمانی، فرآیندی و خدماتی بررسی نخواهد شد، چراکه به‌ندرت در روش‌ها ذکر شده‌اند.

۱-۲-۲ مراحل و گام‌های مشترک

در حالت کلی توالی فعالیت‌های یک روش کیفیت داده از سه گام به‌صورت زیر تشکیل شده است:

۱. وضعیت سازی (که هدف آن جمع‌آوری اطلاعات مبتنی بر اساس فرآیندها و خدمات سازمانی است)؛ این گام بررسی مجموعه داده‌ها و روندهای مرتبط با مدیریت، مسائل کیفیت و هزینه‌های مربوطه است و اگر اطلاعات متنی از تجزیه و تحلیل‌های قبلی در دسترس باشد می‌توان این گام را نادیده گرفت.
۲. ارزیابی / اندازه‌گیری (که مجموعه کیفیت داده‌ها را در ابعاد کیفیت مربوطه اندازه‌گیری می‌کند)؛ اصطلاح اندازه‌گیری برای پرداختن به موضوع اندازه‌گیری ارزش مجموعه ابعاد کیفیت داده‌ها استفاده می‌شود. اصطلاح ارزیابی زمانی استفاده می‌شود که چنین اندازه‌گیری‌هایی با مقادیر مرجع

مقایسه شوند تا قادر به تشخیص کیفیت شوند. در این پژوهش عبارت ارزیابی برای اکثر روش‌هایی اتخاذ شده است که در آن‌ها بر اهمیت علل کیفیت داده‌های ضعیف تأکید شده است.

۳. بهبود فرآیند که مربوط به انتخاب مراحل، استراتژی‌ها و تکنیک‌ها برای رسیدن به اهداف کیفیت داده‌های جدید تعریف می‌شود.

اگر گام ارزیابی بر اساس اسناد و مدارک موجود باشد، انجام آن حالت اختیاری خواهد داشت. از آنجا که معمولاً این فرض در مورد داده‌ها وجود دارد، در مورد گام وضعیت سازی بحث نمی‌کنیم. معمولاً مراحل ارزیابی و بهبود را در مجموعه‌ی مشترک از گام‌های اساسی سازمان‌دهی می‌کنند، هرچند برای هر مرحله نام‌های مختلفی اتخاذ شده است اما به‌طور کلی گام‌های مرحله‌ی ارزیابی عبارت‌اند از:

- تجزیه و تحلیل داده‌ها، که به بررسی طرح داده‌ها و انجام مصاحبه برای رسیدن به درک کاملی از داده‌های مرتبط باهم و قوانین موجود در معماری و مدیریت داده‌ها می‌پردازد؛
- تجزیه و تحلیل الزامات کیفیت اطلاعات، که به بررسی نظر کاربران داده‌ها و مدیران برای شناسایی مسائل کیفیت و تعیین اهداف جدید می‌پردازد.
- شناسایی مناطق بحرانی، که به انتخاب و ارزیابی کمی مرتبط‌ترین پایگاه داده و جریان داده می‌پردازد.
- مدل‌سازی فرآیند، که مدلی از فرآیندهای تولید یا به‌روزرسانی داده‌ها را فراهم می‌کند؛
- اندازه‌گیری کیفیت، در این گام بر اساس مسائل کیفیت مشخص شده و اثرگذار در مرحله‌ی تجزیه و تحلیل الزامات کیفیت اطلاعات، ابعاد کیفیتی انتخاب شده و معیارهای مربوط به آن‌ها تعریف می‌شوند؛ زمانی که این مسائل مبتنی بر معیارهای کمی باشند اندازه‌گیری می‌تواند هدف باشد اما هنگامی که این مسائل مبتنی بر ارزیابی‌های کیفی توسط مدیران داده‌ها و کاربران باشند اندازه‌گیری می‌تواند ذهنی باشد.

مراحل گام بهبود عبارت‌اند از:

- بررسی هزینه‌ها، که هزینه‌های مستقیم و غیرمستقیم کیفیت داده‌ها را برآورد می‌کند؛
 - تعیین مسئولیت‌های فرآیند، که به شناسایی صاحبان فرآیند و تعریف مسئولیت آن‌ها بر اساس فعالیت‌های تولید و مدیریت داده‌ها می‌پردازد؛
 - تعیین مسئولیت داده‌ها، که به شناسایی صاحبان داده‌ها و تعریف مسئولیت‌های مدیریت داده‌ی آن‌ها می‌پردازد.
 - شناسایی علل خطاها، که علت مشکلات کیفیت را شناسایی می‌کند.
 - انتخاب استراتژی‌ها و تکنیک‌ها، که تمام استراتژی‌های بهبود داده‌ها و تکنیک‌های مربوطه را شناسایی کرده و مطابق با دانش موجود، اهداف کیفیت و محدودیت‌های بودجه به بررسی و انتخاب استراتژی می‌پردازد؛
 - طراحی راه‌حل‌های بهبود داده، که مؤثرترین و کارآمدترین استراتژی‌ها، مجموعه تکنیک‌ها و ابزارهای مرتبط به بهبود کیفیت داده‌ها را انتخاب می‌کند؛
 - کنترل فرآیند، که نقاط بررسی را در فرآیندهای تولید داده‌ها، برای نظارت بر کیفیت آن‌ها در طول اجرای فرآیند تعریف می‌کند.
 - طراحی مجدد فرآیند، تعریف اقدامات بهبود فرآیند که می‌تواند پیشرفت‌های مربوط به کیفیت اطلاعات را ارائه دهد؛
 - مدیریت بهبود، که قوانین سازمانی جدید را برای کیفیت داده‌ها تعریف می‌کند.
 - نظارت بر بهبود، این گام به ارائه فعالیت‌های نظارت دوره‌ای می‌پردازد که بازخوردی از نتایج حاصل از فرآیند بهبود و تنظیم پویا را فراهم می‌کند.
- در بخش‌های بعدی روش‌ها با توجه به توانایی‌های ارزیابی و بهبود آن‌ها در طول مجموعه‌ای از مراحل و گام‌های معرفی شده مقایسه می‌شوند. توجه داشته باشید که معمولاً هر یک از روش‌ها به یک قابلیت ارزیابی و یا بهبود خاص در شرایط مختلف اشاره می‌کنند.

۲-۲-۲ استراتژی‌ها و تکنیک‌ها

در مراحل بهبود، روش‌ها دو نوع استراتژی کلی اتخاذ می‌کنند:

۱. استراتژی داده محور

۲. استراتژی فرآیند محور

استراتژی‌های داده محور، کیفیت داده‌ها را با اصلاح مستقیم ارزش داده‌ها بهبود می‌دهند. به عنوان مثال، ارزش داده‌های منسوخ شده با به روزرسانی آن‌ها از یک پایگاه داده همراه با داده‌ها و اعمال آن بر پایگاه داده کنونی بهتر می‌شود. استراتژی‌های فرآیند محور کیفیت را توسط طراحی مجدد فرآیندهایی که داده را ایجاد و یا تغییر می‌دهند، بهبود می‌بخشند. به عنوان مثال، یک فرآیند را می‌توان با گنجاندن فعالیت‌هایی که فرمت داده‌ها را قبل از ذخیره‌سازی کنترل می‌کنند، دوباره طراحی کرد.

استراتژی‌های داده محور و فرآیند محور، انواع مختلفی از تکنیک‌ها مانند الگوریتم‌ها، فن‌آوری هوشمند و فعالیت‌های مبتنی بر دانش را اعمال می‌کنند که هدف آن‌ها بهبود کیفیت داده می‌باشد. فهرستی از تکنیک‌های بهبود مرتبط با استراتژی‌های مبتنی بر داده، به صورت زیر است:

۱. دستیابی به داده‌های جدید، که باعث بهبود داده‌های قدیم و کسب کیفیت بالاتر آن‌ها می‌شود و این

تکنیک بیشتر به جای روش‌هایی است که موجب مشکلات کیفیتی در داده‌ها می‌شوند.

۲. استانداردسازی (یا نرمالیزه کردن)، که مقادیر داده‌های غیراستاندارد را با مقادیری که با استاندارد

مطابقت دارد، جایگزین یا کامل می‌کند. به عنوان مثال، نام مستعار با نام مربوطه جایگزین می‌شود،

مثلاً، باب با رابرت، و عبارات اختصاری متناظر با نام کامل جایگزین می‌شوند، برای مثال،

Channel Str با خیابان کانال جایگزین می‌شوند.

۳. ثبت ارتباط، که نمایش اطلاعات را در دو (یا چند) جدول مشخص می‌کند که ممکن است به هدفی

یکسان در دنیای واقعی اشاره کنند؛

۴. یکپارچه‌سازی داده‌ها و طرح‌ها، که یک دیدگاه یکپارچه از اطلاعات ارائه‌شده توسط منابع داده‌های ناهمگن را تعریف می‌کند. هدف اصلی ادغام و یکپارچه‌سازی، دسترسی یک کاربر به داده‌های ذخیره‌شده توسط منابع داده ناهمگن و از طریق دیدگاه یکپارچه این داده‌ها است.

سیستم‌های اطلاعاتی توزیعی، تعاونی، P2P (بخش ۶,۲) و منابع داده با در نظرگیری انواع حالت‌ها و ناهمگنی‌ها را می‌توان به‌طور کلی به موارد زیر دسته‌بندی کرد:

۱,۴. ناهمگنی در فن‌آوری

ناهمگنی فن‌آوری با توجه به استفاده‌ی محصولات توسط فروشندگان مختلف، در لایه‌های مختلف زیرساخت‌های اطلاعاتی و ارتباطاتی ایجاد می‌شوند.

۲,۴. ناهمگنی در طرح

ناهمگنی در طرح به دو صورت ایجاد می‌شوند:

۱,۲,۴. در درجه اول به خاطر استفاده از مدل‌های داده‌های مختلف ایجاد می‌شود (همانند منبعی که مدل داده رابطه‌ای را اتخاذ کرده و منبع‌های مختلفی که مدل داده‌های XML را پذیرفته‌اند)

۲,۲,۴. نمایش‌های مختلف برای همان شی مانند دو منبع رابطه‌ای که یک شی را به‌عنوان یک جدول و یک ویژگی نشان می‌دهند.

۳,۴. ناهمگنی در سطح نمونه

ناهمگنی در سطح نمونه توسط مقادیر داده مختلف و متضاد ایجاد می‌شود، که توسط منابع مجزا برای اشیاء مشابه ارائه می‌شود. به‌عنوان مثال، این نوع عدم تجانس یا ناهمگنی را می‌توان با فرآیندهای مستقل و ضعیف که منابع داده مختلف را تغذیه می‌کنند هماهنگ کرد. یکپارچه‌سازی داده‌ها باید با تمام این ناهمگنی‌های ذکرشده برخورد کند.

۵. اعتماد به منبع، که منابع داده را بر اساس کیفیت داده‌های خود انتخاب می‌کند.

۶. موضع یابی و تصحیح خطا، که خطاهای کیفیت داده را با تشخیص گزارش‌هایی که مجموعه قوانین باکیفیت را برآورده نکرده‌اند، شناسایی و از بین می‌برد. این تکنیک‌ها عمدتاً در حوزه‌های آماری مورد مطالعه قرار می‌گیرند. در مقایسه با داده‌های ابتدایی، داده‌های آماری مانند میانگین، مجموع، حداکثر و ... حساسیت کمتری نسبت به موضع یابی احتمالی نادرست و اصلاح مقادیر دارند. تکنیک‌هایی برای موضع یابی و تصحیح خطا در جهت رفع تناقضات، اطلاعات ناقص و دورافتاده ارائه شده است (Johnson, Dasu 2003)، (Batini, Scannapieco 2006).

۷. بهینه‌سازی هزینه‌ها، اقدامات بهبود کیفیت را همراه با مجموعه‌ای از ابعاد برای به حداقل رساندن هزینه‌ها تعریف می‌کند.

دو تکنیک اصلی برای مشخص کردن استراتژی‌های فرآیند محور:

- بررسی و چک کردن روندهای کنترل فرآیند تولید داده‌ها هنگامی که:

۱. داده‌های جدید ایجاد می‌شوند

۲. مجموعه داده‌ها به‌روز شده

۳. مجموعه داده‌های جدید با روند مشخص قابل دسترسی باشند.

در این روش، استراتژی واکنشی برای اصلاح داده‌ها، اجتناب از تخریب داده‌ها و انتشار خطا استفاده می‌شود.

- فرآیند طراحی مجدد، به طراحی دوباره فرآیندها، حذف علل پایین بودن کیفیت و معرفی فعالیت‌های جدید می‌پردازد که نهایتاً منجر به تولید داده‌هایی باکیفیت بالاتر می‌شود. اگر فرآیند طراحی مجدد اساس کار باشد، این تکنیک به مهندسی مجدد فرآیند کسب و کار اشاره می‌کند (Champy, Hammer 2001)؛ (Stoica et al. 2003).

در مطالعه (Redman 1996) چند تکنیک معمولی از استراتژی‌های داده محور و فرآیند محور با بحث در مورد بهبودی که هر یک از این روش می‌توانند در ابعاد کیفیتی مختلف و هزینه‌ی اجرای روش داشته باشند،

مقایسه شده‌اند. این مقایسه در دیدگاه‌های کوتاه‌مدت و بلندمدت انجام شده است. مقایسات بر موارد زیر متمرکز می‌باشد:

۱. کسب داده‌های جدید
۲. ثبت ارتباط
۳. موضع یابی (محل‌ی سازی) و تصحیح خطا
۴. کنترل فرآیند
۵. تکنیک‌های طراحی مجدد فرآیند.

به‌طور کلی، تکنیک‌های فرآیند محور عملکرد بهتری نسبت به تکنیک‌های داده محور در درازمدت دارند، چراکه آن‌ها علل ریشه‌ای مشکلات کیفیت را از بین می‌برند. اما از دیدگاه کوتاه‌مدت، طراحی مجدد فرآیند می‌تواند بسیار گران باشد (Redman 1996) (English 1999). در مقابل، استراتژی‌های داده محور در کوتاه‌مدت به صرفه هستند، اما در بلندمدت گران هستند. آن‌ها برای کاربردهای تک زمانه مناسب هستند به همین دلیل برای داده‌های ایستا توصیه می‌شوند.

۲-۲-۳ ابعاد

در تمام روش‌ها، تعریف ابعاد کیفیت و معیارهای ارزیابی داده‌ها، فعالیتی اساسی می‌باشد. به‌طور کلی، معیارهای مختلف را می‌توان با هر یک از ابعاد کیفیت مرتبط کرد. در برخی موارد معیار منحصر به فرد است و تعریف نظری از ابعادی منطبق با تعریف عملیاتی معیار مربوطه است. به همین دلیل، میان تعاریف نظری و عملیاتی ابعاد تنها زمانی که معیارهای متعددی فراهم باشند، تفاوت می‌گذاریم.

ابعاد کیفیت را می‌توان برای گسترش داده‌ها (مقادیر کمی و یا کیفی داده‌ها) بر اساس طرحی که دارند بکار برد. اگرچه طرح کیفیت داده‌های مفهومی و منطقی به عنوان یک حوزه پژوهشی به رسمیت شناخته شده است،

اما اغلب تعاریف ابعاد و معیارهای کیفیت داده‌ها متضاد با تعاریف طرح‌ها تعیین شده‌اند. این پژوهش عمدتاً بر ابعاد کیفیتی و معیارهایی متمرکز است که به مقادیر داده اشاره می‌کنند.

تحقیقات مربوط به کیفیت داده، طبقه‌بندی کاملی از ابعاد کیفیت داده‌ها را فراهم می‌کنند؛ با این حال برخی اختلافات در تعریف مهم‌ترین ابعاد با توجه به ماهیت تعریف شده از کیفیت وجود دارد. شش طبقه مهم از ابعاد کیفیت توسط Wang (۱۹۹۶)؛ Wang و Strong (۱۹۹۶)؛ Redman (۱۹۹۶)؛ Jarke و همکاران (۱۹۹۵)؛ Bovee و همکاران (۲۰۰۱) و Naumann (۲۰۰۲) ارائه شده است. با تجزیه و تحلیل این طبقه‌بندی، تعریف مجموعه‌ای اولیه‌ای از ابعاد کیفیت داده‌ها امکان‌پذیر می‌شود، از جمله آن‌ها دقت، جامعیت، ثبات، و به موقع بودن است که مورد تمرکز اکثر نویسندگان می‌باشد.

با این حال، هیچ توافق کلی بر روی تعریف مجموعه‌ای از ابعاد کیفیت داده‌ها و یا معنای دقیق هر یک از ابعاد وجود ندارد. تعاریف مختلف ارائه شده در متون در زیر مورد بحث قرار گرفته است.

دقت. تعاریف متعددی برای عبارت دقت ارائه شده است. Wang و Strong (۱۹۹۶) دقت را به صورت زیر تعریف کرده‌اند "میزان اطلاعات صحیح، قابل اعتماد و تأیید شده". Ballou و Pazer بیان کردند داده‌ها زمانی دقیق هستند که مقادیر داده ذخیره شده در پایگاه داده با مقادیر دنیای واقعی مطابقت داشته باشد. (Ballou, Pazer 1985) در مطالعه Redman (۱۹۹۶)، دقت میزان نزدیکی یک مقدار داده V ، به برخی مقادیر دیگر V' ، که درست در نظر گرفته شده است، می‌باشد.

جدول ۱-۲ عمده‌ترین تحقیقاتی را ارائه می‌کند که به ارائه تعریف دقت می‌پردازند. با مقایسه این تعاریف، می‌توان توافق قابل توجهی در تعریف انتزاعی جامعیت را مشاهده کرد.

جدول ۲-۱. تعاریف مربوط به دقت در مراجع مختلف

منبع	تعریف
Wand and Wang 1996	توانایی یک سیستم اطلاعات برای نشان دادن هر حالت معنی دار از یک سیستم دنیای واقعی
Wang and Wand 1996	داده ها با توجه به طول، عرض و دامنه برای حوزه مورد نظر کافی باشند
Redman 1996	درجه ای که در آن مقادیر ارقام اطلاعاتی در یک مجموعه داده های مورد نظر قرار می گیرند
Jarke et al. 1995	درصدی از اطلاعات دنیای واقعی که در منابع داده ها و / یا انبار داده وارد شده است
Bovee et al. 2001	میزانی که اطلاعات اجزاء مورد نیاز (ضروری) ^۱ از یک نهاد ^۲ را شامل می شوند
Liu and Chi 2002	نسبتی از مقادیر جمع آوری شده در منابع اطلاعاتی به کل مقادیر دنیای واقعی

جدول ۲-۲. تعاریف مختلف ارائه شده برای جامعیت

منبع	تعریف
Wand and Wang 1996	توانایی یک سیستم اطلاعاتی برای نشان دادن هر حالت معنی دار از یک سیستم دنیای واقعی
Wang and Wand 1996	اینکه داده ها با توجه به طول، عرض و دامنه برای کار در دست کافی باشند
Redman 1996	درجه ای که در آن ارزش ها در یک مجموعه داده قرار می گیرند
Jarke et al. 1995	درصد اطلاعات وارد شده در دنیای واقعی به منابع داده ها و یا انبار داده
Bovee et al. 2001	اطلاعاتی که تمام ابعاد مورد نیاز برای توضیح یک نهاد را دارند
Naumann 2002	نسبت بین تعداد مقادیر غیر تهی در یک منبع و اندازه رابطه جامع
Liu and Chi 2002	تمام ارزش های مورد حمایت که در هر یک نظریه جمع آوری می شود.

^۱ Required

^۲ Entity

به طور کلی دو نوع دقت یعنی دقت نحوی و دقت معنایی را می توان از هم متمایز کرد. روش کیفیت داده تنها دقت نحوی را در نظر گرفته و به عنوان نزدیکی مقدار v ، به عناصر دامنه متناظر، D تعریف می کند. در دقت نحوی به دنبال مقایسه مقدار v با مقدار v' دنیای واقعی آن نیستیم، بلکه علاقه مندیم بررسی کنیم که آیا v یکی از مقادیر D است یا خیر، و یا اینکه چقدر به مقادیر D نزدیک است. به عنوان مثال، $v = \text{'John'}$ از لحاظ دستوری دقیق در نظر گرفته می شود حتی اگر $v = \text{'John'}$ باشد.

جامعیت. جامعیت یعنی تا چه اندازه مجموعه داده های معلوم شامل توصیف داده ی مربوط به مجموعه ای از اشیاء دنیای واقعی هستند.

جدول ۲-۲ تحقیقاتی را بیان می کند که به ارائه ی تعریف جامعیت پرداخته اند. با مقایسه ی این تعاریف، می توان توافق قابل توجهی در تعریف انتزاعی جامعیت مشاهده کرد. تعاریف در حوزه هایی که به آن اشاره کرده اند متفاوت می باشند، به عنوان مثال حوزه ی سیستم اطلاعات در Wand و Wang (۱۹۹۶)، انبار داده ها در Jarke و همکاران (۱۹۹۵)، نهاد در Bovee و همکاران (۲۰۰۱) مورد نظر می باشند.

معمولاً در حوزه ی پژوهش پایگاه داده های رابطه ای، جامعیت مربوط به معنای مقادیر null (باطل - صفر) می باشد. مقدار null در معنی عام ارزش از دست رفته هستند، ارزشی که در دنیای واقعی وجود دارد، اما در یک مجموعه داده های غیر قابل دسترس قرار دارد. به منظور تعیین مشخصات جامعیت، درک اینکه چرا این ارزش از دست رفته است، مهم می باشد. یک مقدار به چند دلیل از دست می رود، یکی اینکه آن ها وجود دارند اما شناخته شده نیستند، یا به دلیل آن که اصلاً وجود ندارد و یا مشخص نشده اند که آیا وجود دارند یا خیر (Antonellis, Atzeni 1993).

فرض کنید در شکل ۱-۲، فرد با ویژگی های نام، نام خانوادگی، تاریخ تولد، و ایمیل در نظر گرفته شده است. اما شخص ۲ درجایی از جدول بدون ایمیل است و درجایی کامل است. حال اگر فرد ۴ شناخته شده نباشد ممکن است جایی دارای ایمیل و جای دیگر ناقص باشد.

ثبات. بعد ثبات به نقض قوانین معنایی تعریف شده بر روی مجموعه ای از اطلاعات اشاره دارد. با اشاره به نظریه رابطه‌ای، محدودیت یکپارچگی با سایر اقلام اطلاعاتی دسته‌ای از قوانین معنایی است. در نظریه رابطه‌ای، دو دسته اساسی مربوط به محدودیت یکپارچگی را می‌توان متمایز کرد، این دو دسته عبارتند از: محدودیت‌های درون رابطه‌ای و محدودیت‌های بین رابطه‌ای. محدودیت درون رابطه‌ای طیف وسیعی از مقادیر مجاز را برای دامنه یک ویژگی تعریف می‌کند. مثال‌ها به صورت زیر هستند "سن باید بین محدوده ۰ و ۱۲۰ باشد"، و یا "اگر سابقه کاری کمتر از ۳ سال است، سپس حقوق می‌تواند بالاتر از ۲۵,۰۰۰ یورو در سال باشد" محدودیت یکپارچگی بین رابطه‌ای شامل انتساب روابط متفاوت است. ارتباط میان دانشگاه، رشته و ... از محدودیت‌های میان رابطه‌ای است.

در نظریه رابطه‌ای، دودسته اساسی مربوط به محدودیت یکپارچگی را می‌توان متمایز کرد، این دودسته عبارت‌اند از:

- محدودیت‌های درون رابطه‌ای
- محدودیت‌های بین رابطه‌ای

محدودیت درون رابطه‌ای^۱ طیف وسیعی از مقادیر مجاز را برای دامنه یک ویژگی تعریف می‌کند. مثلاً داریم:

"سن باید بین محدوده ۰ و ۱۲۰ باشد" و یا "اگر سابقه کاری کمتر از ۳ سال است، سپس حقوق می‌تواند بالاتر از ۲۵,۰۰۰ یورو در سال باشد"

ID	Name	Surname	BrithDate	Email
1	John	Smith	1974/17/3	smith@abc.it
2	Edward	Mornoe	1967/3/2	NULL
3	Anthony	White	1936/1/1	NULL
4	Marianne	Collines	1955/20/11	NULL

not existing

existing but
unknown

not known
if existing

شکل ۲-۱. مقادیر NULL و جامعیت داده‌ها (BATINI ET AL. 2009)

جدول ۲-۳. تعاریف موجود برای ابعاد مرتبط به زمان

منبع	تعریف
Wand and Wang 1996	به موقع بودن تنها به تأخیر بین تغییر یک حالت دنیای واقعی و اصلاح ناشی از حالت سیستم اطلاعات اشاره دارد
Wang and Wand 1996	به موقع بودن یعنی تا چه حد سن داده‌ها برای کار در دست مناسب هستند
Redman 1996	نرخ ارز، درجه‌ای اینکه داده چقدر به روز است. ارزش داده به روز است اگر باوجود اختلاف ممکن ناشی از تغییرات مربوط به زمان تا مقدار صحیح، صحیح باشد
Jarke et al. 1995	ارز زمانی توصیف می‌شود که اطلاعات در منابع و / یا انبار داده‌ها وارد شوند. نوسانات، دوره زمانی را توصیف می‌کند که اطلاعات آن در دنیای واقعی، معتبر باشد
Bovee et al. 2001	به موقع بودن دو مؤلفه دارد: سن و نوسانات. سن و یا ارز معیار چندسالگی اطلاعات، بر اساس مدت سنی است که ثبت شده است. فراریت معیار بی‌ثباتی اطلاعات، فرکانس تغییر مقدار برای یک ویژگی نهادی است
Naumann 2002	به موقع بودن، متوسط سن داده‌ها در یک منبع است
Liu and Chi 2002	به موقع بودن، میزان به روز بودن کافی داده‌ها برای یک کار است

محدودیت یکپارچگی بین رابطه‌ای^۱ شامل انتساب روابط متفاوت است. به عنوان مثال، فرض کنید رابطه فیلم‌هایی را در نظر بگیریم که عنوان، مدیر و سال را داریم و می‌خواهیم رابطه آن‌ها با جایزه اسکاری که برنده می‌شوند را بیابیم. (در واقع دو جدول و پایگاه داده مجزا)

محدودیت‌های بین رابطه‌ای می‌تواند برای هر فیلم ظاهر شده در هر طرف رابطه بیان شود، مثلاً "سال فیلم باید برابر با سال فیلم اسکار گرفته باشد." متون گسترده‌ای در مورد پایگاه داده‌های سازگار وجود دارد، مثلاً Arenas و همکاران (۱۹۹۹). واکنش مداوم در یک پایگاه داده رابطه‌ای مشکل خصوصیات منطقی، که ممکن است محدودیت یکپارچگی داده‌شده را نقض کند، نشان می‌دهند. آن‌ها روشی را برای محاسبه پاسخ‌های سازگار، با اثبات صحت و جامعیت داده‌ها ارائه می‌دهند. محدودیت یکپارچگی را به عنوان توانمندسازی یکپارچه‌سازی داده‌ها نیز ذکر می‌کنند. (Cal and et al. 2004).

در حوزه‌های آماری، داده‌های به دست آمده از پرسشنامه‌های سرشماری دارای ساختار پرسشنامه‌ای هستند. قوانین معنایی ویرایش، می‌توانند در طرح پرسشنامه‌ای برای مشخص کردن صحت پاسخ تعریف شوند. چنین قوانینی معمولاً بر پیشگیری از خطا دلالت دارند. برای مثال، یک ویرایش می‌تواند به این شکل باشد: اگر وضعیت تأهل "ازدواج کرده" باشد، سن نباید کمتر از ۱۴ باشد.

پس از تشخیص گزارش‌های خطا، عمل تصحیح مقادیر شروع می‌شود (Holt, Fellegi 1976).

ابعاد مربوط به زمان: این ابعاد عبارت‌اند از: نرخ ارز، نوسانات، و بهنگام بودن. یک جنبه مهم از داده‌ها، به‌روزرسانی آن‌ها در طول زمان است. ابعاد مربوط به زمان اصلی پیشنهادی در متون، نرخ ارز، نوسانات، و بهنگام بودن می‌باشد. جدول ۲ تعاریف ارائه شده در متون را برای این سه بعد زمان مقایسه می‌کند. Wand و Wang (۱۹۹۶) و Redman (۱۹۹۶) تعاریف بسیار مشابهی را برای به‌نگام بودن و نرخ ارز ارائه می‌دهد. Wang و Strong (۱۹۹۶) و Liu و Chi (۲۰۰۲) همان معنی را برای به‌نگام بودن در نظر گرفتند،

درحالی که Bovee و همکاران (۲۰۰۱) تعریفی را برای بهنگام بودن از نظر ارزش و نوسانات ارائه می دهند. تعریف ارزش توسط Bovee و همکاران (۲۰۰۱) بیان شده اما متناظر با تعریف بهنگام بودن توسط Wang و Strong (۱۹۹۶) و Liu and Chi (۲۰۰۲) می باشد. این مقایسه نشان می دهد که در تعریف انتزاعی ابعاد مربوط به زمان توافق وجود ندارد و معمولاً، نرخ ارزش و بهنگام بودن برای اشاره به یک مفهوم استفاده می شوند.

۲-۲-۴ هزینه ها

هزینه ها مرتبط با مصرف منابعی هستند که در هر روش با توجه به اثرات داده های باکیفیت پایین بر روی فعالیت ها در نظر گرفته می شوند. پس، هزینه کیفیت داده برابر با مجموع هزینه های فعالیت های ارزیابی و بهبود کیفیت داده ها می باشد که به هزینه ی برنامه های کیفیت داده و هزینه ی کیفیت داده های ضعیف مربوط می شود. با تخصیص هزینه به داده هایی باکیفیت پایین، می توان پس از اجرای یک برنامه آن ها را به داده هایی باکیفیت مؤثرتر ارتقا داد که معمولاً بین روش ها گران تر هستند. بنابراین، با افزایش هزینه ی برنامه های کیفیت داده، هزینه های تخصیص یافته برای بهبود کیفیت داده های ضعیف کاهش می یابد. این کاهش می تواند به نفع برنامه ی کیفیت داده ها باشد.

هزینه تخصیص یافته برای برنامه ی بهبود کیفیت داده ها را می توان به عنوان یک هزینه پیشگیرانه در نظر گرفت که توسط سازمان ها برای کاهش خطای داده ها پرداخت می شود. این دسته هزینه شامل هزینه تمام گام ها و مراحل است که فرآیند ارزیابی و بهبود کیفیت داده ها را تشکیل می دهد (بخش ۱,۲ را ببینید).

هزینه های بهبود کیفیت پایین داده ها را می توان به شرح زیر طبقه بندی کرد (English 1999):

۱. هزینه های فرآیند، مانند هزینه های مربوط به اجرای دوباره کل فرآیند به دلیل اشتباه در داده ها؛
۲. هزینه های فرصت با توجه به زیان و درآمد ازدست رفته

هزینه‌ی اصلاح کیفیت داده‌های ضعیف برعکس هزینه‌های برنامه‌ی کیفیت داده‌ها به شدت وابسته به زمینه می‌باشد. این باعث می‌شود ارزیابی آن دشوار باشد چراکه ارزش داده‌های مشابه و سطح کیفیت مربوطه اثرات مختلفی بسته به گیرنده‌ی آن‌ها دارد. مثلاً با دریافت اطلاعات منسوخ در سهام توسط یک معامله‌گر فعال ممکن است خسارات اقتصادی قابل توجهی با توجه به تصمیمات سرمایه‌گذاری ایجاد شود. در مقابل، یک روزنامه همان اطلاعات منسوخ را دریافت کرده و گزارش‌های تجاری ماهانه‌ای منتشر می‌کند و ممکن است از این طریق کسی زیان اقتصادی نبیند.

۲-۲-۵ انواع داده‌ها

هدف نهایی یک روش کیفیت داده، تجزیه و تحلیل داده‌ها می‌باشد که در حالت کلی، اشیاء دنیای واقعی را در فرمتی توصیف می‌کند که بتواند با استفاده از روش‌های نرم‌افزاری و طراحی یک شبکه، ذخیره، بازیابی و پردازش شود. در زمینه‌ی کیفیت داده بسیاری از نویسندگان به طور ضمنی یا صریح سه نوع داده را شناسایی کرده‌اند:

۱. داده‌های ساختاریافته، واحدها یا تعمیمی از وضعیتی هستند که توسط ویژگی‌های ابتدایی تعریف شده در یک دامنه، شرح داده شده‌اند. دامنه نشان‌دهنده طیف وسیعی از مقادیری است که می‌توان به ویژگی‌ها اختصاص داد و معمولاً با انواع اطلاعات ابتدایی زبان‌های برنامه‌نویسی، مانند مقادیر عددی یا رشته متن مطابقت دارند. جداول رابطه‌ای و داده‌های آماری رایج‌ترین نوع داده‌های ساختاریافته هستند.

۲. داده‌های ساختار نیافته، دنباله عمومی از علائم می‌باشند، که معمولاً در زبان طبیعی کدگذاری شده‌اند. نمونه‌های معمولی از داده‌های بدون ساختار، پرسشنامه‌های حاوی سؤالات باز با پاسخ‌های آزاد و یا متن پیام یک ایمیل می‌باشد.

۳. داده نیمه ساختاریافته، داده‌هایی هستند که ساختار آن‌ها دارای درجه انعطاف‌پذیری است. داده‌های نیمه ساختاریافته، داده‌های بدون طرح و یا خود توصیفی نیز نامیده می‌شوند (Abiteboul et al. 2000؛ Buneman 1997؛ Calvanese 1999).

Patrick	Metzisi	Masai Mara	KE	(۱) داده‌های ساختار یافته
Mr Patrick Metzisi, Born in the Masai Mara region that is part of Kenia				(۲) داده‌های ساختار نیافته
<pre><PersonalData> <name> Patrick Metzisi </name> <BornIn> Masai Mara, Kenya </BornIn> </PersonalData></pre>				(۳) داده‌های نیمه ساختار یافته

شکل ۲-۲. نمایش‌های مختلفی از اطلاعاتی یکسان در دنیای واقعی (BATINI ET AL. 2009)

XML زبان نشانه‌گذاری مشترک استفاده شده برای نشان دادن داده‌های نیمه ساختاریافته است. برخی از ویژگی‌های مشترک آن‌ها عبارت‌اند از:

(۱) داده‌ها می‌توانند شامل زمینه‌های ناشناخته‌ای در زمان طراحی باشد؛ مثلاً یک فایل XML

دارای فایل طرح XML مربوطه نیست.

(۲) همان نوع داده‌ها ممکن است در راه‌های متعددی نشان داده شوند؛ مثلاً یک داده ممکن است

توسط یک زمینه یا با زمینه‌های متعدد، حتی در درون یک مجموعه داده واحد نشان داده شوند؛

(۳) در میان زمینه‌های شناخته شده در زمان طراحی، بسیاری از زمینه‌ها بدون ارزش می‌باشند.

تکنیک‌های کیفیت داده در مقابل داده‌های ساختار نیافته همواره پیچیده‌تر هستند. مثلاً فرض کنید یک سیستم ثبت‌کننده‌ی اطلاعات شخصی مانند نام، نام خانوادگی، منطقه، و مکان تولید را در نظر بگیریم. شکل ۲ مثال آقای پاتریک متسیزی که در منطقه ماسایی مارا در کنیا به دنیا آمده است، را با استفاده از انواع حالات ساختاریافته (شکل ۲-۲ (۱))، ساختار نیافته (شکل ۲-۲ (۲)) و نیمه ساختاریافته (شکل ۲-۲ (۳)) را نشان می‌دهد.

دقت نحوی آن گونه که در بخش ۳،۲ در مورد داده‌های ساختاریافته توضیح داده شد، اندازه‌گیری می‌شود. برای محاسبه‌ی تابع فاصله‌ی داده‌های نیمه ساختاریافته باید فاصله کلی مربوط به شکل درخت XML را علاوه بر فاصله محلی زمینه‌ها در نظر بگیریم.

اکثر کارهای پژوهشی در متون مربوط به کیفیت داده‌ها بر روی داده‌های ساختاریافته و نیمه ساختاریافته متمرکز شده‌اند. به همین دلیل، اگرچه ما بر ارتباط داده‌های ساختار نیافته اشاراتی می‌کنیم، اما این پژوهش بر داده‌های ساختاریافته و نیمه ساختاریافته متمرکز شده است.

طبقه‌بندی مبتنی بر مشاهده داده‌ها در متون کیفیت داده‌ها، همچون تولید محصولات می‌باشد (Shankaranarayan et al. 2000). از این منظر، سه نوع داده مشخص شده‌اند:

- داده‌های خام: این داده‌ها به عنوان داده‌هایی تعریف می‌شوند که دست‌خوش هیچ پردازشی نشده‌اند و با ایجاد و ذخیره‌سازی اولیه، می‌توان آن‌ها را برای مدت طولانی نگهداری کرد.
- محصولات داده‌ای: که نتیجه‌ی یک فعالیت تولیدی انجام شده بر روی داده‌ها می‌باشند؛
- اقلام داده‌ای تکمیل کننده: که در زمان ایجاد اطلاعات محصول مربوطه مورد نیاز می‌باشند و موقتاً ذخیره می‌شوند تا محصول نهایی را بسازند.

همان‌طور که در بخش قبل بحث شد، این نوع طبقه‌بندی موجب می‌شود تا نرم‌افزار، داده تکنیک‌های کیفی را به‌طور سنتی در جهت تضمین کیفیت فرآیندهای تولید استفاده کند.

۲-۲-۶ ارتباط کیفیت داده با سیستم‌های اطلاعاتی

روش‌های کیفیت داده توسط نوع سیستم‌های اطلاعاتی که به فعالیت‌های ارزیابی و بهبود مربوط می‌شوند نیز تحت تأثیر قرار می‌گیرند. این متون مفهوم معماری سیستم‌های اطلاعاتی (معماری IS^۱) را برای توصیف مدل

هماهنگی که توسط سیستم اطلاعاتی شرکت پشتیبانی شده است، فراهم می‌کند (Zachman 2006). معماری انواع سیستم‌های اطلاعاتی بر اساس درجه داده، پردازش و ادغام مدیریت پشتیبانی شده توسط یک سیستم فنی، متمایز شده‌اند. همان‌طور که درجه ادغام، پردازش و مدیریت داده کاهش می‌یابد، ارزیابی کیفیت داده و تکنیک‌های بهبود که اعمال می‌شوند، پیچیده‌تر می‌شوند. در همان زمان، ارزیابی و بهبود کیفیت داده‌ها چالش برانگیزتر می‌شود. انواع سیستم‌های اطلاعاتی زیر را می‌توان بر اساس درجه ادغام آن‌ها متمایز کرد:

- در یک سیستم اطلاعات یکپارچه در حقیقت نرم‌افزارها تک لایه‌ای هستند که دسترسی به داده‌ها را ارائه نمی‌دهند. اگرچه معمولاً داده‌ها در پایگاه داده‌ای مرتب می‌شوند که قابل جستجو باشند، اما نرم‌افزارهای جداگانه را به اشتراک نمی‌گذارند. این امر می‌تواند موجب تکرار داده‌ها شود، که احتمالاً بر تمام ابعاد کیفیت نیز مؤثر می‌باشد.
- یک انبار داده مجموعه‌ای متمرکز از اطلاعات بازایی شده از پایگاه داده‌های متعدد می‌باشد. انبار داده به صورت دوره‌ای با داده‌های به روز از پایگاه داده‌ی اصلی به وسیله استخراج روندهای خودکار و هماهنگی داده‌ها، بهنگام سازی می‌شود. داده‌ها از لحاظ فیزیکی نیز یکپارچه می‌شوند، چون آن‌ها با توجه به طرح، ادغام و درنهایت ذخیره در انبار داده‌ها مجدداً فرمت دهی می‌شوند.
- سیستم اطلاعات توزیعی، مجموعه‌ای از ماژول‌های نرم‌افزار هماهنگ شده توسط یک گردش کار می‌باشد. نرم‌افزارها معمولاً به ردیف‌هایی مانند نمایش‌ها، منطق برنامه، مدیریت داده‌ها و ویژگی‌های دسترسی به داده‌های خروجی تقسیم می‌شود. در این حالت داده‌ها را می‌توان در پایگاه داده‌های مختلف ذخیره کرد اما قابلیت همکاری توسط ادغام منطقی طرح آن‌ها تضمین شده می‌ماند.
- سیستم اطلاعات تعاونی (CIS^۱) را می‌توان به عنوان یک سیستم اطلاعاتی در مقیاس بزرگ نیز تعریف کرد که سیستم‌های مختلف در سازمان‌های مختلف و مستقل را برای به اشتراک گذاری اهداف مشترک به صورت داخلی به هم وصل می‌کند (De Michelis et al. 1997). هماهنگی

با دیگر سیستم‌های اطلاعاتی نیاز به توانایی تبادل اطلاعات دارد. در سیستم‌های اطلاعات تعاونی، داده‌های منطقی یکپارچه نیست، چراکه در پایگاه داده‌های جداگانه با توجه به طرح‌های مختلف ذخیره می‌شوند. با این حال، نرم‌افزارها روند انتقال و تبادل داده‌ها را ترکیب می‌کنند تا بتوانند موجب همکاری میان فرآیندهای رایج شوند. به عبارت دیگر، ادغام در سطح فرآیند تحقق می‌یابد.

- در متون علمی برای نشان دادن هر نوع اطلاعات اتخاذ شده از فن‌آوری‌های وب، از اصطلاح سیستم اطلاعاتی وب WIS (Isakowitz et al. 1998) استفاده می‌شود. از دیدگاه فنی سیستم اطلاعاتی وب یک نرم‌افزار سرویس گیرنده / سرویس دهنده است. چنین سیستمی معمولاً از داده‌های ساختاریافته، نیمه ساختاریافته و ساختار نیافته استفاده کرده و توسط ابزارهای توسعه و مدیریت بر اساس تکنیک‌های خاص برای هر نوع داده پشتیبانی می‌شود.

- در یک سیستم اطلاعاتی همتا به همتا (P2P)، تمایزی بین کاربر و سرور وجود ندارد. این سیستم توسط مجموعه‌ای از گره‌های یکسان تشکیل شده است که داده‌ها و خدمات نرم‌افزاری را به منظور برآوردن نیازهای کاربران به اشتراک می‌گذارد. سیستم‌های P2P توسط چند خاصیت مشخص می‌شود:

- نداشتن هماهنگی مرکزی
- نداشتن پایگاه داده مرکزی
- همتا نداشتن در دیدگاه جامع سیستم

همتایان آن‌ها خودمختارند و می‌توانند به صورت پویا به سیستم وصل شده یا از سیستم قطع شوند. با این حال، همتایان معمولاً روش‌های مدیریت مشترک را به اشتراک می‌گذارند.

۳-۲ مقایسه‌ی متدلوژی‌های مختلف کیفیت داده‌ها

در این بخش، روش‌ها بر اساس معیارهای بحث شده در بخش قبلی مقایسه می‌شوند. جدول ۲-۴ لیست روش‌های مدنظر در این پژوهش را توسط کلمات اختصاری همراه با نام توسعه‌یافته روش و مرجع اصلی نشان می‌دهد. همچنین از مخفف‌هایی برای اشاره به هر روش در بقیه این پژوهش استفاده خواهد شد.

جدول ۲-۴. مدل‌های مختلف کیفیت داده‌ها

مخفف روش	نام اصلی	مراجع
TDQM	مدیریت کیفیت داده (Data Quality Management)	Wang 1998
DWQ	روش‌شناسی کیفیت انبار داده (The Datawarehouse Quality Methodology)	Jeusfeld et al. 1998
TIQM	مدیریت کیفیت اطلاعات کلی	English 1999
AIMQ	روشی برای ارزیابی کیفیت اطلاعات	Lee et al. 2002
CIHI	روش موسسه کانادایی برای روش اطلاعات سلامت	Long and Seko 2005
DQA	ارزیابی کیفیت داده	Pipino et al. 2002
IQM	اندازه‌گیری کیفیت اطلاعات	Eppler & Munzenmaier 2002
ISTAT	روش ISTAT	Falorsi et al 2003
AMEQ	روش اندازه‌گیری مبتنی بر فعالیت و ارزیابی کیفیت اطلاعات محصول	Su and Jin 2004
DaQuinCIS Data	کیفیت در سیستم‌های اطلاعات تعاونی	Scannapieco et al. 2004
QAFD	روشی برای ارزیابی کیفیت داده‌های مالی	De Amicis and Batini 2004
CDQ	روشی جامع برای مدیریت کیفیت داده	Batini and Scannapieco 2006

جدول ۲-۵. مقایسه مدل‌های مختلف کیفیت داده‌ها مبتنی بر روش‌های ارزیابی

مخفف روش / مرحله	تحلیل داده‌ها	تحلیل الزامات DQ	شناسایی حوزه بحرانی	مدل‌سازی فرآیند	اندازه‌گیری کیفیت	قابل توسعه به دیگر ابعاد و معیارها
TDQM	+		+	+	+	ثابت
DWQ	+	+	+		+	باز
TIQM	+	+	+	+	+	ثابت
AIMQ	+		+		+	ثابت
CIHI	+		+			ثابت
DQA	+		+		+	باز
IQM	+				+	باز
ISTAT	+				+	ثابت
AMEQ	+		+	+	+	باز
COLDQ	+	+	+	+	+	ثابت
DaQuinCIS Data	+		+	+	+	باز
QAFD	+	+	+		+	ثابت
CDQ	+	+	+	+	+	باز

هزینه‌ها، ابعاد و گام‌ها به‌عنوان متمایزترین معیار که منجر به شناسایی چهار روش متفاوت از هم می‌شود، نشان داده می‌شوند، که در ادامه در مورد آن‌ها بحث شده و تفاوت‌های اساسی میان آن‌ها را برجسته می‌شوند. در بقیه این بخش، روش‌ها در ادامه‌ی دیدگاه‌های مشخص شده در بخش ۲ مقایسه شده‌اند.

۲-۳-۱ روش‌ها، گام‌ها و مراحل

جداول ۲-۵ و ۲-۶ مراحل و گام‌های مورد توجه قرار گرفته توسط هر یک از روش‌ها بمنظور ایجاد بهبود در کیفیت را نشان می‌دهد. در این جداول اگر در مدل مربوطه به آن گام یا روش ایجاد بهبود اشاره شده باشد از علامت "+" استفاده شده است و در غیر اینصورت آن سلول خالی خواهد بود. به‌عنوان‌مثال، روش‌شناسی کیفیت انبار داده عموماً به مرحله مدل‌سازی فرآیند اشاره دارد، اما جزئیات اجرا را ارائه نمی‌دهد، لذا در جدول ۲-۵ این گام برای مدل روش‌شناسی کیفیت انبار داده در نظر گرفته نشده است. با دلیل مشابهی، مرحله اندازه‌گیری کیفیت موسسه کانادایی برای روش اطلاعات سلامت در جدول ۲-۵ در نظر گرفته نشده است. ارزیابی و بهبود به‌طور جداگانه در دو بخش بعدی بحث شده‌اند.

جدول ۲-۶. مقایسه مدل‌های مختلف کیفیت داده‌ها مبتنی بر گام‌های بهبود

شناسایی علت خطاها	استراتژی و تکنیک‌های انتخاب	ارزیابی پاسخگویی داده	ارزیابی قابلیت پاسخگویی روش	ارزیابی هزینه	مخفف روش / مرحله
+	+	+	+	+	TDQM
+	+	+		+	DWQ
+	+	+	+	+	TIQM
+					DQA
+	+				ISTAT
+					AMEQ
+	+			+	COLDQ
+	+				DaQuinCIS Data
+	+	+	+	+	CDQ

جدول ۲-۷. مقایسه مدل های مختلف کیفیت داده ها مبتنی بر گام های بهبود

مخفف روش / مرحله	کنترل فرآیند	طراحی راه حل های بهبود داده	طراحی مجدد فرآیند	مدیریت بهبود	نظارت بر بهبود
TDQM			+	+	+
DWQ		+		+	
TIQM		+	+		+
DQA					
ISTAT		+	+		
AMEQ					+
COLDQ	+	+	+		+
DaQuinCIS Data					
CDQ	+	+	+		

۲-۳-۲ فاز ارزیابی

جدول ۲-۷ مراحل زیر را با روش های مختلف در مرحله ارزیابی مقایسه می کند. به طور کلی، روش ها با مراحل طبقه بندی شده در بخش ۱،۲، با نام های مختلف نشان داده شده اند. با این حال، تشخیص ارتباط نام مراحل با تجزیه و تحلیل اهداف مراحل مشکل نیست. به عنوان مثال، روش موسسه کانادایی برای روش اطلاعات سلامت در مورد یک گام روش شناسی بحث می کند که به شناسایی پایگاه های داده با سطح کیفیت پایین تر از یک آستانه قابل قبول می پردازد. این مرحله دارای گام های روشن همراه با یافتن نقاط بحرانی بخش ۱،۲ می باشد.

اکثر مراحل برای ارزیابی گام‌ها، تجزیه و تحلیل داده‌ها و اندازه‌گیری کیفیت طراحی شده‌اند. باین حال، آن‌ها با توجه به روش‌های مختلف انجام می‌شوند. مثلاً مرحله اندازه‌گیری کیفیت به صورت پرسشنامه در روشی برای ارزیابی کیفیت اطلاعات با ترکیبی از معیارهای ذهنی و عینی در ارزیابی کیفیت داده و یا با تجزیه و تحلیل آماری در روشی برای ارزیابی کیفیت داده‌های مالی انجام می‌شود. روش‌های اندازه‌گیری، شرایط خاص محتوای سازمانی، فرآیندها، کاربران و یا خدمات مختلف را برآورده می‌کنند.

جدول ۲-۸. روش‌ها و انواع استراتژی‌ها

مخفف روش/استراتژی	داده محور	فرآیند محور
TDQM		طراحی مجدد فرآیند
DWQ	ادغام داده‌ها و طرح	
TIQM	پاک کردن داده نرمالیزه کردن محلی سازی و تصحیح خطا	طراحی مجدد فرآیند
ISTAT	نرمالیزه کردن ثبت ارتباطات	طراحی مجدد فرآیند
COLDQ	بهینه سازی هزینه	کنترل فرآیند ثبت ارتباطات
DaQuinCIS	اعتماد منبع ثبت ارتباطات	
CDQ	نرمالیزه کردن ثبت ارتباطات	

تنها چند روش هستند که مرحله‌ی تجزیه و تحلیل الزامات کیفیت داده، شناسایی مسائل کیفیت داده و جمع‌آوری سطح کیفیت هدف جدید از کاربران را انجام می‌دهند. این مرحله به طور ویژه به ارزیابی و حل اختلافات در سطوح هدف کیفیت داده ذینفعان مختلف مرتبط می‌باشد. مثلاً روشی برای ارزیابی کیفیت

داده‌های مالی مجموعه‌ای از سطوح کیفیت هدف را با انواع مختلف برای کارشناسان توصیه می‌کند، از جمله کارشناسان کسب و کار و اپراتورهای مالی، اما به اصلاح سطوح کیفیت داده ناسازگار کمکی نمی‌کند. چند روش از مدل‌سازی فرآیند پشتیبانی می‌کنند. توجه داشته باشید که به‌غیر از روش اندازه‌گیری مبتنی بر فعالیت و ارزیابی کیفیت اطلاعات محصول، روش‌ها از مدل‌سازی فرآیند پشتیبانی کرده و همچنین یک استراتژی فرآیند محور را برای مرحله بهبود اتخاذ می‌کنند.

۲-۳-۳ گام بهبود

جداول ۲-۶ و ۲-۷ مراحل بهبود روش‌های مختلف را مقایسه می‌کند. معمولاً در گام بهبود به شناسایی علل خطاها پرداخته می‌شود. روش ارزیابی کیفیت داده بر اهمیت گام شناسایی علل خطاها تأکید می‌کند، اما در مورد اجرای آن‌ها بحث نمی‌کند. همچنین روش شناسی کیفیت انبار داده به مدل ریاضی بر اساس مفهوم وابستگی به‌منظور حمایت از گام شناسایی علل خطاها اشاره می‌کند، اما تعریف مدل به‌عنوان کار در حال انجام و بدون فراهم کردن آن انجام شده است.

فعالیت‌های بهبود در روش‌های مختلف به‌غیر از روش روش شناسی کیفیت انبار داده، عمدتاً بر اساس طراحی مجدد فرآیند هستند، که در آن فرمت معیار مدنظر، سؤال متریک هدف^۱ را فراهم می‌کنند (Basili et al. 1994). این رویکرد ابتدا در زمینه‌ی مهندسی نرم‌افزار ارائه شده بود. مرحله ارزیابی هزینه، معمولاً در روش کیفیت داده الزامی است. این مرحله برای اندازه‌گیری مزیت اقتصادی راه‌حل‌های بهبود و انتخاب کارآمدترین تکنیک‌های بهبود در نظر گرفته می‌شود. در مقابل، مدیریت مرحله‌ی راه‌حل بهبود تنها با مدیریت کیفیت داده انجام می‌شود. روش‌های دیگر به طیف گسترده‌ای از تکنیک‌ها و بهترین شیوه‌های مدیریت تغییر اشاره دارند (Kettinger, Grover 1995). علاوه بر این، ممکن است مرحله ارزیابی روش

^۱ Goal Question Metric

به منظور ارزیابی نتایج حاصل از گام بهبود تکرار شود. مثلاً روش ارزیابی کیفیت داده گام‌های روش قبلی را به منظور بررسی اثربخشی بهبود توصیه می‌کند.

در نهایت، رابطه بین کیفیت داده، فرآیند و سازمان توسط روش‌های مدیریت کیفیت اطلاعات کلی، مدیریت کیفیت داده و روشی جامع برای مدیریت کیفیت داده در نظر گرفته می‌شود. این روش‌ها به بحث تخصیص مسئولیت در فرآیندها و داده‌ها می‌پردازند. این مراحل توسط نتایج حاصل از گام بازسازی حالت پشتیبانی می‌شوند. روش روشی جامع برای مدیریت کیفیت داده در مورد مجموعه‌ای از ماتریس‌ها برای نشان دادن رابطه بین فرآیندها، واحدهای سازمانی و پایگاه داده بحث می‌کند که در طول مرحله بازسازی حالت ایجاد شده و پس از آن در مراحل انتساب مسئولیت استفاده می‌شود.

۲-۴ روش‌ها، استراتژی‌ها و تکنیک‌ها

جدول ۲-۸، استراتژی‌ها و تکنیک‌های به خدمت گرفته شده توسط روش‌های متفاوت را نشان می‌دهد. یک روش با یک استراتژی در ارتباط است اگر راهنمایی را برای انتخاب و طراحی تکنیک‌های متناظر، فراهم کند.

توجه داشته باشید که ستون‌های جدول ۲-۸ با برجسب فرآیند محور^۱، همان اطلاعات ستون‌های کنترل فرآیند و طراحی مجدد فرآیند در جدول ۲-۷ را نشان می‌دهند و ستون‌هایی با برجسب داده محور صراحتاً به تکنیک‌های داده محور ضمنی در جداول ۲-۶ و ۲-۷ اشاره می‌کنند.

جدول ۲-۸، روشهای مختلف کیفیت داده برای استراتژی‌های ترکیبی را به نمایش می‌دهد و ترکیب‌های متنوعی از روش‌های داده محور و تکنیک‌های فرآیند محور را ارائه می‌دهد. روش مدیریت کیفیت اطلاعات کلی برای طیف وسیع‌تری از تکنیک‌های داده محور و فرآیند محور استفاده می‌شود. در مقابل آن، روش مدیریت کیفیت داده دستورالعمل‌هایی برای اعمال استراتژی‌های فرآیند محور با استفاده از اطلاعات

^۱ Process-driven

"ماتریس تحلیل تولید" (Ballou et al. 1998) فراهم می‌کند، که نشان می‌دهد چه زمانی و چگونه باید داده‌ها را بهبود داد.

شایان‌ذکر است روشی که به‌طور انحصاری داده محور (مثلاً روش‌شناسی کیفیت انبار داده برای ارزیابی کیفیت در سیستم‌های اطلاعات تعاونی) و یا استراتژی فرآیند محور را اتخاذ کند ممکن است برای سازمان‌هایی که دارای شیوه‌های کیفیت داده هستند انعطاف‌پذیر نباشد. تنها روشی که صراحتاً به این مسئله می‌پردازد روشی جامع برای مدیریت کیفیت داده می‌باشد، که به‌طور مشترک تکنیک داده محور و فرآیند محور را انتخاب می‌کند. انتخاب مناسب‌ترین استراتژی و تکنیک مبتنی بر متغیرهای تصمیم وابسته به دامنه می‌باشد (Batini et al. 2008).

نرمالیزه کردن، ثبت ارتباطات داده‌ها و یکپارچه‌سازی طرح نشان‌دهنده‌ی تکنیک‌های داده محوری است که به‌طور گسترده در روش کیفیت داده اتخاذ شده‌اند، درحالی‌که طراحی مجدد فرآیند به روش فرآیند محور مربوط می‌شود. اکنون به نوآوری‌های خاص مربوط به تکنیک‌های داده محور و فرآیند محور می‌پردازیم.

۲-۴-۱ تکنیک‌های داده محور

تکنیک‌های نرمالیزه کردن در حوزه‌های مختلف، از جمله سرشماری و داده‌های منطقه‌ای مطرح شده است. روش ISTAT و روش جامع برای مدیریت کیفیت داده تکنیک‌های نرمالیزه سازی را برای بهبود کیفیت داده از طریق مقایسه داده‌ها با نگاه کردن^۲ به جداول و تعریف یک طرح جامع ارائه می‌دهند. مثلاً روش ISTAT از سیستم ثبت ملی خیابان به‌عنوان یک جدول برای داده‌های منطقه‌ای استفاده می‌کند.

ثبت ارتباطات در تحقیقات پایگاه داده از دهه‌ی ۵۰ بررسی شد و در بسیاری از زمینه‌ها مانند برنامه‌های مراقبت بهداشتی، اداری و سرشماری استفاده شده است. در چنین شرایطی، ایجاد روش‌های کارآمد منطبق با

^۱ Manufacturing Analysis Matrix

^۲ Look-Up

کامپیوتر می تواند استفاده از منابع دفتری را کاهش دهد و خطاها را به حداقل برساند. روش جامع برای مدیریت کیفیت داده در مورد سه نوع از تکنیک های ثبت ارتباطات بحث می کند:

۱. تکنیک های احتمالی، که بر اساس مجموعه ی گسترده ای از روش ها که در طول دو قرن گذشته، در آمار و نظریه احتمالات توسعه یافته اند، مثل شبکه های بیزی در داده کاوی، استفاده می کند.
 ۲. تکنیک های تجربی که از تکنیک های الگوریتمی مانند طبقه بندی، تجزیه و تحلیل درخت، مقایسه همسایه و هرس استفاده می کنند.
 ۳. روش های مبتنی بر دانش، که بر اساس استخراج دانش از زمینه ها و استفاده از استراتژی های استدلال بر مبنای آن ها بکار می روند.
- معیارهای انتخاب بین این سه نوع تکنیک در روش جامع برای مدیریت کیفیت داده مورد بحث قرار گرفت. پروژه ارزیابی کیفیت در سیستم های اطلاعات تعاونی، تکنیک ثبت ارتباطات خاص را توسعه داده است (Bertolazzi et al. 2003). در پروژه ارزیابی کیفیت در سیستم های اطلاعات تعاونی ارتباط رکوردها در دو گام انجام شد:

۱. اول، ثبت ارتباطات با نسخه های مختلف از نهادهای مشابه در منابع داده مختلف هم راستا می شود؛
 ۲. دوم، ثبت ارتباطات با استفاده از پرس و جو و شناسایی موارد مشابه در نتایج پرس و جو که توسط هر منبع داده بازگردانده شده، صورت می گیرد.
- روش ثبت ارتباطات مبتنی بر روش همسایگی طبقه بندی شده است (Hernandez, Stolfo 1998)، اما برخی از ویژگی های جدید زیر برای آن معرفی شده اند:

- تطبیق کلیدی به جای انتخاب طراح بوده و به طور خود کار انتخاب می شود.
- الگوریتم تطبیق مبتنی بر تابعی است که تابع فاصله کلاسیک را بر حسب طول رشته ها، نرمالیزه می کند.

یکپارچه سازی داده ها و طرح ها (Lenzerini 2002) حوزه گسترده ای از تحقیقات می باشد که تا حدودی با کیفیت داده ها دارای همپوشانی هستند. اغلب تکنیک های بهبود داده محور مورد استفاده در روش ها مبتنی بر استفاده از داده های جدید هستند که به منظور بهبود کیفیت جمع آوری داده ها ارائه شده اند. در نتیجه تکنیک های بهبود کیفیت داده در درجه اول بر روی ناهمگنی سطح نمونه به منظور شناسایی ثبت های مشابه، ارزش های متضاد و انتخاب یک نمونه ی نهایی منحصر به فرد متمرکز هستند.

روش های روش شناسی کیفیت انبار داده، ISTAT، ارزیابی کیفیت در سیستم های اطلاعات تعاونی و روش جامع برای مدیریت کیفیت داده به ناهمگنی سطح نمونه می پردازند. در روش روش شناسی کیفیت انبار داده، منابع اطلاعاتی ناهمگن در دسترس، ابتدا به شیوه ای یکنواخت از طریق مکانیسم های استخراج به نام الگوریتم های پوششی ساخته می شوند. سپس در کار یکپارچه سازی اطلاعات و حل منازعات میانجی گری می کنند. در نتیجه داده های استاندارد و یکپارچه شده به صورت موردی در انبار داده ها ذخیره می شوند.

روش ISTAT نشان می دهد که چگونه ناهمگنی در میان داده های مدیریت شده با مدل مشترک برای نشان دادن فرمت داده های تبدیل شده بر اساس زبان نشانه گذاری XML، توسط سازمان های دولتی مختلف حل می شود. در این روش درک ناهمگنی میان سازمان ها آسان تر شده است، در حالی که برای حل این ناهمگنی به توافق دو یا چند جانبه نیاز است.

در روش ارزیابی کیفیت در سیستم های اطلاعات تعاونی به بررسی ناهمگنی سطح نمونه در میان منابع داده مختلف توسط کار گزار کیفیت داده پرداخته می شود. نسخه های مختلفی از داده های مشابه دریافت شده به عنوان پاسخ به درخواست، توسط کار گزار کیفیت داده تنظیم می شوند و بهترین مقدار کیفیتی انتخاب می شود.

روش جامع برای مدیریت کیفیت داده مشابه روش ISTAT می باشد چرا که تأکید بیشتری بر استقلال سازمان در یک سیستم همکاری دارد. در این روش حل ناهمگنی مطالعات موردی که به عنوان بهترین شیوه

پیشنهاد شده است، از طریق ثبت ارتباطات در یک لایه بسیار نازک از داده‌ها، یعنی شناسه انجام شده است. سایر داده‌ها تنها در صورت تصمیم‌گیری مستقل سازمان‌های درگیر استفاده می‌شوند.

۲-۴-۲ تکنیک‌های فرآیند محور

روش‌هایی که به گام طراحی مجدد فرآیند می‌پردازند تمایل به استفاده از تکنیک‌های مربوط به فرآیند مهندسی مجدد کسب و کار دارند (BPR) (Muthu et al. 1999) (Hammer 1990). روش مدیریت کیفیت داده در این رابطه استثناست به طوری که رویکردی برای کنترل فرآیند طراحی مجدد را پیشنهاد می‌دهد و به یک "سیستم تولید محصول اطلاعاتی" اشاره می‌کند (Ballou et al. 1998). این روش مدل نقشه تولید اطلاعات (IP-MAP) را پیشنهاد می‌دهد (Shankaranarayan et al. 2000) که در مدل-های محصولات اطلاعاتی مدیریت شده توسط فرآیندهای تولید استفاده می‌شود. "نقشه تولید اطلاعات" یک مدل گرافیکی طراحی شده برای کمک به تجسم تحلیلگران فرآیند تولید اطلاعات، شناسایی مالکیت مراحل فرآیند، درک اطلاعات و مرزهای سازمانی، برآورد زمان و کیفیت معیارهای مرتبط با فرآیند تولید فعلی است. توصیف فرآیندها، یک فعالیت اجباری و هم‌راستا با جهت‌گیری کلی استراتژی‌های فرآیند محور است. پس از مدل‌سازی و ارزیابی فرآیند تولید اطلاعات، فعالیت‌های کنترل فرآیند جدید مشخص می‌شوند و یا در مورد طراحی مجدد فرآیند تصمیم گرفته می‌شود.

راه‌حل‌های پیچیده مانند IP-MAP به دلیل هزینه‌های بالای پیاده‌سازی همیشه قابل استفاده نیستند و در برخی موارد، کاربرد عملی از طریق اجرای تنه‌ای مدل‌سازی فرآیند امکان‌پذیر نیست. به همین دلیل، روش‌های دیگر و راه‌حل‌های کمتر رسمی ولی امکان‌پذیر انتخاب می‌شود. مثلاً روش جامع برای مدیریت کیفیت داده مبتنی بر مجموعه‌ای از ماتریس‌هاست که روابط اصلی را در میان داده‌ها، جریان اطلاعات، فرآیندها و واحدهای سازمانی توصیف می‌کنند. رابطه بین واحدهای سازمانی و فرآیندها نیز در توسعه IP-MAP مدل‌سازی شده‌اند. (Scannapieco et al. 2002).

۵-۲ روش‌ها و ابعاد

ابعاد کیفی در نظر گرفته شده برای روش‌های بررسی شده در این پژوهش در جدول ۸ نشان داده شده‌اند. در این جدول هر بعد در ارتباط با یک روش است، اگر آن روش تعریف متناظری در مورد آن ارائه کند و در جدول ۴-۲ نیز متناظر با هر بعد و روش آدرس‌دهی انجام گرفته است (جدول ۴-۲ را ببینید).

توجه کنید که ابعاد گوناگونی در روش‌ها تعریف شده که پیچیدگی مفهوم کیفیت داده را تأیید می‌کند. این امر غافلگیرکننده نیست چرا که امروزه تعداد زیادی از پدیده‌ها می‌توانند برحسب داده‌ها تعریف شوند. دسته‌بندی چندگانه ابعاد کیفی با روش‌ها ارائه شده است. روش مدیریت کیفیت اطلاعات کلی ابعاد را به صورت ذاتی^۱ و عملی^۲ دسته‌بندی می‌کند. روش ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها میان طرح، داده، واسطه‌ها و ابعاد کاربردی اطلاعات تمایز ایجاد می‌کند. روش موسسه کانادایی برای روش اطلاعات سلامت یک دسته‌بندی دوحلای برحسب ابعاد و خصوصیات مرتبط ارائه می‌کند. روش جامع برای مدیریت کیفیت داده نیز ابعاد طرح و داده را ارائه می‌کند.

در جدول ۱۰-۲ معیارهای فراهم شده برای بررسی ابعاد کیفی روش‌های مختلف آورده شده است. معیارها برای دقت معنایی بکار نمی‌روند زیرا دو روش ارزیابی کیفیت داده‌های مالی و روش جامع برای مدیریت کیفیت داده که آن را آدرس‌دهی می‌کنند، روش‌های اندازه‌گیری مشخصی را ارائه نمی‌کنند. با توجه به این ویژگی‌ها، چند معیار یا شاخص اندازه‌گیری برای هر بعد تعریف شده و هر بعد دارای چند ورودی^۳ در جدول است. توجه کنید معیارهای ذهنی^۴ نظیر پیمایش‌های^۵ کاربر تقریباً برای تمام ابعاد کیفی تعریف شده است.

^۱ Inherent

^۲ Pragmatic

^۳ Entry

^۴ Subjective

^۵ Survey

معیارهای متفاوت برای ابعاد یکسان با مخفف‌هایی شناسایی می‌شوند که در جدول ۲-۱۱ آمده‌اند تا آن‌ها را با روش‌هایی مرتبط کنند که در آن‌ها استفاده و یا تعریف شده‌اند.

آخرین ستون جدول ۲-۱۱ برای هر بعد و هر سنجه مرتبط با آن، دو مورد را مشخص می‌کند:

- تعداد روش‌هایی که از سنجه‌ها استفاده می‌کنند

- تعداد کل روش‌هایی که به بعد متناظر با آن سنجه اشاره می‌کنند

نسبت میان این دو عدد (حاصل تقسیم این دو عدد) در سنجه‌های مختلف به عنوان درجه اجماع^۱ آن سنجه در نظر گرفته می‌شود. چنین معیاری برای ابعادی مثل دقت، جامعیت^۲ و ثبات^۳ معمولاً بالا است در حالی که این مقدار به شکل چشمگیری برای دو بعد مرتبط به زمان یعنی بموقع بودن^۴ و نرخ ارز^۵ پایین است.

در متدولوژی اندازه‌گیری کیفیت اطلاعات که به کیفیت اطلاعات وب می‌پردازد تقریباً برای هر بعد یک سنجه وجود دارد. چنین معیارهایی با در نظر گرفتن ابزارهای اندازه‌گیری که در زمینه‌ی وب هستند، تعریف می‌شود. مثلاً امکان دارد با استفاده از یک تحلیلگر سایت، ابعادی نظیر قابلیت دسترسی، پایداری، بجا بودن، اختصار^۶ و قابلیت نگهداری^۷ را ارزیابی کرد. تحلیلگران می‌توانند برای ارزیابی عملی بودن و سهولت اجرا از ارزیابی حجم محاسبات ترافیکی استفاده کنند در حالی که اسکنرهای درگاه^۸ برای ارزیابی امنیت، مفید هستند. تعداد بالای ابزارهای اندازه‌گیری در حوزه وب، منجر به وجود تعداد بالای معیارها در متدولوژی اندازه‌گیری کیفیت اطلاعات می‌شود.

^۱ Consensus

^۲ Completeness

^۳ Consistency

^۴ Timeliness

^۵ Currency

^۶ Conciseness

^۷ Maintainability

^۸ Port

متدولوژی ارزیابی کیفیت اطلاعات دارای معیارها و ابعاد مشخص گوناگونی است. این امر به علت رویکرد بالا به پایین به خدمت گرفته شده در روش ارزیابی کیفیت اطلاعات برای تعریف ابعاد و معیارهاست که از دو دسته بندی متفاوت استفاده می کند:

۱. کیفیت محصول در مقابل کیفیت خدمات

۲. مطابقت با مشخصات در مقابل برآورده سازی نیازهای مشتری یا عدم تأمین آنها

که این ها منجر به مجموعه ی به شدت پراکنده ای از ابعاد و معیارهای مرتبط می شود.

جدول ۲-۹. روش‌ها و ابعاد کیفی

مخفف	ابعاد کیفی داده‌ها
TDQM	قابلیت دسترسی، مناسب بودن، باورپذیری، جامعیت، ارائه همسان، سهولت تغییر، دارای ارزش افزوده، بدون خطا، تفسیرپذیری، عینی بودن، مربوط بودن، شهرت، امنیت، بجا بودن، درک پذیری
DWQ	صحت، تمامیت، به‌اندازه بودن، قابلیت ردگیری، قابلیت تفسیر، تکامل‌پذیری فراداده‌ها، قابلیت دسترسی (سیستم، تراکنش، امنیت)، مفید بودن (قابلیت تفسیر)، بجا بودن (رایج بودن، فراریت)، پاسخ‌دهی، تمامیت، اعتبار، دقت، همسانی، قابلیت تفسیر
TIQM	ابعاد ذاتی: متابعت تعریف (همسانی)، جامعیت، متابعت قوانین تجاری، دقت (در منبع جایگزین)، دقت (در واقعیت)، درستی، بدون تکثیر، هم ارزی داده‌های زائد، هم‌زمانی داده‌های زائد ابعاد عملی: قابلیت دسترسی، بجا بودن، وضوح متنی، یکپارچگی اشتقاق، قابلیت استفاده، راستی (تمامیت حقایق)، هزینه
AIMQ	قابلیت دسترسی، مناسب بودن، باورپذیری، جامعیت، ارائه موجز/همسان، سهولت اجرا، بدون خطا، قابلیت تفسیر، عینی بودن، مربوط بودن، شهرت، امنیت، بجا بودن، درک پذیری
CIHI	ابعاد: دقت، قابلیت مقایسه بجا بودن، قابلیت استفاده، مربوط بودن خصوصیات: پوشش زیاد، پوشش کم، قابلیت اطمینان، مجموعه و ضبط، واحد/آیتم بدون پاسخ، ویرایش‌پذیری و سندیت داشتن، پردازش، تخمین، بجا بودن، جامعیت، یکپارچگی، استانداردسازی، هم ارزی، توانایی پیوند، قابلیت مقایسه محصول/تاریخی، قابلیت دسترسی، مستندسازی، قابلیت تفسیر، انطباق‌پذیری، ارزش
DQA	قابلیت دسترسی، حجم مناسبی از داده، باورپذیری، جامعیت، بدون خطا، همسانی، ارائه موجز/همسان، مربوط بودن، سهولت دست‌کاری، قابلیت تفسیر، عینی بودن، شهرت، امنیت، بجا بودن، درک‌پذیری، ارزش افزوده

مخفف	ابعاد کیفی داده‌ها
IQM	قابلیت دسترسی، همسانی، بجا بودن، موجز بودن، نگهداشت پذیری، رایج بودن، عملی بودن، سهولت، سرعت، جامعیت، وضوح، دقت، قابلیت ردگیری، امنیت، صحت، تعاملی
ISTAT	دقت، تمامیت، همسانی
AMEQ	ارائه همسان، قابلیت تفسیر، وضعیت درک پذیری، ارائه موجز، بجا بودن، تمامیت ارزش افزوده، مناسب بودن، معناداری، فقدان پریشانی، تنظیم، قابل خواندن، معقولیت، درستی، قابلیت اطمینان، بدون تورش، کارایی داده، کاستی طراحی، عملیات، کاستی‌ها، دقت، هزینه، عینی بودن، باورپذیری، شهرت، قابلیت دسترسی، صحت، عدم ابهام همسانی
COLDQ	طرح: روشنی تعریف، جامعیت، انعطاف‌پذیری، نیرومندی، ضرورت، ویژگی دانه‌ای، دقت حوزه‌ها، همگنی، قابلیت شناسایی، قابلیت اکتساب، مربوط بودن، سادگی/پیچیدگی، همسانی معنایی، همسانی نحوی داده: دقت، مقادیر تهی، تمامیت، همسانی، رایج بودن، بجا بودن، توافق مصرف، هماهنگی، پیوستگی ارائه: مناسب بودن، صحیح قابلیت تفسیر، انعطاف‌پذیری، دقت قالب، قابلیت حمل، همسانی، استفاده از انباره، کاربرد اطلاعاتی: قابلیت دسترسی، فراداده، حریم خصوصی، امنیت، حشو، هزینه
DaQuinCIS	دقت، تمامیت، همسانی، رایج بودن، قابلیت اعتماد
QAFD	دقت معنایی/نحوی، همسانی داخلی/خارجی، تمامیت، رایج بودن، منحصربه‌فرد بودن
CDQ	طرح: درستی برحسب مدل، درستی برحسب الزامات، تمامیت، اقتضا، قابلیت خواندن، نرمال‌سازی، داده: دقت معنایی/نحوی، دقت معنایی، تمامیت، همسانی، رایج بودن، بجا بودن، فراریت، قابلیت تکمیل، شهرت، قابلیت دسترسی، هزینه

توجه کنید که روش ارزیابی کیفیت اطلاعات فقط از معیارهای ذهنی برای ارزیابی ابعاد کیفی استفاده می‌کند. در این روش کیفیت داده‌ها عمدتاً برحسب پرسشنامه‌هایی که دارای ۴ تا ۶ فیلد مستقل برای هر بعد کیفی هستند، ارزیابی می‌شود. فیلدها دارای ساختار کلی به شرح زیر می‌باشند. "این اطلاعات (وصفی^۱ یا توضیحی^۲) هستند." مثلاً تمامیت مرتبط با شش فیلد است که برخی از آن‌ها عبارت‌اند از:

۱. این اطلاعات تمام مقادیر ضروری را فراهم می‌کند.
۲. این اطلاعات برای نیازهای ما کافی است.
۳. این اطلاعات، نیازهای فعالیت‌های ما را تأمین می‌کند.

^۱ Attribute

^۲ Phrase

جدول ۲-۱۰. ابعاد و معیارها

ابعاد	نام	تعریف معیارها
دقت	Acc1	دقت نحوی: این به صورت فاصله میان مقدار ذخیره شده در پایگاه داده و درستی، اندازه گیری می شود
	Acc2	دقت معنایی = تعداد مقادیر صحیح/تعداد کل مقادیر
	Acc3	تعداد عناصر دقیق تحویل داده شده
تمامیت	Acc3	پیمایش کاربر - پرسشنامه
	Compl1	تمامیت = تعداد مقادیر غیر تهی/تعداد کل مقادیر
	Compl2	تمامیت = تعداد عناصر تحویل داده شده/تعداد مورد انتظار
	Compl3	تمامیت داده های وب = $(Completeness_{max} - (T_{max} - T_{min}) * Completeness_{min}) / 2$
همسانی	Compl4	پیمایش کاربر - پرسشنامه
	Cons1	همسانی = تعداد مقادیر همسان/تعداد کل مقادیر
	Cons2	تعداد عناصر نقض کننده محدودیت ها، تعداد تفاوت های برنامه نویسی
	Cons3	تعداد صفحات با انحراف از راهنمای سبک
بجا بودن	Cons4	پیمایش کاربر - پرسشنامه
	Time1	بجا بودن = $(\max(0; 1 - \text{Currency/Volatility}))$
	Time2	درصد اجرای فرآیندها قابل انجام در چارچوب زمانی مورد نظر
	Time3	پیمایش کاربر - پرسشنامه
رایج بودن	Curr1	رایج بودن = زمانی که در آن داده ها در سیستم ذخیره می شود - زمانی که در آن داده ها در دنیای واقعی به روزرسانی می شود
	Curr2	زمان آخرین به روزرسانی
	Curr3	رایج بودن = زمان درخواست - آخرین به روزرسانی
	Curr4	رایج بودن = سن + (زمان درخواست - زمان ورود)
	Curr5	پیمایش کاربر - پرسشنامه

فراریت	Vol1	طول زمان برای آنکه داده‌ها معتبر باقی بمانند
ابعاد	نام	تعریف معیارها
منحصربه‌فردی	Uni1	تعداد تکثیرها
حجم مناسب داده	Appr1	حجم مناسبی از داده‌ها = حداقل ((تعداد واحدهای داده‌ای فراهم‌شده/تعداد واحدهای داده‌ای موردنیاز؛ (تعداد واحدهای داده‌ای موردنیاز/تعداد واحدهای داده‌ای فراهم‌شده))
	Appr2	پیمایش کاربر - پرسشنامه
قابلیت دسترسی	Access1	قابلیت دسترسی = ماکزیمم (۰؛ ۱- (زمان تحویل - زمان درخواست)/(ضرب‌الاجل - زمان درخواست))
	Access2	تعداد پیوندهای نامعتبر - تعداد مهارهای نامعتبر
	Access3	پیمایش کاربر - پرسشنامه
اعتبار	Cred1	تعداد عناصر با مقادیر پیش‌فرض
	Cred2	پیمایش کاربر - پرسشنامه
قابلیت تفسیر	Inter1	تعداد عناصر با داده‌های قابل تفسیر، مستندسازی مقادیر کلیدی
	Inter2	پیمایش کاربر - پرسشنامه
قابلیت استفاده	Usa1	پیمایش کاربر - پرسشنامه
یکپارچگی اشتقاق	Integr1	درصد محاسبات صحیح داده‌های مشتق شده بنا به تعریف محاسبه یا فرمول مشتق
موجز بودن	Conc1	تعداد صفحات عمیق (به شدت وابسته به سلسله‌مراتب)
	Conc2	پیمایش کاربر - پرسشنامه
قابلیت نگهداشت	Main1	تعداد صفحات با فرا اطلاعات از دست‌رفته
عملی بودن	App1	تعداد صفحات جدا افتاده
	App2	پیمایش کاربر - پرسشنامه
سهولت	Conv1	مسیرهای ناوبری دشوار: تعداد مسیرهای ناوبری گم‌شده/معیوب

سرعت	Speed1	زمان پاسخ سرور و شبکه
جامعیت	Comp1	پیمایش کاربر - پرسشنامه
ابعاد	نام	تعریف معیارها
روشنی	Clar1	پیمایش کاربر - پرسشنامه
قابلیت ردگیری	Trac1	تعداد صفحات بدون مؤلف یا منبع
امنیت	Sec1	تعداد ورودهای ضعیف
	Sec2	پیمایش کاربر - پرسشنامه
درستی	Corr1	پیمایش کاربر - پرسشنامه
عینی بودن	Obj1	پیمایش کاربر - پرسشنامه
مربوط بودن	Rel1	پیمایش کاربر - پرسشنامه
شهرت	Rep1	پیمایش کاربر - پرسشنامه
سهولت اجرا	Ease1	پیمایش کاربر - پرسشنامه
تعامل پذیری	Interact1	تعداد فرم‌ها - تعداد صفحات شخصی سازی شده

جدول ۲-۱۱. روش‌ها و معیارهای کیفی

مقدار / بعد	CDQ	QAfD	DaQuinCI _S	COLDQ	AMEQ	ISTAT	IQM	DQA	CIHI	AIMQ	TIQM	DWQ	TDQM	
۹/۱۳	+	+	+	+	+	+		+			+		+	Acc1
۱/۱۳												+		Acc2
۲/۱۳							+		+	+				Acc3
۷/۱۲	+	+	+	+	+			+		+			+	Compl1
۲/۱۲											+	+		Compl2
۱/۱۲	+													Compl3
۲/۱۲											+			Compl4
۶/۱۰								+					+	Cons1
۱/۱۰												+		Cons2
۱/۱۰							+							Cons3
۲/۱۰									+	+				Cons4
۳/۷	+			+				+						Time1
۲/۷				+							+			Time2
۲/۷									+	+				Time3
۲/۸			+				+							Curr1
۲/۸		+					+						+	Curr2
۱/۸									+	+				Curr3
۱/۸	+													Curr4
۲/۸									+	+				Curr5
۲/۲	+												+	Vol1
۱/۲		+									+			Uni1
۱/۲								+						Appr1

	TDQM	DWQ	TIQM	AIMQ	CIHI	DQA	IQM	ISTAT	AMEQ	COLDQ	DaQuinCIS	QAfD	CDQ	مقدار / بعد
Appr2				+										۱/۲
Access1							+							۱/۴
Access2								+						۱/۴
Access3			+	+										۲/۴
Cred1		+												۱/۲
Cred2				+										۱/۲
Inter1		+												۱/۲
Inter2				+										۱/۲
Usa1			+											۱/۱
Integr1			+											۱/۱
Conc1														۱/۲
Conc2							+							۱/۲
Main1				+										۱/۱
App1														۱/۱
App2							+							۱/۱
Conv1							+							۱/۱
Speed1							+							۱/۱
Comp1				+			+			+				۳/۳
Clar1			+				+			+				۳/۳
Trac1							+							۱/۱
Sec1							+							۱/۱
Sec2				+										۱/۱
Corr1							+							۱/۱

	TDQM	DWQ	TIQM	AIMQ	CIHI	DQA	IQM	ISTAT	AMEQ	COLDQ	DaQuinCIS	QAFD	CDQ	مقدار / بعد
Obj1				+										۱/۱
Rel1				+										۱/۱
Rep1				+										۱/۱
Ease1				+										۱/۱
Interact1							+							۱/۱

معیارهای گردآوری شده با روش روش شناسی کیفیت انبار داده دارای رابطه‌ی تنگاتنگی با فرآیندهای استفاده کننده از داده‌ها هستند. تمام ابعاد بر اساس میزان انتظارات کاربران فرآیندی خاص ارزیابی شده‌اند. مثلاً جامعیت به صورت نسبت مقادیر غیر تهی به کل تعداد مقادیر مورد نیاز یک فرآیند خاص تعریف می شود و نه با کل تعداد مقادیر ذخیره شده در پایگاه داده. (بخش ۲,۳ را ببینید).

جدول ۲-۱۲. مقایسه میان دسته‌بندی ENGLISH و LOSHIN

روش‌ها/دسته‌بندی هزینه‌ها	TIQM	COLDQ
هزینه ارزیابی فعالیت‌های کیفیت داده	هزینه ارزیابی یا بازرسی: • هزینه نرم‌افزار تحلیل کیفیت اطلاعات • زمان افراد در فرآیند ارزیابی	هزینه‌های کشف
هزینه بهبود فعالیت‌های کیفیت داده	هزینه بهبود فرآیند و جلوگیری از عیب	هزینه‌های اصلاح داده‌ها هزینه‌های نگهداشت داده‌ها: • هزینه‌های اکتساب بالاسری • هزینه‌های واپاشی • هزینه‌های زیرساخت هزینه‌های بهبود فرآیند
هزینه‌های فرآیند کیفیت پایین داده‌ها	هزینه‌های خرابی فرآیند: • هزینه‌های غیرقابل‌بازیابی • هزینه‌های دیون و افشا • هزینه‌های بازیابی مشتریان ناراضی ضایعات اطلاعاتی و استفاده مجدد: • هزینه‌های سامان دادن داده‌های زائد و پشتیبانی • هزینه‌های شکار یا تعقیب اطلاعات ازدست‌رفته • هزینه‌های کار مجدد تجاری • هزینه‌های رفع مشکلات و کاهش بهره‌وری • هزینه‌های تأیید داده‌ها • هزینه‌های بازنویسی نرم‌افزار • هزینه‌های حذف و اصلاح داده‌ها • هزینه‌های نرم‌افزار حذف داده‌ها	تأثیرات اجرایی: • هزینه‌های عقب‌گرد • هزینه‌های کار مجدد • هزینه‌های جلوگیری • هزینه‌های وارانتی اثرات تاکتیکی و استراتژیک: • هزینه‌های تأخیر • هزینه‌های حق تقدم • هزینه‌های بی‌باری • هزینه‌های افزایش دشواری • هزینه‌های از بین رفتن دشواری • هزینه‌های بدگمانی سازمانی • هزینه‌های ناهمترازی

روش‌ها/دسته‌بندی هزینه‌ها	TIQM	COLDQ
هزینه‌های فرصت کیفیت پایین داده‌ها	هزینه‌های فرصت ازدست‌رفته یا نادیده انگاشته شده: • هزینه‌های فرصت ازدست‌رفته • هزینه‌های فرصت نادیده انگاشته شده • ارزش ازدست‌رفته سهامدار	هزینه‌های منفعت ازدست‌رفته: • هزینه‌های تنیدگی • هزینه‌های کاهش • هزینه‌های استهلاک • هزینه‌های توقیف

۶-۲ روش‌ها و هزینه‌ها

بعد هزینه فقط در روش‌های مدیریت کیفیت اطلاعات کلی، ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها و روش جامع برای مدیریت کیفیت داده در نظر گرفته شده است. در این بخش هزینه‌ها را از دونقطه نظر متفاوت تحلیل می‌کنیم:

۱. دسته‌بندی‌های هزینه

۲. معیارهای کمی کردن^۱ هزینه

۱-۶-۲ دسته‌بندی هزینه‌ها

روش مدیریت کیفیت اطلاعات کلی (English 1999) و روش ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها (Loshin 2004) جزئیاتی را برای دسته‌بندی هزینه‌ها ارائه کردند. دسته‌بندی سوم توسط Eppler و Helfert (۲۰۰۴) ارائه شده است.

جدول ۱۲-۲، دسته‌بندی روش‌های مدیریت کیفیت اطلاعات کلی و ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها را مقایسه می‌کند. در روش مدیریت کیفیت اطلاعات کلی به علت کیفیت پایین داده‌ها، هزینه کیفیت داده‌ها متناظر با هزینه فرآیندهای تجاری و مدیریت داده است. هزینه‌های ارزیابی یا بازرسی کیفیت

^۱ Quantification

اطلاعات، ابعاد کیفیت داده‌ها را برای تأیید اینکه فرآیندها در حال عملکرد مطلوب هستند یا خیر، اندازه‌گیری می‌کند. سرانجام هزینه‌ی بهبود فرآیند یا جلوگیری از خرابی مطرح می‌شود که شامل فعالیت‌هایی برای بهبود کیفیت داده‌ها باهدف حذف یا کاهش هزینه‌های کیفیت پایین داده‌ها است. در روش مدیریت کیفیت اطلاعات کلی هزینه‌ها به علت کیفیت پایین داده‌ها تحلیل شده و به سه زیرگروه تقسیم می‌شوند:

۱. **هزینه‌های خرابی فرآیند.** زمانی تحمیل می‌شوند که داده‌های باکیفیت پایین موجب فرآیندی نادرست شوند. مثلاً نشانی پستی غیردقیق منجر به تحویل اشتباه محموله‌ها می‌شود.
 ۲. **ضایعات اطلاعاتی و کار مجدد.** زمانی که داده‌ها دارای کیفیت پایینی باشند، نیازمند انواع مختلف فعالیت‌های مدیریت خرابی مانند کار مجدد، حذف یا بازگردانی می‌شوند.
 ۳. **هزینه‌های فرصت از دست رفته.** این هزینه‌ها متناظر با عواید و سودهای از دست رفته ناشی از کیفیت پایین داده‌ها است. مثلاً به علت دقت پایین نشانی ایمیل مشتری، کمپین‌های تبلیغات دوره‌ای نمی‌توانند به درصدی از مشتریان دست پیدا کنند که این امر منجر به کاهش عواید می‌شود.
- روش ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها هزینه‌های کیفیت پایین داده‌ها را تحلیل کرده و آن‌ها را با توجه به اثر حوزه‌ای دسته‌بندی می‌کند:

- **حوزه عملیاتی،** که شامل اجزای سیستم پردازش اطلاعاتی و هزینه‌های عملیاتی سیستم است؛
 - **حوزه تاکتیکی،** که تلاش می‌کند مسائل را پیش از قوت گرفتن آن‌ها، آدرس‌دهی و حل کند؛
 - **حوزه استراتژیک،** که تصمیمات بلندمدت را حمایت می‌کند.
- برای اثرات عملیاتی و اثرات تاکتیکی/استراتژیک، دسته‌بندی هزینه‌ها ارائه شده است. در اینجا ما هزینه‌های عملیاتی را تشریح می‌کنیم:
- **هزینه‌های عقب‌گرد،** زمانی تحمیل می‌شود که کاری که انجام شده است باید به حالت قبلی بازگردانده شود؛

- **هزینه‌های کار مجدد**، زمانی تحمیل می‌شود که یک مرحله پردازش باید تکرار شود؛
 - **هزینه‌های جلوگیری**، زمانی به وجود می‌آید که فعالیت جدیدی پیاده‌سازی می‌شود تا فعالیت ضروری برای جلوگیری از خرابی عملیاتی به علت مسئله کیفیت داده‌ها را انجام دهد؛
 - **هزینه‌های وارانتی**، مربوط به تضمین بازگشت خسارات ناشی از ضرر و زیان است.
- روش جامع برای مدیریت کیفیت داده دسته‌بندی ارائه می‌کند که ناهمگنی میان روش‌های مدیریت کیفیت اطلاعات کلی، ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها و دیدگاه Eppler و Helfert (۲۰۰۴) را برطرف می‌کند. برای جزئیات بیشتر، مقاله Batini و Scannapieco (۲۰۰۶) را ببینید.

۲-۶-۲ معیارهای کمی سازی هزینه‌ها

ارزیابی مجموع هزینه‌ی کیفیت داده‌ها با توجه به انواع فعالیت‌های کیفیت داده که باید انجام شوند و اولویت‌های آن‌ها صورت می‌گیرد (بخش ۲,۲ را ببینید). روش‌های مدیریت کیفیت اطلاعات کلی، ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها و روش جامع برای مدیریت کیفیت داده تنها روش‌هایی هستند که معیارها را برای این نوع فعالیت‌ها فراهم می‌کنند.

انتخاب و اولویت‌بندی در روش مدیریت کیفیت اطلاعات کلی با طی مراحل زیر حاصل می‌شود:

- کاربران فعلی که از داده‌ها استفاده می‌کنند را شناسایی می‌کند؛
- خطاهایی را که اثر منفی بر کیفیت داده‌ها می‌گذارد را فهرست می‌کند؛
- واحدهای تجاری که اغلب اوقات تحت تأثیر داده‌هایی با کیفیت پایین قرار می‌گیرند، شناسایی می‌کند؛
- هزینه مستقیم برنامه‌های فعلی کیفیت داده‌ها را تخمین می‌زند؛
- هزینه خطاهای داده‌ها را برای تمام کاربران تخمین می‌زند و از داده‌ها استفاده می‌کند؛
- از هزینه‌ها برای توجیه و اولویت‌بندی شروع برنامه‌های کیفیت داده استفاده می‌کند که شامل ساختاربندی یک برنامه مستمر بهبود کیفیت است.

هر نوع خطا در یک تناوب زمانی رخ می‌دهد و شامل هزینه‌ای است. به یاد داشته باشید که هزینه‌ی دسته‌های مختلف خطا یک مقدار مشروط است و با تغییر فرآیندی که از داده‌ها استفاده می‌کند، تغییر می‌کند. مدل‌های ارائه‌ی فرآیند، اجازه‌ی شناسایی فعالیت‌های تحت تأثیر خطاهای داده‌ای را می‌دهند. از آنجایی که فعالیت‌ها با مجموع هزینه‌های سازمانی در ارتباط هستند، هزینه کار مجدد می‌تواند تخمین‌های کمی و عینی ایجاد کند.

روش ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها با حمایت از سرمایه‌گذاری در برنامه‌ی مدیریت دانش، بر ارزیابی هزینه‌های مرتبط با کیفیت پایین داده تمرکز دارد. این ارزیابی با طی مراحل زیر حاصل می‌شود:

- نقشه‌ی زنجیره اطلاعاتی را برای درک چگونگی جریان یافتن اطلاعات در سازمان، ترسیم می‌کند؛
- کاربران فعلی که از داده‌ها استفاده می‌کنند را شناسایی می‌کند؛
- اثر کیفیت داده‌ها را بر مشتریان شناسایی می‌کند؛
- داده‌های معیوب را با تخصیص به نواحی حساس ایزوله می‌کند؛
- حوزه‌های متأثر و مرتبط با هر نمونه از کیفیت پایین داده را شناسایی می‌کند؛
- هزینه‌های اقتصادی را بر اساس اثرات نهایی داده‌های بد، مشخص می‌کند؛
- همه‌چیز را برای ارزیابی اثر کلی اقتصادی گردآوری می‌کند؛
- فرصت‌های بهبود را شناسایی می‌کند.

نتیجه این کار با عنوان کارت امتیاز^۱ کیفیت داده‌ها یاد می‌شود. این نتیجه‌گیری هزینه‌های مرتبط با کیفیت پایین داده را خلاصه می‌کند و می‌تواند به عنوان ابزاری برای یافتن بهترین راه‌حل بهبود استفاده شود.

در روش جامع برای مدیریت کیفیت داده معیار اصلی برای انتخاب میان فرآیندهای بهبود، حداقل سازی هزینه برنامه کیفیت داده است. ابتدا فرآیندهای بهبود متفاوت در چارچوب روش‌های داده محور و فرآیند محور اعمال شده و سپس بر اساس داده‌ها، جریان داده‌ها و اساس روش بهبود، روش متناسب شناسایی می‌شود. سپس هزینه فرآیندهای متفاوت ارزیابی و مقایسه شده و فرآیند با حداقل هزینه انتخاب می‌شود.

جدول ۲-۱۳. روش‌ها و انواع داده‌ها

مخفف نوع داده/روش	ساختاریافته	نیمه ساختاریافته
TDQM	X	X
DWQ	X	
TIQM	X	در نظر گرفته شده به طور ضمنی
AIMQ	X	در نظر گرفته شده به طور ضمنی
CIHI	X	X
DQA	X	
IQM	X	X
ISTAT	X	X
AMEQ	X	در نظر گرفته شده به طور ضمنی
COLDQ	X	در نظر گرفته شده به طور ضمنی
DaQuinCIS	X	X
QAFD	X	
CDQ	X	X

۲-۲ روش‌ها و انواع داده‌ها

در بخش ۵,۲ مشاهده کردیم که نوع داده‌ها بر ابعاد کیفیت داده و تکنیک‌های ارزیابی و بهبود اثر می‌گذارند. در جدول ۲-۱۳، سه نوع داده دسته‌بندی شده در بخش ۲,۵ با روش‌های کیفیت داده مرتبط شدند. اکثر روش‌ها، داده‌های ساختاریافته را آدرس‌دهی می‌کنند درحالی که فقط تعداد کمی از روش‌ها داده‌های نیمه ساختاریافته را آدرس‌دهی می‌کنند. در جدول ۲-۱۳ مشاهده می‌شود که هنگامی که روش دقیقاً به نوع داده اشاره نکرده باشد از عبارت "در نظر گرفته شده به طور ضمنی" استفاده می‌شود اما مراحل و قدم‌ها می‌تواند بر آن اعمال شود. مثلاً روش ارزیابی کیفیت اطلاعات از عبارات عمومی استفاده می‌کند و ارزیابی کیفی را از طریق پرسش‌هایی که بر داده‌های ساختاریافته اعمال می‌شود انجام می‌دهد ولی ممکن است به هر نوع داده شامل داده‌های ساختار نیافته نیز اشاره داشته باشد.

درروش ISTAT داده‌های نیمه ساختاریافته، استانداردسازی قالب داده‌ها و بیان آن‌ها را در طرح رایج XML در نظر گرفته می‌شود. این طرح برای حداقل کردن اختلافات میان سازمان‌ها و نحوه‌ی همکاری آن‌ها با قسمت‌های دیگر پیاده‌سازی شده است. درروش کیفیت در سیستم‌های اطلاعات تعاونی مدلی برای ایجاد ارتباط میان مقادیر کیفی و اسناد XML ارائه شده است (Scannapieco et al. 2004). سازمان‌های مختلف در یک سیستم اطلاعاتی تعاونی از مدل داده و کیفیت داده در حوزه جریان‌های داده‌ای تبادل شده استفاده می‌کنند. کیفیت جریان داده برای جلوگیری از انتشار خطا در تبادل داده‌ها حیاتی می‌شود. روش می‌تواند به منظور تصدیق دقت، جامعیت، رایج بودن و همسانی داده‌ها استفاده شود. مدل‌های نیمه ساختاریافته به هر سازمان اجازه می‌دهند، کیفیت داده‌های خود را با درجه مشخصی از انعطاف‌پذیری، تنظیم کنند. کیفیت می‌تواند با عناصر متنوعی از مدل داده‌ها در ارتباط باشد که بازه‌ی آن از مقادیر داده‌های منفرد تا کل منبع داده‌ها متغیر است.

روش جامع برای مدیریت کیفیت داده قدم‌ها و تکنیک‌هایی را برای داده‌های ساختاریافته، توسعه داده و برای داده‌های نیمه ساختاریافته نیز بسط داده است. مثلاً هر نوع داده اعم از ساختاریافته و نیمه ساختاریافته در فاز بازسازی وضعیت، پیمایش شده است. یک تکنیک مشخص بهبود داده محور نیز برای داده‌های غیر ساختاریافته ارائه شده است که نام آن ثبت وضعیت داده‌ها^۱ است. این تکنیک برای مرتبط کردن یک فایل متنی با طرح یک پایگاه داده به وسیله کشف الگوهای تکرارشونده در داخل متن، استفاده می‌شود (Aiken 1996).

جدول ۲- ۱۴. روش‌ها و سیستم‌های اطلاعاتی

وب	تعاونی	انبار داده	توزیع شده	یکپارچه	مخفف نوع سیستم اطلاعاتی/روش
			در نظر گرفته شده به طور ضمنی	متمرکز	TDQM
		به شدت متمرکز			DWQ
			متمرکز	متمرکز	TIQM
			در نظر گرفته شده به طور ضمنی	متمرکز	AIMQ
			متمرکز	متمرکز	CIHI
			در نظر گرفته شده به طور ضمنی	متمرکز	DQA
به شدت متمرکز					IQM
	به شدت متمرکز		متمرکز	متمرکز	ISTAT
				متمرکز	AMEQ
				متمرکز	COLDQ
	به شدت متمرکز		متمرکز	متمرکز	DaQuinCIS
				متمرکز	QAFD
	به شدت متمرکز		متمرکز	متمرکز	CDQ

۸-۲ روش‌ها و انواع سیستم‌های اطلاعاتی

جدول ۱۴-۲ نشان می‌دهد که روش‌های مختلف تا چه حدی با انواع سیستم‌های اطلاعاتی معرفی شده در بخش ۲،۶، سروکار دارند. جدول از یک مقیاس چهار متغیره استفاده می‌کند که:

۱. **کاملاً متمرکز** بدان معناست که کل سازمان درک شده و روش مطابق با آن نوع سیستم اطلاعاتی

طراحی شده است، درحالی که راهنماهای کلی برای سایر سیستم‌های اطلاعاتی، نیز فراهم است،

۲. **متمرکز** بدان معناست که روش، راهنماها و تکنیک‌های دقیقی را برای انواع سیستم اطلاعاتی فراهم می‌کند،

۳. **در نظر گرفته شده به طور ضمنی** دارای همان معنای جدول ۱۲ است

۴. **مقدار خالی در جدول** نشان می‌دهد که روش، راهنماهای کلی فراهم می‌کند و به طور دقیق نوع سیستم اطلاعاتی متناظر را آدرس‌دهی نمی‌کند.

از آنجایی که هر روشی بر روی ابعادی متمرکز است پس سیستم‌های اطلاعاتی وب نیز در جدول گنجانده شده و دیگر مسائل مرتبط با چنین سیستم‌هایی در بخش ۴ مورد بحث قرار خواهد گرفت. ولی هیچ نوع روشی به سیستم‌های P2P اشاره نکرده است، لذا به آن‌ها در بخش نتیجه‌گیری و مسائل آتی پرداخته خواهد شد (بخش ۴).

می‌توان مشاهده کرد که روش‌های اندازه‌گیری مبتنی بر فعالیت و ارزیابی کیفیت اطلاعات محصول، ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها و ارزیابی کیفیت داده‌های مالی بر سیستم‌های اطلاعاتی یکپارچه^۱ تمرکز دارند. این روش‌ها داده‌های ساختاریافته را در یک سیستم منفرد در نظر گرفته و مسائل کیفیت داده‌ی ناشی از تبادل داده میان سازمان‌های مجزا را نادیده می‌گیرند.

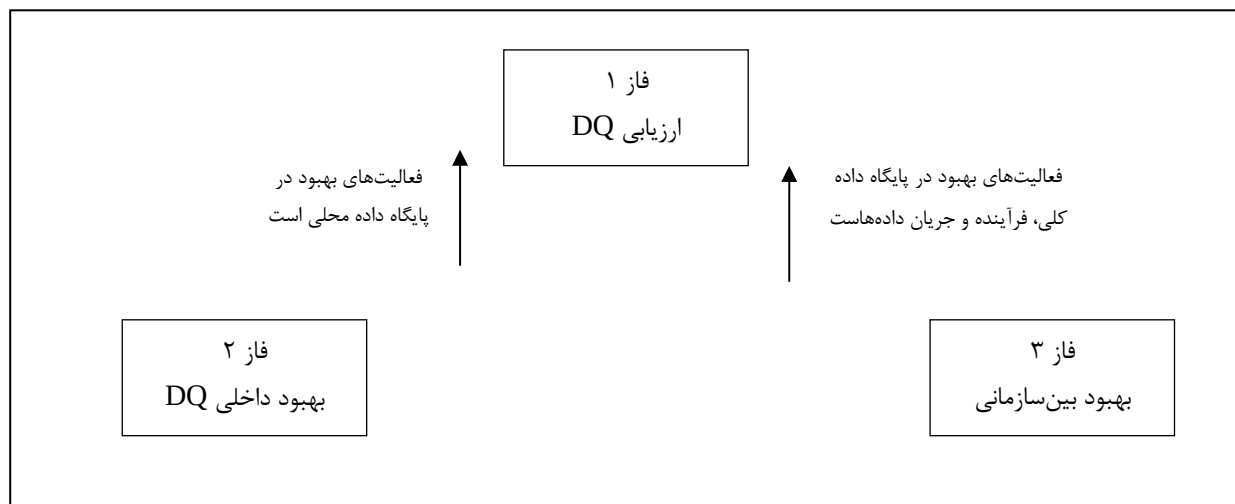
همانگونه که در جدول ۱۴-۲ مشخص است برخی روش‌ها مانند روش DWQ و IQM در یک نوع سیستم اطلاعاتی متمرکز است. بعنوان مثال مدل DWQ فقط بر انبار داده‌ها متمرکز بوده و در این دسته از

سیستم‌های اطلاعاتی کاربرد دارد. از سوی دیگر برخی از روش‌ها مانند TDQM، ISTAT و CDQ در انواع مختلفی از سامانه‌های اطلاعاتی متمرکز هستند. به عنوان مثال روش TDQM در سیستم اطلاعاتی یکپارچه و توزیع شده کاربرد دارد و این نشاندهنده جامعیت و پوشش بالای این روش است.

روش مدیریت کیفیت اطلاعات کلی می‌تواند بر سیستم‌های یکپارچه و سیستم‌های توزیعی اعمال شود زیرا هنگامی که ساختار داده‌ها تحلیل شود، سیستم مرکزی و توزیعی فراهم می‌شود. روش موسسه کانادایی برای روش اطلاعات سلامت بر سیستم‌های توزیعی تمرکز دارد و استفاده از آن در پایگاه‌های داده ملی است.

سایر روش‌ها یعنی ISTAT، ارزیابی کیفیت در سیستم‌های اطلاعات تعاونی و روش جامع برای مدیریت کیفیت داده می‌توانند بر چند نوع سیستم اعمال شوند. هر روش راهنماهای دقیقی را برای پیچیده‌ترین نوع سیستم‌ها (یعنی سیستم اطلاعات تعاونی) فراهم می‌کند. بنابراین می‌تواند بر سیستم‌های یکپارچه و توزیعی نیز اعمال شود. این روش‌ها از رویکردهای متفاوتی برای سامان دادن چند جریان داده و پایگاه‌های داده‌ی ناهمگن استفاده می‌کنند.

روش ISTAT دارای رویکرد قوی بین سازمانی بوده و برای مدیریت دولتی ایتالیا استفاده شده است. این روش دارای ساختار توزیعی با مدیریت کاملاً خودمختار برای شناسایی و انجام روابط است (این حالت کلی برای بسیاری از کشورها است). سه فاز اصلی روش در شکل ۲-۳ آمده است.



شکل ۲-۳. طرح کلی روش ISTAT (Batini et al. 2009)

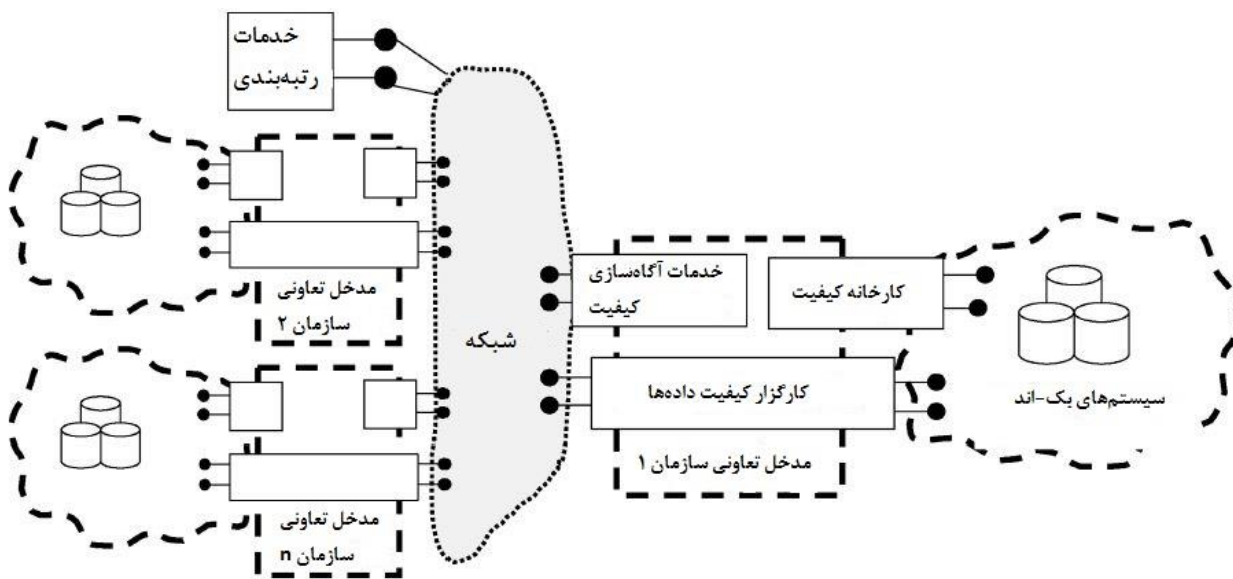
فاز ارزیابی (فاز ۱) مرتبط‌ترین فعالیت‌هایی که باید در فاز بهبود (فاز ۲ و ۳) انجام شوند را به شرح زیر شناسایی می‌کند:

۱. فاز ۲ بر روی پایگاه‌های داده‌ی محلی که برای مدیریت‌های متفاوت طراحی شده عمل می‌کند. در این فاز ابزارهایی طراحی شده تا فعالیت‌های کیفیت داده را به‌طور خودمختار انجام دهند و روندهایی نیز برای بهبود مهارت‌های محلی کیفیت داده ارائه شده‌اند.
۲. فاز ۳ سیستم کلی اطلاعاتی تعاونی را برحسب جریان داده تبادل شده در نظر می‌گیرد و پایگاه داده-ی مرکزی برای اهداف تعاونی، تنظیم می‌کند. این فعالیت‌ها به‌طور مرکزی، برنامه‌ریزی شده و هماهنگ انجام می‌گیرد.

روش ارزیابی کیفیت در سیستم‌های اطلاعات تعاونی چارچوبی فراهم می‌کند که در آن خدمات و ابزارهای متنوعی برای سیستم‌های اطلاعاتی تعاونی ارائه می‌شوند (شکل ۲-۴). کارگزاران^۱ کیفیت داده‌ها که در بخش ۳،۲ توصیف شده، هسته معماری و مسئول اصلاح پاسخ‌های ناهمگن به تفسیرهاست. خدمات آگاه‌سازی کیفیت یک موتور انتشاری/اشتراکی است که به‌عنوان پیام عمومی میان اجزای معماری سازمان‌های همکار

استفاده می‌شود (Scannapieco et al. 2004). در سازمان‌ها به خاطر تغییرات کیفیت داده‌ها از این موتور استفاده می‌شود و این موتور اجازه می‌دهد اشتراکی مبتنی بر کیفیت صورت گیرد. سپس بهترین مقدار کیفیت انتخاب شده توسط کارگزار در سازمان‌ها انتخاب می‌شود تا سازمان داده‌های خود را با داده‌های با کیفیت بالاتر جایگزین کند (تابع بهبود کیفیت). چارچوب کارخانه کیفیت (شکل ۲-۴) مسئول ارزیابی کیفیت داده‌های داخلی هر سازمان است (برای مطالعه بیشتر، مقاله (Cappiello et al. 2003) را بخوانید). در این چارچوب درخواست از کاربران بیرونی یا سیستم‌های اطلاعاتی در کارخانه کیفیت به وسیله تحلیلگر کیفیت پردازش می‌شود و یک تحلیل ایستا از ابعاد کیفیت داده‌های مرتبط با داده‌های درخواستی ارائه می‌شود. سپس آن‌ها را با پارامترهای کیفیت محک^۱ موجود در انبار کیفیت مقایسه می‌کنند. خدمات رتبه‌بندی نهایی در بخش ۴ بررسی خواهند شد.

پژوهش‌های دیگری که از ابزارهای پردازش تفسیری مبتنی بر کیفیت و حل تعارض سطح نمونه‌ای برای سیستم‌های اطلاعاتی تعاونی ارائه کرده‌اند، فیوژن پلکس (Motro and Anokhin 2005)، آی‌فویس (Rahm et al. 2005) و هامر (Bilke et al. 2005) هستند.



شکل ۲-۴. معماری چارچوب ارزیابی کیفیت در سیستم‌های اطلاعات تعاونی (BATINI ET AL. 2009)

در این قسمت به روش‌هایی که بر انواع سیستم‌های اطلاعاتی خاص یعنی روش‌شناسی کیفیت انبار داده و اندازه‌گیری کیفیت اطلاعات کاملاً متمرکز هستند پرداخته می‌شود. روش روش‌شناسی کیفیت انبار داده مخصوص انبارهای داده است. در تحقیق (Jarke et al. 1995) مشاهده می‌شود که اکثر محققان در بررسی‌های خود بر روی روش کیفیت انبار داده، به نقش میانجی‌های میان سیستم‌های OLTP و پشتیبانی تصمیم تمرکز کرده‌اند. این مسئله نقش سازمانی انبار کردن داده‌ها را برای کنترل مرکزی جریان‌های اطلاعاتی نادیده می‌گیرد. به تبع آن، تعداد زیادی از مسائل کیفیت مربوط به روش کیفیت انبار داده نمی‌توانند با فرا مدل‌های سنتی روش کیفیت انبار داده بیان شوند. روش روش‌شناسی کیفیت انبار داده دو تغییر در نحوه-ی حل این نوع مسائل انجام می‌دهد. ابتدا فراداده‌ها در معماری روش کیفیت انبار داده با مدل‌های صریح بنگاهی، غنی می‌شوند. سپس تکنیک‌های ریاضیاتی برای اندازه‌گیری یا بهینه‌سازی جنبه‌های متنوع کیفیت روش کیفیت انبار داده ارائه می‌شود.

ابزار آرتکوس نیز می‌تواند برای آدرس‌دهی مسائل کیفیت داده مشارکت کند (Vassiliadis et al. 2001). آرتکوس، فرا مدلی تشکیل شده از چند عنصر ارائه می‌دهد که عبارت‌اند از:

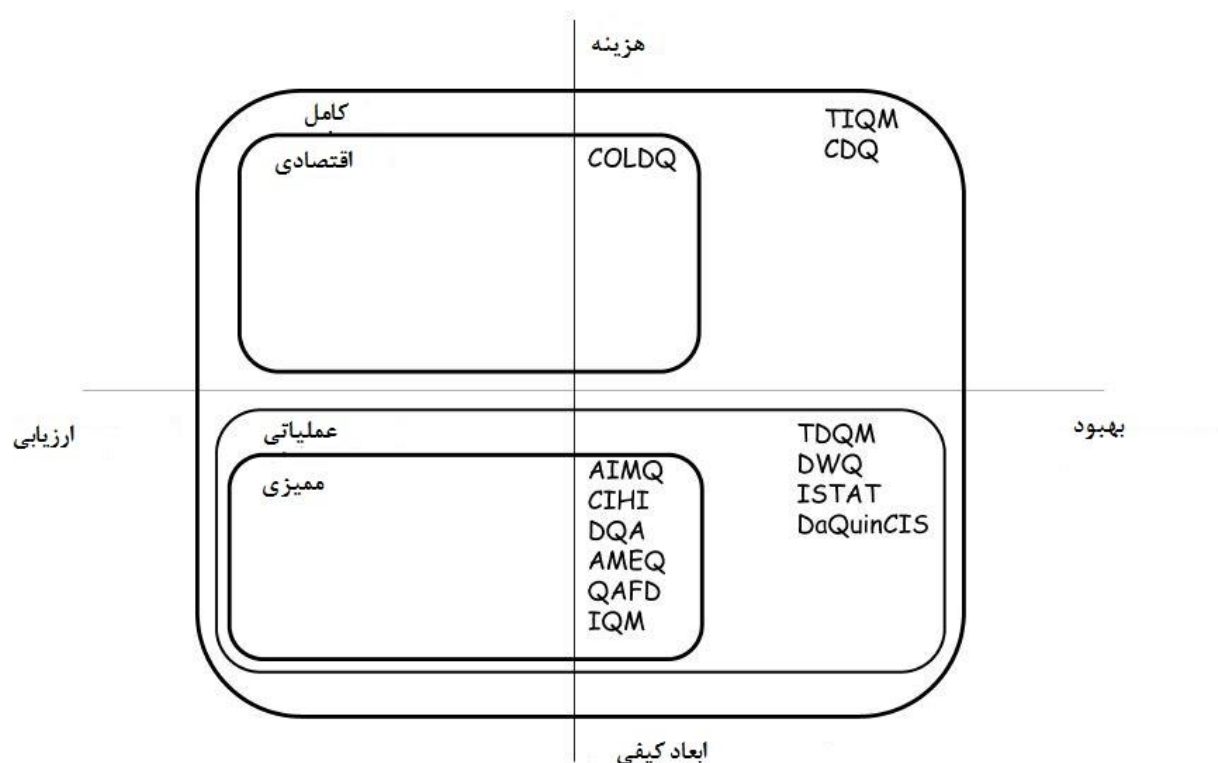
- (۱) **فعالیت‌ها**، واحدهای اتمی فعالیت پردازش داده‌ها؛
- (۲) **سناریوها**، مجموعه از فعالیت‌های که باهم اجرا می‌شوند؛
- (۳) **کارخانه کیفیت**، که برای هر فعالیت، متناظر با ابعاد و سنجه‌های تعریف‌شده در بخش ۲،۳ تعریف‌شده است.

روش اندازه‌گیری کیفیت اطلاعات بر سیستم‌های اطلاعاتی وب کاملاً متمرکز است، چرا که مجموعه‌ی وسیعی از ابزارهای موجود را برای ارزیابی کیفیت اطلاعات در حوزه‌ی وب در نظر می‌گیرد. ازجمله این ابزارها می‌توان به تحلیلگران وب، تحلیلگران ترافیک، اسکنرهای درگاه، سیستم‌های رصد عملکرد، ابزارهای وب کاوی و ابزارهای پیمایش برای تولید بازخورد مبتنی بر عقاید کاربر اشاره کرد. معیارهای کیفیت اطلاعاتی متنوعی می‌توانند با کمک این ابزارها اندازه‌گیری شوند. روش اندازه‌گیری کیفیت اطلاعات قدم‌های پی‌درپی سیستماتیکی را برای تطبیق دادن معیارهای کیفیت اطلاعات با ابزارهای اندازه‌گیری فراهم می‌کند.

۲-۹ مقایسه‌ی مجمل روش‌ها

مقایسه‌ی جزئی روش‌ها در بخش‌های قبلی به‌روشنی نشان داد که روش‌ها بر روی زیرمجموعه‌ای از مسائل کیفیت داده تمرکز دارند. تفاوت‌های وسیع در نقاط مورد تمرکز روش‌ها را می‌توان از دیدگاهی در چهار گروه تقسیم‌بندی کرد که در شکل زیر دیده می‌شود. این گروه‌ها عبارت‌اند از:

- **روش‌های کامل** که فاز ارزیابی و فاز بهبود را پشتیبانی می‌کنند، همچنین مسائل تکنیکی و اقتصادی را هم پوشش می‌دهند؛
- **روش‌های ممیزی** که بر فاز ارزیابی تمرکز دارند و پشتیبانی محدودی از فاز بهبود ارائه می‌دهند؛
- **روش‌های عملیاتی** که بر مسائل تکنیکی فازهای ارزیابی و بهبود متمرکز هستند اما مسائل اقتصادی را پوشش نمی‌دهند.
- **روش‌های اقتصادی** که بر هزینه‌های ارزیابی تمرکز دارند.



شکل ۲-۵. یک دسته‌بندی از روش‌ها

از نظر تاریخی، میان ابعاد کیفی و تکامل تکنولوژی‌های ICT همبستگی وجود دارد. نسل اول سیستم‌های اطلاعاتی (دهه‌های ۵۰ و ۶۰) دارای معماری تکنولوژیک ویژه بودند و شامل یک پردازنده‌ی مرکزی و یک پایگاه داده‌ی یکپارچه بودند. جریان‌های اطلاعاتی در آن‌ها ساده و تکراری بوده و اکثر خطاها به وسیله داده‌های نادرست به وجود می‌آمد. از جمله کاربردهای مهم و در مقیاس بزرگ آن‌ها مدیریت سرشماری و تحلیل داده‌های درمانی بود و کیفیت داده‌ها بر دقت، همسانی، جامعیت و ابعاد مرتبط با زمان تمرکز داشت (Catarci and Scannapieco 2002). حیاتی‌ترین مسائل در مدیریت کیفیت این داده‌ها، مکان‌یابی و اصلاح خطا در منابع داده و ثبت پیوند میان منابع داده جدید و منابع موجود بود.

تکامل سیستم‌های اطلاعاتی از یکپارچه به حالت مبتنی بر شبکه، سبب رشد تعداد منابع داده هم در اندازه و هم در دامنه شد و به تبع آن، پیچیدگی مدیریت کیفیت داده‌ها به شدت افزایش یافت. روش‌های کیفیت داده بر ابعاد جدید کیفیت نظیر جامعیت منبع داده، رایج بودن داده‌ها و همسانی منابع جدید داده با پایگاه داده

بنگاهی تمرکز می‌کنند. با ظهور وب و سیستم‌های اطلاعاتی نظیر آن ارزیابی و کنترل منابع داده دشوار شده است؛ چراکه جستجو از طریق وب نامحدود است. به خاطر چنین تکامل سریعی، روش‌ها شروع به آدرس‌دهی ابعاد کیفی جدید نظیر قابلیت دسترسی و شهرت کردند. قابلیت دسترسی، توانایی کاربران را برای دسترسی به داده‌ها بر اساس فرهنگ و وضعیت فیزیکی آن‌ها و تکنولوژی‌های موجود اندازه می‌گیرد و این معیار در سیستم‌های اطلاعاتی تعاونی و مبتنی بر شبکه مهم است. شهرت (یا قابلیت اعتماد) معیاری در منابع داده است که توانایی آن‌ها را برای فراهم کردن اطلاعات صحیح را اندازه می‌گیرد و مشخصاً مرتبط با سیستم‌های مبتنی بر وب و هم‌تا به هم‌تا^۱ است که در بخش ۴ بیشتر توضیح داده شده است.

تکامل ICT و رشد تبعی پیچیدگی‌های کیفیت داده دلایل اصلی برای تمرکز روش‌ها در زیرمجموعه‌ای از مسائل کیفیت داده است، همان‌طور که در شکل ۵ نشان داده شده است. اکثر روش‌ها متعلق به دسته‌ی ممیزی هستند و بر مسائل تکنیکی فاز ارزیابی تمرکز دارند. این امر نیاز به ارزیابی کیفیت داده‌ها را ضروری می‌سازد؛ چراکه این امر به عنوان بخشی از فعالیت‌های بهبود برای ارزیابی اثربخشی تکنیک‌های بهبود است. بحث ارزیابی کیفیت جزو مسائل پیچیده و بدیع در حوزه‌ی کیفیت داده است که مشخصاً از دید علمی جالب توجه هستند. مثلاً تعریف معیارهایی نظیر جامعیت و دقت، به شدت وابسته به تئوری پایگاه داده است و دامنه مسائل طراحی پایگاه داده سنتی نظیر دید یکپارچه و نرمال‌سازی طرح را وسیع می‌کند. در مقابل، مسائل اقتصادی ارزیابی و بهبود نیازمند یک رویکرد تجربی است. آزمایش تکنیک‌ها در جهان واقعی می‌تواند چالش برانگیز باشد و فقط تعداد کمی از روش‌ها بر شواهد تجربی و اقتصادی کیفیت داده تمرکز دارند.

باید یادآور شد که روش‌های ممیزی در فاز ارزیابی، دقیق‌تر از روش‌های کامل و عملیاتی هستند. آن‌ها با فراهم کردن مثال‌های بیشتر و دانشی مرتبط، اطلاعات دقیق‌تری را درباره‌ی نحوه انتخاب تکنیک‌های مناسب ارزیابی ارائه می‌دهند. سپس آن‌ها هر نوع مسئله را صرف‌نظر از تکنیک‌های بهبودی که می‌تواند اعمال شود، شناسایی می‌کند. مثلاً روش‌های ارزیابی کیفیت اطلاعات و ارزیابی کیفیت داده‌های مالی به‌طور دقیق توضیح داده‌اند که ارزیابی عینی و غیر عینی چگونه می‌تواند انجام شود و راهنمایی را برای تفسیر نتایج ارائه

می‌دهد. روش ارزیابی کیفیت داده درباره‌ی تعاریف عملیاتی که می‌تواند برای اندازه‌گیری ابعاد متفاوت کیفیت داده برای ارزیابی معیارهای تجمعی در پایگاه‌های داده و منابع داده استفاده شود، بحث می‌کند.

روش‌های عملیاتی بر ارزیابی کیفیت داده و شناسایی مسائلی که رویکرد بهبودشان بهترین عملکرد را دارد، تمرکز می‌کنند. یکی از وظایف اصلی، شناسایی مجموعه‌ای از ابعاد مرتبط با بهبود و توصیف روش‌هایی برای ارزیابی آن‌ها است. مثلاً روش مدیریت کیفیت داده یک روش باهدف عمومی است و مجموعه‌ی کاملی از ابعاد مرتبط و روش‌های بهبودی را که می‌توانند در حوزه‌های مختلف اعمال شوند، ارائه می‌کنند. وقتی که روش‌ها بر حوزه‌های مشخصی تمرکز می‌کنند، جامعیت کاهش می‌یابد. مثلاً روش روش‌شناسی کیفیت انبار داده حوزه انبار داده را تحلیل کرده و ابعاد کیفی جدیدی را مطابق با معماری انبار داده تعریف می‌کند. تهیه‌ی فهرست ابعاد مرتبط، نقطه شروع مهمی برای فرآیند بهبود است چراکه از شرکت‌ها در شناسایی مسائل کیفیت داده که بر انبار داده آن‌ها اثر دارند، پشتیبانی می‌کند. توجه کنید که روند ارزیابی به‌طور دقیق در روش‌های عملیاتی که بر حوزه‌ی مشخصی تمرکز دارند توصیف شده‌اند. بنابراین تخصصی کردن روش‌های عملیاتی، جامعیت و عملیاتی بودن آن‌ها را در مقایسه با روش‌های کامل کاهش می‌دهد اما کارایی تکنیک‌های ارائه‌شده را افزایش می‌دهد.

همچنین روش‌های عملیاتی زمانی که بر یک مسئله‌ی کیفیت داده خاص تمرکز می‌کنند، مؤثرتر هستند (خصوصاً برای فرآیندهای داده محور). مثلاً روش ISTAT بر محلی کردن داده‌ها تمرکز دارد که آدرس‌های موجود در یک قلمروی مدیریتی را با داده‌های شخصی افراد مقیم، مرتبط می‌کند. چنین داده‌هایی به‌طور متوالی در میان آرانس‌ها تبادل می‌شوند. برای این کار روش ISTAT قالب‌های ناهمگن متفاوتی را به خدمت می‌گیرد تا بتواند برای موضع‌یابی داده‌ها و صاحبان داده‌های حوزه‌های متفاوت، اطلاعاتی نظیر کد پستی را مصاحبه کرده و یک قالب متحد و رویکردی جدید را برای تبادل داده‌هایی بر اساس زبان XML ارائه کند.

به عنوان مثال دیگری، اگر مسائل کیفیت داده به دقیق بودن و جامعیت داده های شخصی مرتبط باشد، روش های بهبود می توانند در هدف گیری تکنیک های ثبت، دقیق تر عمل کنند؛ این روشی است که بر اساس آن روش های ISTAT و ارزیابی کیفیت در سیستم های اطلاعات تعاونی به فراهم کردن الگوریتم هایی برای کشف موارد مشابه در پیوند ثبت و یکپارچه کردن منابع مختلف می پردازند. مثلاً حذف نام تکراری خیابان ها در نواحی دوزبانی نظیر ناحیه آلتو آدیج در ایتالیا با به خدمت گرفتن توابع مشابهت و خطاهای پارادایمی نظیر استفاده از u به جای $\bar{u}$ ، انجام شده است.

مسئله دیگری که در روش های عملیاتی آدرس دهی می شود، نظارت بر فرآیندهای کیفیت داده، مسائل شدنی و مدیریت ریسک های مرتبط است. تفاوت اصلی روش ISTAT نسبت به سایر روش ها، تمرکز آن بر روی مسائلی با مقیاس بزرگ در تمام حوزه های مدیریت های مرکزی یک کشور است. مثلاً واحدهای سازمانی پراکنده و توزیع شده در کل قلمرو که وابسته به آژانس های مرکزی هستند. معمولاً چنین گروهی از مدیریت ها دارای خصوصیات کلیدی زیر هستند:

۱- پیچیدگی بالا برحسب رابطه متقابل، فرآیندها و خدماتی که آژانس های متفاوت به علت جدا بودن

دارند؛

۲- سطح بالایی از خودمختاری که اجرای قوانین را دشوار می کند؛

۳- ناهمگنی زیاد معنا و ارائه داده و جریان داده؛

۴- همپوشانی زیاد میان ثبت ها و موارد ناهمگن.

بهبود کیفیت داده در چنین ساختار پیچیده ای معمولاً یک فعالیت بسیار پر ریسک و هزینه بر است. این مطلب خود دلیلی است برای آنکه چرا هدف اصلی روش ISTAT، شدنی بودن راه حل است. بنابراین توجه اصلی متمرکز بر متداول ترین نوع داده های تبادل شده میان آژانس ها است. در اولین مرحله بر روی مرتبط ترین داده ها و جریان داده ها برای کشف مهم ترین نواحی ارزیابی صورت می گیرد تا پیچیدگی های مسئله کاهش یابند. در مرحله بعد ارزیابی های عمیق تری بر پایگاه های داده محلی مدیریت شده توسط مدیریت های

متفاوت تحت مسئولیت آن‌ها انجام می‌شود و این کار با استفاده از تکنیک‌ها و ابزارهای متداول انجام می‌شود. درنهایت، تبادل داده میان آژانس‌هایی با مرکزیت مشخص، طرح‌ریزی و هماهنگی می‌شود.

روش‌های کامل در فراهم کردن یک چارچوب جامع برای هدایت برنامه‌های بزرگ کیفیت داده در سازمان‌هایی که داده‌های کلیدی را پردازش می‌کنند، مؤثر هستند. از سوی دیگر، آن‌ها تعامل کلاسیکی میان قابلیت اجرایی روش و عدم هماهنگی‌سازی آن در حوزه‌های مشخص یا زمینه‌های تکنولوژیک را نشان می‌دهند. روش‌های کامل فقط به‌طور حاشیه‌ای تحت تأثیر نوع، منابع و جریان داده قرار می‌گیرند که بخشی از سیستم‌های مدرن هستند. مثلاً مدل مدیریت کیفیت داده IP-MAP به‌منظور مدیریت رشد پیچیدگی سیستم‌ها برحسب فرآیندها و عملگرها تا IP-UML رشد یافته است. هرچند نقش و استفاده آن در چارچوب کلی روش مدیریت کیفیت داده به شکل زیادی تغییر پیدا نکرده است.

تکنیک‌های اقتصادی سایر روش‌ها را تکمیل می‌کنند و می‌توانند به آسانی در چارچوب کلی ارائه‌شده توسط هر روش کاملی جای گیرد. اکثر روش‌های ممیزی و بهبود دارای هزینه‌ی ارزیابی هستند (جدول ۱۱ را ببینید). اگرچه آن‌ها عمدتاً بر هزینه آغازین کیفیت داده تمرکز می‌کنند و این در حالی است که تحلیل کامل هزینه-منفعت باید هزینه‌ی انجام ندادن کار را نیز در نظر بگیرد. هزینه‌ی کیفیت پایین داده‌ها معمولاً به خاطر طبیعت سازمانی است. روش‌های اقتصادی بر هر دو جنبه تمرکز دارند. مثلاً روش ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها بر ارزیابی هزینه‌های مرتبط با کیفیت پایین داده‌ها تمرکز دارد و اثر اقتصادی مبتنی بر اثرات نهایی داده‌های بد را ساختاربندی می‌کند. نتیجه‌ی آن با نام کارت امتیاز کیفیت داده شناخته می‌شود و می‌تواند به‌عنوان ابزاری برای یافتن بهترین راه‌حل برای بهبود استفاده شود. در روش جامع برای مدیریت کیفیت داده ارزیابی کلی هزینه‌های کیفیت پایین توسعه‌ی بیشتری یافته است. در این روش به این واقعیت توجه نشده است که همان بهبود کیفیت می‌تواند با اولویت‌ها و مسیرهای متفاوت نیز به دست آید و حداقل سازی هزینه‌ی برنامه‌ی کیفیت داده، معیار اصلی برای انتخاب فرآیندهای بهبود جایگزین است. این روش بیان می‌کند که فرآیندهای بهبود متفاوتی می‌توانند به‌عنوان تکنیک‌های داده محور و فرآیند محور، بر پایه‌ی داده‌ها و جریان داده‌های مستند اعمال شوند و سپس روند بهبود شناسایی شود.

در ادامه، تائید روش‌ها در حوزه اجرا و مهم‌ترین تجربیات گزارش شده در تحقیقات پیشین را شرح می‌دهیم. به علت پراکندگی تجربیات گزارش شده، این قسمت به‌طور مورد به مورد تشریح می‌شود. درواقع، تائید تجربی روش‌ها غالباً مورد غفلت واقع شده یا بر اساس مطالعه‌ی موردی کوچکی است و به‌ندرت با مقیاس بزرگ آزمایش‌های علمی، سنجیده شده است. چراکه این پیاده‌سازی‌ها معمولاً در حوزه صنایع اتفاق می‌افتد و بخشی از اهداف آن برای مشاوره گرفتن در آن حوزه است و نحوه تست فرآیند منتشر نشده باقی می‌ماند.

تجربه استفاده از نسخه‌های اولیه روش مدیریت کیفیت داده در اسناد وزارت دفاع ایالات متحده گزارش شده است (US Department of Defense 1994). هم‌اکنون استفاده از ابزارهای توسعه‌یافته‌ی کیفیت داده در بستر رویکردهای برنامه‌نویسی و SQL برای بررسی کیفیت داده‌ها مورد توجه است. اخیراً از این روش‌ها پایه‌ای در حوزه‌های قانونی نیز استفاده شده است (Sessions 2007). سایر کاربردهای روش مدیریت کیفیت داده در تسهیلات درمانی ارتش و حوزه‌ی درمان وزارت دفاع ایالات متحده گزارش شده است (Wang 1998) (Corey et al. 1996). یک مزیت روش مدیریت کیفیت داده این است که می‌تواند بر اساس مسائل حیاتی کیفیت داده و داده‌های متناظر با آن‌ها به‌صورت کارآمد، نمایندگی و جامع بودن معیارهای کیفیت داده و اینکه آیا مجموعه‌ی درستی از معیارها ارزیابی شده است یا خیر را مشخص کند. به همین دلیل از این روش در حوزه‌های مختلف نظیر شرکت‌های بیمه استفاده شده است (Nadkarni 2006) (Bettschen 2005). امروزه با بهبود IP-MAP (Scannapieco et al. 2005; Shankaranarayanan and Wang 2007) یا با ارائه مدل کامل مبتنی بر مدیریت کیفیت داده (Baskarada et al. 2006) تلاش‌های بیشتری برای توسعه‌ی مدیریت کیفیت داده انجام شده است. روش مدیریت کیفیت اطلاعات کلی یک روش حرفه‌ای است و کاربردهای بسیار متنوعی از این روش در زندگی واقعی و مطالعات موردی زیادی در مقاله (English 1999) آمده است که از مدیریت رابطه با مشتری تا بازاریابی از راه دور و حوزه سلامت را در برمی‌گیرد. تعداد زیادی از ابزارهای کیفیت داده نیز مورد مقایسه قرار گرفته‌اند. قواعد مدیریتی و داستان‌های موفق بی‌شماری در وبسایت www.infoimpact.com نیز

آمده است که مثال‌هایی از تجارب حرفه‌ای مرتبط با حوزه‌ی استفاده از کیفیت داده برای کاهش هزینه، بهبود ارزش اطلاعاتی و اثربخشی تجاری ارائه می‌کند.

روش موسسه کانادایی برای روش اطلاعات سلامت سال‌های زیادی برای سازمان‌دهی، نگهداشت و بهبود در ۲۰ پایگاه داده‌ی مدیریت درمانی اعمال شده است که بسیاری از آن‌ها مبتنی بر جمعیت هستند (Long and Seko 2005). شواهدی در سراسر کانادا وجود دارد که اثر روش موسسه کانادایی برای روش اطلاعات سلامت در حمایت از مدیریت کارآمد حوزه سلامت و توسعه سیاست‌های عمومی را بازتاب می‌دهد (Chapman et al. 2006). روش موسسه کانادایی برای روش اطلاعات سلامت همچنین به عنوان مدل مرجع در سایر کشورها نیز استفاده می‌شود؛ مثلاً این روش در کنار روش مدیریت کیفیت داده برای استقرار یک استراتژی کیفیت داده برای وزارت بهداشت در نیوزیلند استفاده شده است (Kerr and Norris 2004). پس از اجرای بلندمدت این روش، مؤلفان داستان‌های موفق‌ی ارائه کردند که نشان می‌دهد فرآیند ارزیابی در زمان استفاده از روش موسسه کانادایی برای روش اطلاعات سلامت برای دستیابی به هدف اولیه شناسایی و رتبه‌بندی مهم‌ترین مسائل بهبود کیفیت داده، موفق عمل کرده است. روش موسسه کانادایی برای روش اطلاعات سلامت نه تنها نواحی مناسب را برجسته می‌کند، بلکه نواحی مشکل‌دار در پایگاه داده را نیز هدف‌گیری می‌کند و کنترل وضعیت کیفیت داده برای داده‌های بیشمار را نیز تسهیل می‌کند. همچنین ارزیابی امتیازات می‌تواند برای توصیف کیفیت داده‌ها در یک موسسه استفاده شود. به نظر می‌رسد که طرح-ریزی استراتژیکی و مشارکتی می‌تواند در زمانی که نتایج ارزیابی خلاصه شوند، تسهیل یابند. اگرچه در این حالت تعداد ارزیابی‌ها اندک است، اما بهبودهای زیادی در پایگاه داده به دست آمده و ممکن است بسیاری از این بهبودها کشف نشده باشند. محدودیت‌های اصلی تجربه‌شده در روش موسسه کانادایی برای روش اطلاعات سلامت این است که خواص اندازه‌گیری فرآیند ارزیابی، هنوز شناخته شده نیست و فقط داده‌های اولیه در زمان مطالعه، در دسترس هستند و پس از اجرای طرح نتایجی در ارتباط با آن داده‌ها ارائه می‌شوند.

مؤلفان روش ارزیابی کیفیت داده‌های مالی ادعا می‌کنند که کار آن‌ها نتیجه‌ی سال‌ها تجربه در حوزه‌ی کیفیت داده‌های مالی است و توسعه‌ی تکنیک‌های تحلیل کیفیت در روش ارزیابی کیفیت داده‌های مالی ساده

است و نیازمند استفاده از پکیج‌های گران‌قیمت نیست. آن‌ها همچنین یک سناریوی واقعی را ارائه کردند که به دلایل امنیتی، تمام مدارک آن را ارائه نکردند.

در کار (Batini and Scannapieco 2006) تجربه بزرگ‌مقیاسی از اجرای روش جامع برای مدیریت کیفیت داده گزارش شده است که به سازمان‌دهی مجدد روابط دولتی تجارت در ایتالیا اشاره دارد. بررسی نحوه‌ی تعامل با دولت برای چندین رویداد تجاری نظیر آغاز یک کسب‌وکار جدید و رشد و نمو آن که شامل نوساناتی در وضعیت قانونی، ترکیب هیئت‌مدیره، مدیریت ارشد و تعداد کارکنان می‌شود، موردنیاز است. آژانس‌ها در تعامل خود با کسب‌وکارها، اطلاعات مشترکی را مدیریت می‌کنند که نوعاً شامل نام رئیس، مقر شرکت، نشانی شعبه‌ها و فعالیت اصلی اقتصادی است. از آنجایی که هر کسب‌وکار به صورت مستقل به هر آژانس گزارش می‌دهد، نسخه‌ها دارای سطوح متفاوتی از دقت هستند. به این منظور روش جامع برای مدیریت کیفیت داده در این حوزه اعمال شده است که موجب نتایج زیر می‌شود:

- اجرای فعالیت‌های مبتنی بر پیوند رکوردها باهدف ایجاد ارتباط میان شناساگرهای کسب‌وکار در مدیریت‌های متفاوت؛
- مهندسی مجدد رابطه‌ی دولت با کسب‌وکار

در سناریوی مهندسی مجدد، کسب‌وکارها فقط با یک مدیریت که مسئول ارسال مقدار یکسانی از اطلاعات به سایر مدیریت‌هاست، در تعامل هستند. با در نظر گرفتن یک دوره‌ی سه‌ساله با محدودیت‌های هزینه‌ای و صرفه‌جویی در ارتباط با کیفیت داده‌ها، منافع به کارگیری روش جامع برای مدیریت کیفیت داده حدود ۶۰۰ میلیون یورو تخمین زده شده است. این امر اعتبار روش را اثبات می‌کند که در آن فعالیت‌های داده محور و فرآیند محور در نظر گرفته شده و برنامه‌ی کیفیت داده‌ها با بهینه‌سازی نسبت هزینه به کیفیت، انتخاب شده است. سایر کاربردهای روش جامع برای مدیریت کیفیت داده در مقاله (Basile et al. 2007) گزارش شده است. مطالعاتی که از روش جامع برای مدیریت کیفیت داده به عنوان روش مرجع برای ارزیابی،

اندازه‌گیری و تقلیل ریسک‌های عملیاتی مرتبط با فرآیند پایه‌ای استفاده کرده‌اند نشان می‌دهند که این روش در برخی موارد منجر به صرفه‌جویی‌های شگرف سالیانه شده است.

۲-۱۰ از کیفیت داده به سمت کیفیت اطلاعات

در این پژوهش به موضوع روش کیفیت داده اشاره شد و سیزده نوع از آن را مقایسه و شرح دادیم. کلیه‌ی تحقیقات کیفیت داده اخیر را بررسی کردیم، هرچند نمی‌توان آن را کامل تصور نمود. روش‌ها برای سنجش و بهبود کیفیت داده‌ها در چندین مسیر در حال تکامل هستند:

(۱) در نظر گرفتن تعداد بیشتری از انواع داده‌ها، حرکت از سمت کیفیت داده به سوی کیفیت اطلاعات،

(۲) ربط دادن بیشتر موضوعات کیفیت داده به موضوعات فرآیند کسب و کار،

(۳) در نظر گرفتن انواع جدید سیستم‌های اطلاعاتی مخصوصاً وب و سیستم‌های اطلاعات P2P.

راجع به این سه حوزه به‌طور جداگانه بحث می‌شود.

تجزیه و تحلیل‌های پیشین بر این حقیقت تأکید دارند که موضوعات کیفیت اطلاعات تنها حوزه‌ی داده‌های نیمه ساختاریافته، داده‌های ساختار نیافته و داده‌های چندرسانه‌ای را به‌طور کلی مورد توجه قرار داده‌اند. در واقع، روش‌ها و تکنیک‌ها از مدت‌ها پیش برای این نوع داده‌ها در بخش‌های تحقیقاتی ایجاد شده‌اند، به‌عنوان نمونه می‌توان درک زبان طبیعی برای اسناد و مدارک در مواردی که غنی‌سازی و بهبود خاصی به کمک تکنیک‌های کیفیت داده روی آنها رخ نداده است.

تحقیقات انجام‌شده‌ی متمرکز بر روی داده‌های نیمه ساختاریافته و ساختار نیافته در حوزه‌ی کیفیت داده حاکی از ارتباط تاریخی محکم بین کیفیت داده و طرح پایگاه داده می‌باشند. حتی روش‌های کامل کیفیت داده دارای سوگیری (بایاس) هستند؛ اما تمرکز اساسی آن‌ها بر مجموعه‌ای از داده‌های ساختاریافته است که ارائه‌دهنده‌ی بیشترین منابع اطلاعاتی بالغ در بیشتر سازمان‌ها هستند (Batini et al. 2009). امروزه حرکت

به داده‌های نیمه ساختاریافته و فاقد ساختار به عنوان منابع سازمانی بسیار رایج تر است. مدیریت دانش، حفاظت وب و سیستم‌های اطلاعات جغرافیایی بیانگر حوزه‌های تحقیقاتی مهمی است. همین‌طور تکنیک‌های کیفیت داده برای داده‌های نیمه ساختاریافته و فاقد ساختار اخیراً مورد تحقیق و بررسی قرار گرفته‌اند (Batini and Scannapieco 2006). بهبود تکنیک‌های کیفیت داده برای داده‌های فاقد ساختار و نیمه ساختاریافته در این حوزه‌ها (فیلدها) نیازمند درجه بالاتری از ارتباط بین‌رشته‌ای است که البته ممکن است با تأخیر بسیار زیادی انجام گیرد.

۲-۱۱ کیفیت داده‌ها و کیفیت فرآیند

ارتباط بین کیفیت داده و کیفیت فرآیند به دلیل ارتباط و تنوع ویژگی‌های فرآیندهای کسب‌وکار در سازمان‌ها، بخش عمده‌ای از تحقیقات را دربر می‌گیرد. تأثیرات متفاوت کیفیت داده‌ها در سه سطح مشخص به نام‌های عملیاتی، تاکتیکی و سطوح استراتژیک در پژوهش (Redman 1998) مورد تحقیق و بررسی قرار گرفته است. در این کتاب مصاحبه‌ها و خروجی‌های چندین مطالعه‌ی خاص گزارش شده است. کیفیت داده و ارتباط آن با کیفیت سرویس‌ها، محصولات، عملیات کسب‌وکار و رفتار مصرف‌کننده در قالب اصطلاحات عمومی زیادی در (Sheng and Mykytyn 2002) و (Sheng 2003) مورد بررسی قرار گرفته‌اند. در این بررسی‌ها گزاره‌های عام مانند «کیفیت اطلاعات یک شرکت به صورت مثبت با عملکرد شرکت رابطه دارد» مبتنی بر شواهد تجربی استفاده شدند. همچنین این مسئله که چگونه بهبود فرآیندهای تولید اطلاعات به صورت مثبت بر داده‌ها و کیفیت اطلاعات تأثیر می‌گذارد نیز مورد تجزیه و تحلیل قرار گرفته است (English 2002).

تعدادی از مقالات به موضوعات خاصی اشاره دارند و در نتیجه نتایج محکم‌تری را نسبت به موضوعاتی ارائه می‌دهند. (Vermeer 2000) موضوع تبادل الکترونیکی داده (EDI) را بررسی می‌کند که مرتبط با تبادل داده‌ها در میان سازمان‌هایی است که از فرمت‌های استاندارد استفاده می‌کنند و چگونگی تأثیر آن بر بازدهی و

اثربخشی فرآیندهای کسب و کار و کیفیت داده در نظر گرفته شده است. EDI به صورت عمومی تر، زبان‌های نشانه گذاری به عنوان یک توانمندسازی تکنولوژی استفاده می کند، زیرا آن‌ها به طور بالقوه فعالیت‌های جابجایی کاغذ، خطا وارد کردن داده‌ها و وظایف وارد کردن داده‌ها را کاهش می دهند.

تأثیر کیفیت داده‌ها بر روی اثربخشی فعالیت‌های تصمیم گیری در پژوهش (Raghunathan 1999) با تمرکز بر روی بعد دقت بررسی شده است. این تجزیه و تحلیل نشان می دهد که اثربخشی تصمیمات با کیفیت داده‌ی بهتر فقط زمانی که تصمیم گیرنده در مورد ارتباط میان متغیرهای مسئله دانش داشته باشد، بهبود می یابد و در غیر این صورت در جهت مخالف تنزل خواهد کرد.

تأثیر کیفیت داده‌ها در شرایط بحرانی مانند بلایا و مواقع اضطراری در پژوهش (Fisher and Kingma 2001) مورد بحث قرار گرفته است. نقص‌ها در دقت، جامعیت، سازگاری و به هنگام بودن با رجوع به موقعیت‌های بحرانی مورد توجه قرار گرفته اند. مثلاً زمانی که ناوگان دریایی Vincennes آمریکا به ایرباس ایرانی شلیک کرد و باعث کشته شدن ۲۹۰ نفر شد.

نقش اطلاعات در زنجیره تأمین در کتاب (Dedeke 2005) مورد بحث قرار گرفته است. در این بررسی توان کیفی یک زنجیره‌ی اطلاعات برای اندازه گیری قابلیت فرآیند تولید و همچنین برای ساخت محصول نهایی در مواردی که تهدیدات باعث به هم ریختن اطلاعات، تغییر در حین تبدیل و نقص اطلاعات شده پیشنهاد می شود. در این بررسی یک چارچوب به نام طراحی توان کیفیت فرآیند برای تشخیص، تجویز و ساخت کیفیت در میان زنجیره‌های اطلاعاتی پیشنهاد شده است.

۲-۱۲ کیفیت داده‌ها و انواع جدید سیستم‌های اطلاعاتی

با سیر تکاملی تکنولوژی، ماهیت سیستم‌های اطلاعاتی هم تغییر کرده و انواع جدیدی از سیستم‌های اطلاعاتی در حال ظهور هستند. در این پژوهش موضوعاتی را تجزیه و تحلیل می کنیم که مربوط به وب و سیستم‌های اطلاعاتی P2P هستند.

نگرانی در مورد سیستم‌های اطلاعاتی باعث شده که روش‌شناسان چند مشکل عمده را مطرح نمایند:

۱- کیفیت داده‌های فاقد ساختار به‌ویژه اسناد و مدارک رو به افزایش است؛

۲- انواع جدیدی از ابعاد کیفیت مانند در دسترس بودن در حال تغییر است.

(Pernici and Scannapieco 2003) مدلی را پیشنهاد می‌دهند که بازدهی اطلاعات را با پیوند داده‌های وب، ذکر نام هر آیتم در یک صفحه وب و ارتباط آن با صفحات و با گروهی از صفحات بالا می‌برد. این مدل درون روش طراحی کیفیت داده و مدیریت سیستم‌های اطلاعات وب به کار گرفته می‌شود. نویسندگان در مورد چگونگی غنی کردن روش‌ها برای طراحی سیستم اطلاعات وب بحث می‌کنند. مثلاً (Mecca et al. 1998) و (Isakowitz et al. 1995) با گام‌های اضافی مخصوص طراحی کیفیت داده را تشریح کرده‌اند و در این بررسی‌ها، چند بعد از جمله ناپایداری، کامل بودن، دقت محتوایی و معنایی مورد توجه قرار گرفته‌اند.

ارتباط کیفیت مستندات وب رو به افزایش است، زیرا تعداد مستنداتی که در فرمت وب مدیریت می‌شوند به‌طور پیوسته در حال افزایش هستند. مثلاً مقاله‌ی (Rao 2003) نشان داده که ۴۰٪ مطالب روی شبکه در طول یک سال از بین می‌روند، ۴۰٪ تغییر می‌یابند و تنها ۲۰٪ به شکل اولی باقی می‌مانند. مطالعات دیگر مانند (Lyman and Variam 2003) نشان داده که متوسط طول عمر یک صفحه وب ۴۴ روز است و هر سال چهار بار، تغییر کامل وب رخ می‌دهد، در نتیجه حفظ داده‌های وب مهم و حیاتی است. اصطلاح حفظ وب نشان‌دهنده توانایی جلوگیری از دست رفتن اطلاعات در وب با ذخیره‌سازی همه ورژن‌های مهم مستندات وب است.

مقاله (Cappiello et al 2003) یک روش برای پشتیبانی از حفظ فرایند در طول کل چرخه عمر اطلاعات (ایجاد، کسب، ذخیره و دسترسی) ارائه می‌دهد. فازهای اصلی این روش عبارت‌اند از:

(۱) هر وقت یک صفحه جدید منتشر می‌شود داده‌ها با فراداده‌هایی که توصیف‌کننده کیفیت آن‌ها بر حسب دقت، کامل بودن، سازگاری، جدید بودن و ناپایداری است مرتبط شوند. همان‌طور که در بخش ۲,۳ تعریف شد.

(۲) در فاز اکتساب، کاربر مقادیر قابل قبولی برای همه ابعاد کیفیت را مشخص می‌کند. اگر داده‌های جدید الزامات کیفیت را برآورده کنند، داخل یک آرشیو جای داده می‌شوند و در غیر این صورت داده‌ها به صاحبانشان بازگردانده می‌شوند.

(۳) در مرحله انتشار وقتی یک صفحه جدید جایگزین یک صفحه قدیمی می‌شود، ناپایداری داده قدیمی ارزیابی می‌شود، اگر داده قدیمی معتبر باشند حذف نمی‌شوند و با یک صفحه URL جدید پیوند می‌خورند.

روش‌های ارزیابی برای کیفیت‌های خاص سایت‌های وب در مطالعات (Atzeni et al 2001), (Fraterrali et al 2005), (Mecca et al 1999) بررسی شده‌اند و به‌طور خاص روی میزان دسترس‌پذیری آن‌ها، طور که در بخش ۲,۳ تعریف شده تمرکز دارند. نهایتاً این معیارها بر اساس روشی ترکیبی کمی و کیفی ارزیابی می‌شوند. در ارزیابی کمی، خطوط راهنمای ارائه‌شده توسط سازمان جهانی وب را چک می‌کنند. ارزیابی کیفی بر اساس آزمایش‌های انجام‌شده با کاربران ناتوان است.

سیستم‌های P2P کاملاً باز هستند و اهمیت ارزیابی و فیلتر داده‌ها در آن‌ها مطرح می‌شود. ابعاد کیفیت مرتبط با اعتماد در سیستم‌های peer-to-peer (دوطرفه) هنوز موضوعاتی پایان نیافته در بررسی‌ها هستند. یک راه حل ممکن برای حل آن تأکید بر شهرت تا مورد اعتماد بودن هر طرف است. مثلاً سرویس رتبه‌بندی ارزیابی کیفیت در سیستم‌های اطلاعات تعاونی (شکل ۴)، مقادیر اعتماد را با هر منبع داده در سیستم اطلاعات پیوند می‌زند.

سرویس رتبه‌بندی متمرکز بوده و فرضش بر این است که توسط یک سازمان ثالث اجرا شود. معیار اعتماد برای تعیین میزان قابلیت اعتماد ارزیابی‌های کیفی انجام‌شده توسط سازمان‌ها به کار می‌روند.

تمرکز معیارها بر اساس مطالعه (De Santis et al 2003) روی شکایت کاربران از مجموعه داده‌ها که توسط آن‌ها اعلام شده می‌باشد.

معیارهای پیشنهادشده در مطالعه‌ی (Gackowski 2006)، منابع داده‌ای دقیقی که داده‌ها را با منابع باکیفیت بالای دقیق مبادله می‌کنند، در نظر می‌گیرند. این معیارهای همچنین شامل یک مؤلفه موقتی هستند که دقت گذشته‌ی منابع داده را به حساب می‌آورد. سایر پیشنهادها شامل بررسی تکنیک‌هایی است که می‌توانند به معتبر بودن مقادیر کمک کند (Gackowski 2006).

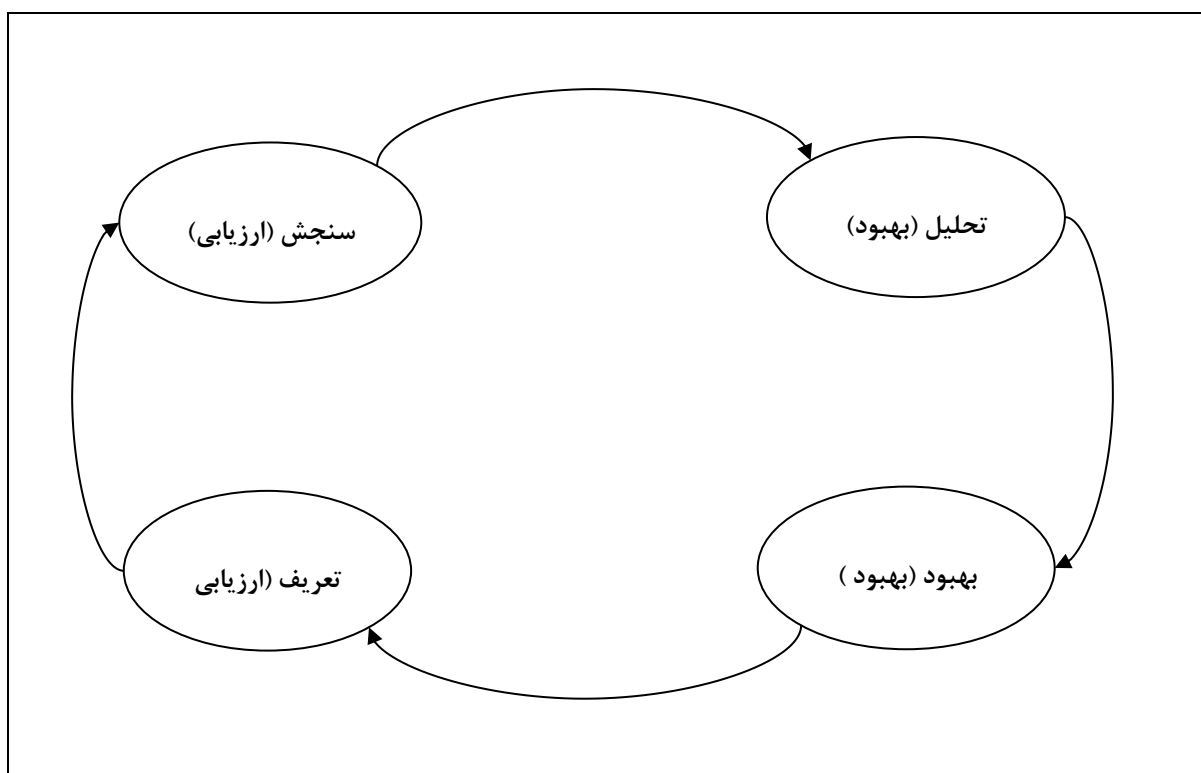
۱۳-۲ خلاصه‌ای از روش‌ها

در این پیوست توصیفی از هر روش ارائه می‌شود. کارت‌های توصیف ساختار مشترکی دارند که مشتمل بر سه بخش هستند:

- ۱) فازهای هر روش و وابستگی متقابل آن‌ها و تصمیم‌های مهم که با دیاگرام و توصیف متنی بیان شده است و شامل ورودی‌ها، خروجی‌ها و گام‌ها است
- ۲) توصیف کلی و تأکید روی تمرکز هر روش و مشارکت اصلی هر کدام در ارزیابی کیفیت داده‌ها و فرآیند بهبود
- ۳) توضیحات تفصیلی که در مورد میزان کاربردی بودن هر روش

۲-۱۳-۱- روش TDQM (مدیریت جامع کیفیت داده)

توضیحات کلی: روش مدیریت کیفیت داده اولین روش عمومی بود که در ادبیات موضوعی کیفیت داده منتشر شد (Wang 1998). این روش در نتایج به دست آمده از پژوهش‌های دانشگاهی گرفته شده است. هدف اصلی روش مدیریت کیفیت داده گسترش کیفیت داده و اصول مدیریت جامع کیفیت است (Oaklan 1989).



فاز ۱: تعریف

ورودی ← محصول اطلاعاتی

تحلیل داده : تعریف ویژگی‌ها در دو سطح : ۱) سطح بالا. توصیف عملکردها برای مصرف کنندگان اطلاعات .

۲) توصیف سطح پایین از واحدها و مؤلفه‌های پایه‌ای محصول اطلاعاتی و روابط آنها .

تحلیل الزامات کیفیت داده : تعریف الزامات IQ از دیدگاه عرضه کنندگان IP، تولید کنندگان و مصرف کنندگان

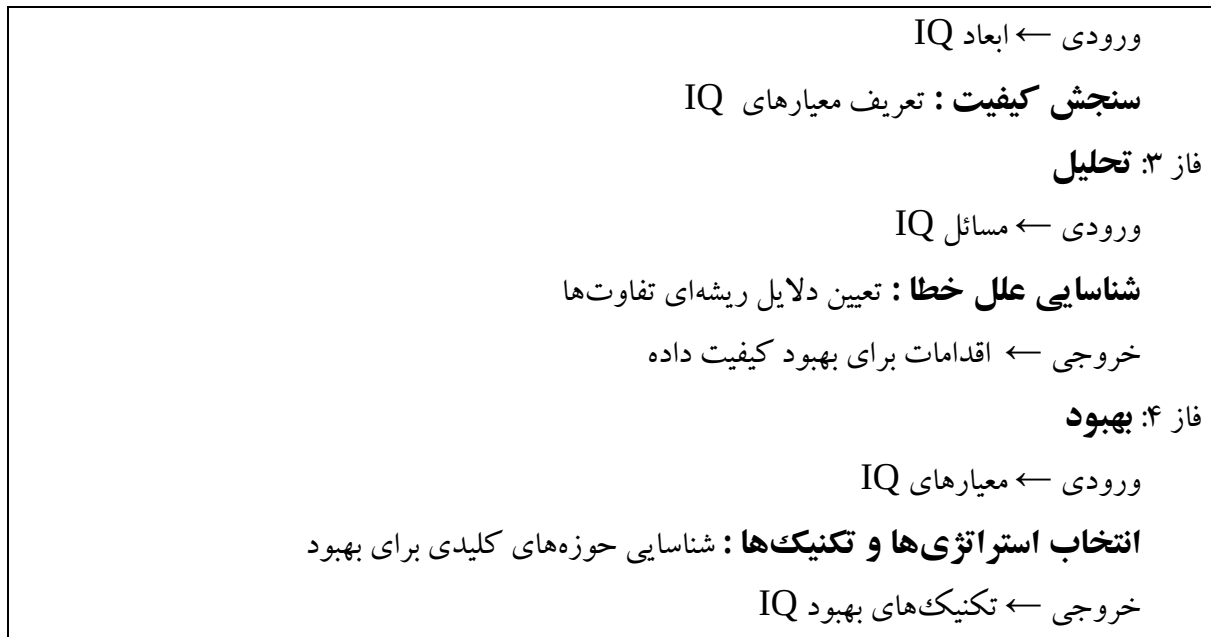
مدیریت مدل کردن فرایند : تعریف سیستم تولید اطلاعات.

خروجی ← طراحی فیزیکی و منطقی محصول اطلاعاتی با ویژگی‌های کیفی ضروری.

یک مدل کیفیت رابطه - موجودیتی که الزامات IP و IQ آن را تعریف می‌کنند.

یک سیستم مدیریت اطلاعات که توصیف می‌کند چگونه IP تولید شده است.

فاز ۲: سنجش



شکل ۲-۶. فازهای مدیریت کیفیت داده

روش مدیریت جامع کیفیت باعث جهت‌گیری تمرکز فعالیت‌های مهندسی مجدد از بازدهی به سمت اثربخشی شده است و این کار از طریق ارائه‌ی خطوط راهنمای روش که هدفشان حذف اختلاف‌ها و تفاوت‌های بین خروجی فرایندهای عملیاتی و الزامات مشتریان است، انجام می‌شود. با توجه به این الزامات، مهندسی مجدد باید از مدل کردن فرایندهای عملیاتی شروع شود. همچنین روش مدیریت جامع کیفیت زبانی برای توصیف فرایندهای تولید اطلاعات یا IP که سازگار با اصول قبل نیز هستند، پیشنهاد می‌دهد که IP-MAP نامیده می‌شود (Shankaranarayan et al. 2000).

زبان IP-MAP به صورت متنوعی به سمت UML و برای پشتیبانی از طراحی سازمانی گسترش یافته است. IP-MAP تنها زبان برای مدل کردن فرایند اطلاعات است و نشان‌دهنده‌ی یک استاندارد عملی است. به عنوان مثال تجربیات عملی در ارتباط با روش مدیریت کیفیت داده در مطالعه Kovac and Weickert (2002) گزارش شده است.

توضیحات تفصیلی: هدف روش مدیریت کیفیت داده پشتیبانی کامل از فرایند بهبود کیفیت (از تحلیل الزامات تا اجرا) است. همان‌طور که در شکل ۶ می‌بینیم، چرخه روش مدیریت کیفیت داده شامل چهار فاز است که یک فرایند پیوسته بهبود کیفیت را اعمال می‌کند و شامل تعریف، سنجش، تحلیل و بهبود است. نقش‌های مسئول برای فازهای مختلف فرایند کیفیت همچنین در روش مدیریت کیفیت داده تعریف می‌شوند. چهار نقش متمایز وجود دارند:

۱. تأمین‌کنندگان اطلاعات که داده‌ها را برای IP ایجاد یا جمع‌آوری می‌کنند،
۲. تولیدکنندگان اطلاعات که داده و زیرساخت سیستم‌های مربوطه را طراحی، توسعه و حفظ می‌کنند،
۳. مصرف‌کنندگان اطلاعات که از داده‌ها در کارشان استفاده می‌کنند،
۴. مدیران پردازش اطلاعات که مسئول مدیریت کل فرایند تولید اطلاعات در چرخه عمر اطلاعات هستند.

همچنین روش مدیریت کیفیت داده از دیدگاه اجرا هم جامع است زیرا خطوط راهنمایی در مورد چگونگی به‌کارگیری روش ارائه می‌دهد. در به‌کارگیری روش مدیریت کیفیت داده، یک سازمان باید:

- (۱) به‌روشنی IP را درک کند.
- (۲) یک تیم IP ایجاد کند که شامل یک مدیر ارشد به‌عنوان رهبر روش مدیریت کیفیت داده باشد، یک مهندس IP که با روش مدیریت کیفیت داده آشنا است و اعضای که تأمین‌کننده، تولیدکننده، مصرف‌کننده‌ی اطلاعات و مدیر IP هستند.
- (۳) ارزیابی و مدیریت IQ را به همه زیرمجموعه‌های تشکیل‌دهنده IP آموزش دهد.
- (۴) بهبود مستمر IP را نهادینه کند.

روش مدیریت کیفیت داده، به ادبیات موضوعی کیفیت اطلاعات برای معیارها و تکنیک‌های بهبود IQ متکی است و به‌صورت صریح به آن‌ها اشاره می‌کند (wang and Strong 1996) و برای مشخصه‌های ابعاد کیفیت روندی مشخص دارد. روش مدیریت کیفیت داده موضوعات کیفیت را به تکنیک‌های متناظر بهبود

مرتبط می‌سازد. اما در ادبیات اخیر به هیچ‌گونه تکنیک خاصی در صنعت اشاره نشده است و هیچ‌گونه پشتیبانی برای تخصصی کردن تکنیک‌های عمومی بهبود کیفیت ارائه نشده است.

۲-۱۳-۲- روش DWQ (کیفیت انبار داده)

توضیحات کلی: روش روش‌شناسی کیفیت انبار داده در پروژه‌ی کیفیت انبار داده‌های اروپا توسعه داده شده است (Jeusfeld et al. 1998). این روش رابطه بین اهداف کیفیت و گزینه‌های طراحی را در انبار داده مطالعه می‌کند. این روش ذهنی بودن مفهوم کیفیت را در نظر می‌گیرد و یک دسته‌بندی از اهداف کیفیت بر طبق گروه ذینفعان ارائه می‌دهد و بر آن‌ها متمرکز می‌شود. سپس تنوع اهداف کیفیت را در نظر می‌گیرد و فراداده‌های متناظر را تعریف می‌کند.

توضیحات تفصیلی: روش روش‌شناسی کیفیت انبار داده بیان می‌کند که در طراحی فراداده‌ها در انبار داده باید سه دیدگاه را لحاظ کنند:

۱) دیدگاه مفهومی کسب‌وکار که روی مدل سازمان تمرکز دارد.

۲) دیدگاه منطقی که روی طرح انبار داده تمرکز دارد.

۳) یک دیدگاه فیزیکی که بیانگر لایه انتقال داده فیزیکی است.

این سه دیدگاه متناظر با سه لایه سنتی انبار داده به نام‌های منابع، انبار داده و مشتریان هستند. این روش از طریق یک دیدگاه فراداده‌ای متناظر به نام سنجش کیفیت، با هر دیدگاه ارتباط پیدا می‌کند.

از دیدگاه کیفیت داده، چهار فاز اصلی مشخصه‌های روش هستند: تعریف، ارزیابی، تحلیل و بهبود (شکل ۷).

در نظرگیری طبقه‌بندی داده‌ها و ابعاد کیفیت نرم‌افزار در بافت انبار داده‌ها، یکی از بخش‌های عمده‌ی این روش است. سه دسته از داده‌ها و فراداده‌ها به صورت زیر تعریف می‌شوند:

- کیفیت طراحی و مدیریت: اولی اشاره به توانایی یک مدل در ارائه‌ی کافی و مؤثر اطلاعات دارد، درحالی‌که دومی به روشی اشاره دارد که مدلی را در طول عملیات انبار داده تکمیل می‌کند.

- کیفیت اجرای نرم افزار: در این روش ابعاد کیفیت استاندارد ISO 9126 در نظر گرفته می شوند، زیرا اجرای نرم افزار یک وظیفه با مشخصه های خاص انبار داده نیست.
 - کیفیت استفاده از داده: که اشاره به ابعادی دارد که مشخص کننده ی استفاده و جستجوی داده در یک انبار داده هستند.
- روش های سنجش مناسب برای هر بعد طبقه بندی شده شناسایی می شوند. لیست این روش ها همواره با درجه ی ارتباطی که با هر ذینفع ورودی مرحله سنجش دارند، تحت تأثیر قرار می گیرند. در فاز ارزیابی کیفیت، مجموعه ای از اطلاعات در مورد هر بعد کیفیت داده وجود دارد که به شرح زیر است:

۱) الزامات کیفیت، محاسبه ی فاصله از مقادیر مورد انتظار

۲) سنجش کیفیت به دست آمده

۳) معیار استفاده شده برای محاسبه اندازه گیری

۴) وابستگی های علی معلولی به دیگر ابعاد کیفیت.

اطلاعات در مورد وابستگی های بین ابعاد کیفیت برای رد گیری و تحلیل مسائل کیفیت استفاده می شوند. شناسایی حوزه های اصلی، مرحله آخر تحلیل روش است. در واقع این مورد تنها فاز بهبود را ذکر می کند اما شامل دانش سازنده ی پایگاه در مورد چگونگی بهبود کیفیت انبار داده نیست.



شکل ۲-۷. فازهای روش شناسی کیفیت انبار داده

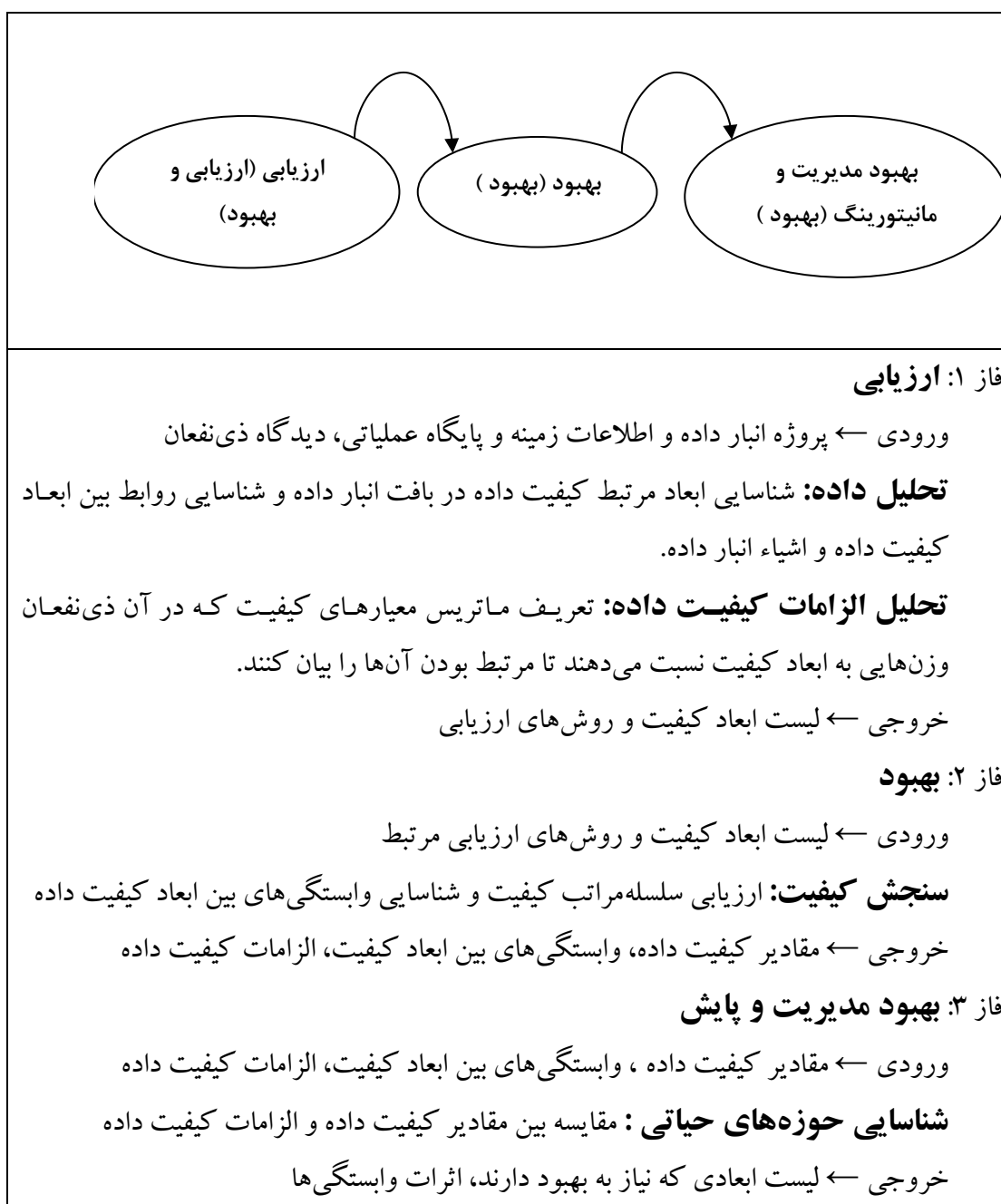
۲-۱۳-۳- روش TIQM (مدیریت جامع کیفیت اطلاعات)

توضیحات کلی: روش مدیریت کیفیت اطلاعات کلی برای پشتیبانی از پروژه‌های انبار داده طراحی شده است. این روش فرض یکپارچه کردن منابع داده عملیاتی به شکل یک پایگاه داده‌ی واحد و یکپارچه را داشته و در انواع یکپارچه‌سازی‌های انجام گرفته برای ایجاد انبار داده استفاده می‌شود. این یکپارچه‌سازی‌ها باعث حذف خطاها و ناهمگونی‌های پایگاه داده‌های منبع می‌شود. روش مدیریت کیفیت اطلاعات کلی روی فعالیت‌های مدیریتی تمرکز دارد که این فعالیت‌ها مسئول یکپارچه‌سازی منابع داده‌ی عملیاتی از طریق بحث در مورد استراتژی‌های سازمان برای انتخاب‌های مؤثر تکنیکی هستند. مثلاً در این روش تحلیل‌های هزینه - فایده از یک دید مدیریتی پشتیبانی می‌شوند. این روش یک دسته-بندی تفصیلی از هزینه‌ها و منافع ارائه می‌دهد (بخش ۳,۴).

توضیحات تفصیلی: شکل ۸ فازهای روش مدیریت کیفیت اطلاعات کلی را نشان می‌دهد. اگر از دیدگاه مدیریتی به روش مدیریت کیفیت اطلاعات کلی بنگریم سه فاز اصلی وجود دارد:

- ارزیابی
- بهبود و مدیریت بهبود
- پایش کردن.

یکی از کمک‌های باارزش این روش تعریف فاز آخر است که خطوط راهنمایی برای مدیریت تغییر در ساختار سازمان مطابق الزامات مدیریت کیفیت داده فراهم می‌کند. به علاوه در رویکرد اقتصادی، ارزیابی هزینه-فایده را برای توجیه مداخلات در کیفیت داده ارائه می‌دهند. هدف این روش دستیابی به سطح بالاتری از کیفیت داده نیست بلکه هدف اصلی به عهده گرفتن اقدامات بهبود در صورت امکان‌پذیر بودن آن است. بنابراین فقط هنگامی که منافع از هزینه‌ها بیشتر شود توجیه دارد.



شکل ۲-۸. فازهای مدیریت کیفیت اطلاعات کلی

۲-۱۳-۴- روش AIMQ (روشی برای ارزیابی کیفیت اطلاعات)

توضیحات کلی: روش ارزیابی کیفیت اطلاعات تنها روش کیفیت اطلاعاتی است که روی مقایسات برای بهینه‌گزینی تمرکز دارد (Lee et al. 2002) و یک تکنیک مستقل از حوزه‌های مختلف برای ارزیابی کیفیت است. پایه‌ی روش ارزیابی کیفیت اطلاعات یک جدول دو در دو به نام مدل PSP/IQ است (شکل ۹) که ابعاد کیفیت را مطابق اهمیتشان از دیدگاه کاربران و مدیران دسته‌بندی می‌کند. محورهای جدول "همخوانی با مشخصه‌ها" و "همخوانی با انتظارات کاربران" هستند. در ادامه روش، چهار دسته ابعاد متمایز شناسایی می‌شوند (محکم، وابسته، مفید و قابل استفاده). سپس ابعاد کیفیت مطابق مقاله‌ی (Wang and Strong 1996) شناسایی شده و در راستای این دسته‌ها طبقه‌بندی می‌شوند. مقایسات باید اطلاعات را درون هر طبقه رتبه‌بندی کند.

مدل PSP/IQ یک ورودی به روش ارزیابی کیفیت اطلاعات است که فازهای آن به‌طور خلاصه در شکل ۱۰ آمده است. مطالب منتشرشده در توصیف روش ارزیابی کیفیت اطلاعات عمدتاً تمرکزشان بر فعالیت‌های ارزیابی است ولی خطوط راهنما، تکنیک‌ها و ابزارهای بهبود ارائه نشده‌اند.

توضیحات تفصیلی: IQ عمدتاً با استفاده از پرسش‌نامه‌ها ارزیابی می‌شود. یک پرسشنامه‌ی اولیه‌ی آزمایشی برای شناسایی ابعاد کیفیت مرتبط و ویژگی‌های مقایسه شده به کار می‌رود. سپس یک پرسش‌نامه‌ی دوم از قبل شناسایی شده را به‌منظور بررسی معیارهای IQ موردتوجه قرار می‌دهند. درنهایت، این معیارها با شاخص‌ها مقایسه می‌شوند. در مقاله‌ی (Lee et al. 2002) فهرستی از ابعاد کیفیت استاندارد و ویژگی‌هایی ارائه شده که به تعریف پرسش‌نامه‌ها کمک می‌کند.

	از انتظار مشتریان فراتر می‌رود یا آن را برآورده می‌کند	همخوانی با مشخصات
کیفیت محصول	اطلاعات مفید	اطلاعات دقیق و درست
کیفیت خدمت	اطلاعات قابل استفاده	اطلاعات وابسته

شکل ۲-۹. مدل PSP/IQ



شکل ۲-۱۰. فازهای ارزیابی کیفیت اطلاعات

۲-۱۳-۵- روش CIHI (موسسه کانادایی اطلاعات سلامت)

توضیحات کلی: روش موسسه کانادایی برای روش اطلاعات سلامت، روشی است که موسسه‌ای کانادایی برای ارزیابی و بهبود کیفیت اطلاعات سلامت به کار گرفته است (Long and Seko 2005). در این روش موضوع اصلی اندازه پایگاه‌های داده و ناهمگونی آن‌هاست. روش موسسه کانادایی برای روش اطلاعات سلامت از انتخاب یک زیرمجموعه از داده‌ها برای تمرکز روی فاز ارزیابی کیفیت پشتیبانی می‌کند. همچنین مجموعه‌ی وسیعی از معیارهای کیفیت را برای ارزیابی ناهمگونی‌ها پیشنهاد می‌دهد.

توضیحات تفصیلی: استراتژی کیفیت داده‌ها در روش موسسه کانادایی برای روش اطلاعات سلامت یک رویکرد دومارحله‌ای پیشنهاد می‌دهد (شکل ۱۱). فاز نخست آن تعریف چارچوب کیفیت داده است و فاز دوم آن یک تحلیل عمیق از داده‌هایی است که بیشترین دسترسی را دارند.

چارچوب کیفیت داده در سه مرحله تعریف می‌شود:

- ۱) استانداردسازی کیفیت اطلاعات داده
 - ۲) توسعه یک استراتژی مشترک برای ارزیابی کیفیت داده
 - ۳) تعریف یک فرایند کاری برای مدیریت داده‌های روش موسسه کانادایی برای روش اطلاعات سلامت که اولویت‌های کیفیت داده را شناسایی می‌کند و رویه‌های بهبود مستمر داده را اجرا می‌کند.
- اجرای چارچوب موسسه کانادایی برای روش اطلاعات سلامت به صورت چرخه‌ای است و طبق رویکرد بهبود مستمر است. موارد زیر برای اجرای موفق لازم هستند:
- تعریف دوره‌ی زمانی برای چرخه
 - تعریف اهداف زمانی برای کیفیت‌های متفاوت مدنظر
 - تخصیص منابع خاص و ویژه برای تحلیل کیفیت داده، ارزیابی و مستندسازی
 - تخصیص منابع ویژه برای بهبود کیفیت داده
- تحلیل داده‌هایی با استفاده از موارد ذکر شده قبلی در سه مرحله انجام می‌شود:
- تحلیل کیفیت داده
 - ارزیابی
 - مستندسازی (در آن مستندات مسائل کیفیت که با ارزیابی و تحلیل شناسایی شده‌اند را گزارش می‌دهند).

ارزیابی کیفیت داده‌ها بر اساس یک مدل سلسله‌مراتبی چهار سطحی است. در سطح اول ۸۶ معیار اولیه کیفیت تعریف می‌شوند. این معیارها توسط الگوریتم‌های ترکیبی به ۲۴ مشخصه‌ی کیفیت در سطح بعدی تبدیل می‌شوند و بعداً دوباره به ۵ بعد درمی‌آیند (در سطح سوم) که شامل دقت، به‌موقع بودن، قابل‌مقایسه بودن، قابل‌استفاده و مرتبط بودن است. درنهایت این پنج بعد به شکل یک ارزیابی پایگاه داده در سطح چهارم درمی‌آیند.

ارزیابی اولیه‌ی ۸۶ معیار کیفیت داده، توسط پرسش‌نامه‌هایی که معیارهای آن با نمره دهی به چهار حالت، عدم به‌کارگیری، نامعلوم، برآورد نشده و برآورد شده هستند، صورت می‌گیرد. سپس نمره در هر سطح جمع شده و ارزیابی اعتبار سنجی صورت می‌گیرد. این فرایند اعتبارسنجی اطمینان می‌دهد که تفسیر و نمره دهی به هر معیار تا حد ممکن استاندارد است.



شکل ۲- ۱۱. فازهای روش موسسه کانادایی برای روش اطلاعات سلامت

۲-۱۳-۶- روش DQA (ارزیابی کیفیت داده‌ها)

توضیحات کلی: روش ارزیابی کیفیت داده (Pipino et al. 2002) برای ارائه اصول کلی راهنمای تعریف معیارهای کیفیت داده طراحی شده است. در ادبیات موضوعی، معیارهای کیفیت داده عمدتاً به صورت ویژه تعریف می‌شوند تا مسائل خاصی را حل کنند؛ بنابراین وابسته به سناریوهای در نظر گرفته شده هستند. روش ارزیابی کیفیت داده هدفش شناسایی اصول سنجش عمومی کیفیت است که در پژوهش‌های پیشین مشترک می‌باشند.

توضیحات تفصیلی: دسته‌بندی معیارهای روش ارزیابی کیفیت داده در شکل ۱۲ خلاصه شده است. معیارهای ذهنی و معیارهای عینی شامل دسته‌بندی وابسته به وظیفه و مستقل از وظیفه هستند. اولین دسته کیفیت داده را بدون دانش زمینه‌ای کاربردی ارزیابی می‌کند درحالی که دومی برای زمینه‌های کاربردی خاص تعریف می‌شود و شامل قواعد کسب و کار، مقررات شرکت‌ها و دولت‌ها و محدودیت‌های ارائه شده توسط مدیریت پایگاه داده است. هر دو معیار به سه دسته‌ی زیر تقسیم‌بندی می‌شوند:

- نسبت ساده
- مقادیر حداقل یا حداکثر
- متوسط وزنی.



شکل ۲-۱۲. فازهای ارزیابی کیفیت داده

۲-۱۳-۲- روش IQM (اندازه‌گیری کیفیت اطلاعات)

توضیحات کلی: هدف اساسی از روش اندازه‌گیری کیفیت اطلاعات ارائه یک چهارچوب جهت کیفیت اطلاعات استفاده‌شده در داده‌های وب می‌باشد. به‌طور خاص، روش اندازه‌گیری کیفیت اطلاعات به انتخاب بر اساس کیفیت و شخصی‌سازی ابزارهای وب در خلاقیت، مدیریت و نگهداری وب‌سایت‌ها می‌پردازد.

توضیحات تفصیلی: روش اندازه‌گیری کیفیت اطلاعات دستورالعملی برای حصول اطمینان که نرم‌افزارهایی که تمامی ابعاد اساسی کیفیت اطلاعات را بررسی می‌کنند، فراهم می‌کند.

این روش مجموعه‌ای از دستورالعمل‌ها ارائه می‌دهد؛ چهارچوب کیفیت اطلاعات و شاخص کیفیت را تعریف کرده و سپس دستورالعمل و چگونگی اندازه‌گیری کیفیت را شرح می‌دهد.

فازهای اصلی روش اندازه‌گیری کیفیت اطلاعات در شکل ۱۳ توضیح داده‌شده‌اند. فاز اول روش اندازه‌گیری معیارها را تعریف می‌کند. سپس چهارچوب کیفیت اطلاعات به‌صورت یک لیست از شاخص‌های مربوط به کیفیت اطلاعات شناخته‌شده می‌شود. محاسبه و تعریف چهارچوب، ورودی محاسبه کیفیت اطلاعات می‌باشد و شاخص کیفیت اطلاعات با روش‌ها و ابزارهایی که در روند اندازه‌گیری مورد استفاده قرار خواهند گرفت، تکمیل می‌شوند.

بعضی شاخص‌ها نیازمند چندین روش اندازه‌گیری می‌باشند و روش‌اندازه‌گیری کیفیت اطلاعات استفاده از روش‌های اندازه‌گیری متعدد را هماهنگ‌سازی می‌نماید.



شکل ۲-۱۳. فازهای اندازه‌گیری کیفیت اطلاعات

۲-۱۳-۸- روش ISTAT (اداره ملی سرشماری ایتالیا)

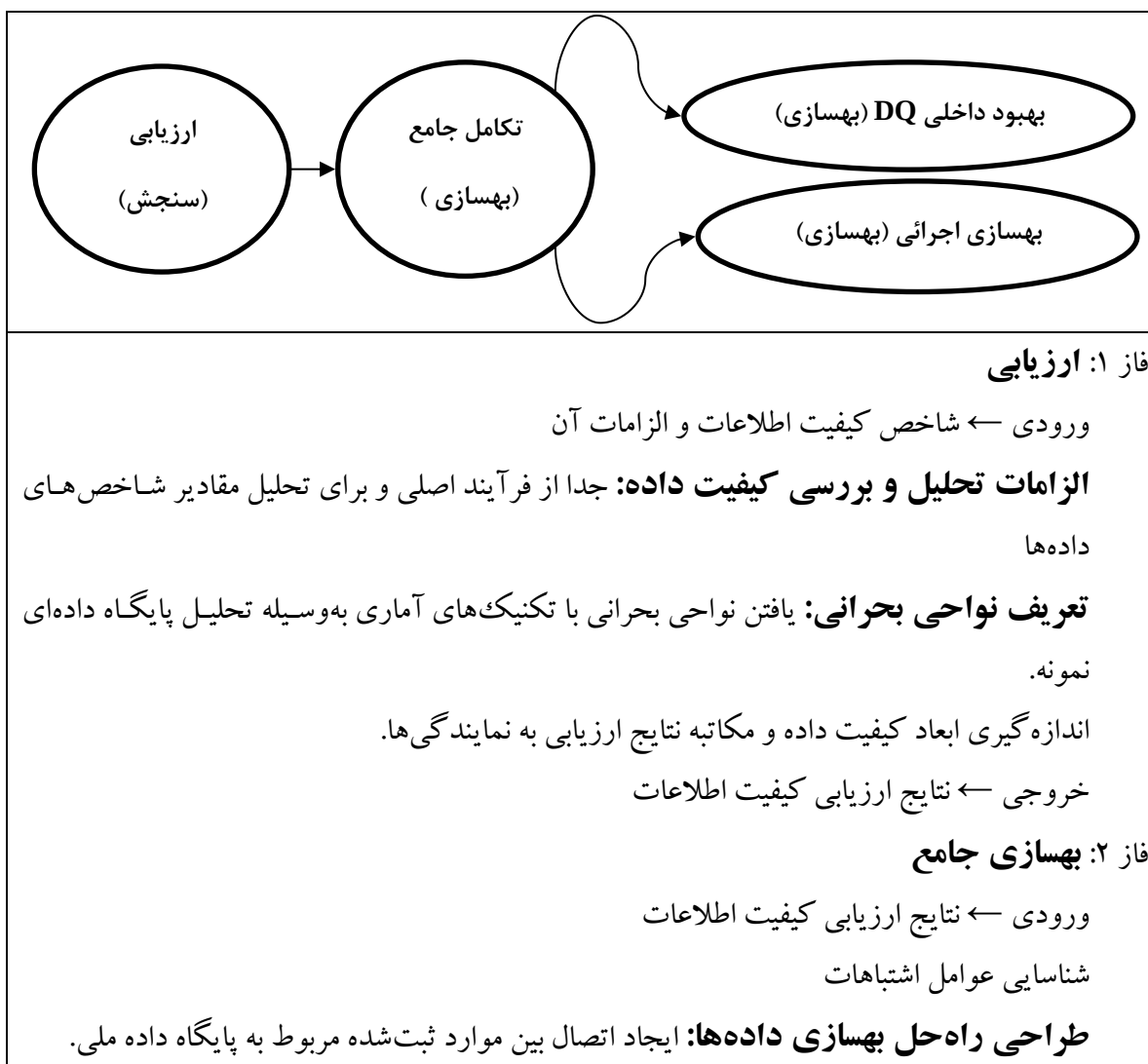
توضیحات کلی: روش ISTAT (Istat 2004; Falorsi et al. 2003) در اداره ملی سرشماری ایتالیا جهت به دست آوردن اطلاعات دقیق و باکیفیت بالا از شهروندان و مشاغل طراحی گردید. مسئله اساسی درباره این روش، چگونگی تضمین به دست آوردن کیفیت داده‌ها از پایگاه داده‌های متعدد عمومی و محلی است. این مسئله یک چالش خاص در مطالعات و پژوهش‌های ایتالیایی است، چراکه مدیریت عمومی به سه ناحیه‌ی جغرافیائی تقسیم شده (مرکزی، منطقه‌ای، و پیرامونی) و هر مدیریت دارای داده‌های مستقل می‌باشد. روش ISTAT بر رایج‌ترین نوع تبادل اطلاعات در میان سطوح مختلف مدیریت عمومی (داده‌های خاص) تمرکز دارد. این روش به شدت بر هنجارها متمرکز بوده و هدف آن مدیریت داده‌ها می‌باشد و در این حالت داده‌ها می‌توانند کیفیت موردنیاز را تأمین نمایند.

توضیحات تفصیلی: مراحل اساسی این روش (شکل ۱۴ را مشاهده نمایید) عبارت‌اند از:

- فاز ارزیابی، که در ابتدا در پایگاه داده مرکزی ISTAT انجام می‌شود، جهت تشخیص مسائل مربوط به کیفیت از دیدگاه یکپارچه‌سازی داده‌ها می‌باشد.
- فاز بهسازی کلی، که مسئول انجام ارتباط رکوردها در پایگاه داده‌های ملی و طرح راه‌حل بهسازی در فرآیندهای ساخت، خرید یا انطباق راه‌حل‌های موجود را دارد.
- فعالیت‌های بهبود در پایگاه داده تحت مالکیت و مدیریت دولت‌های محلی، این فعالیت‌ها باید به وسیله خود دولت‌های محلی با ابزار و دوره‌های ارائه شده توسط ISTAT انجام گردد.
- فعالیت‌های بهسازی که نیازمند همکاری بخش‌های مختلف می‌باشد. این فعالیت‌ها معمولاً پروسه‌های جهت‌دار می‌باشند، چراکه داده‌ها را در طول اجرای فعالیت‌های عملیاتی خاص ردوبدل می‌کنند. پایگاه داده‌ی مرکزی ممکن است جهت اهداف هماهنگی موردنیاز باشد. این فعالیت‌ها به صورت مرکزی برنامه‌ریزی و هماهنگ‌سازی می‌شوند.

روش ISTAT تکنیک‌های ساده ولی تأثیر گذاری را در اندازه‌گیری کیفیت ارائه می‌دهد. همچنین ابزاری برای فعالیت‌های پاک‌سازی داده‌ها در نظر می‌گیرد. در روش ISTAT، مالکین داده‌ها به‌وسیله جزئیات بالا (مثل کد شهروندی و...) تعریف می‌گردند. این روش از استانداردسازی فرمت داده‌ها پشتیبانی می‌کند و بیان آن‌ها با طرح XML این اجازه را می‌دهد که پایگاه‌های داده‌ای محلی باهم ادغام گردند.

ردوبدل شدن اطلاعات بین شرکت‌های محلی به‌وسیله نرم‌افزارها از نو طراحی گشته و بر اساس انتشار و اشتراک عمل می‌کنند.





شکل ۲-۱۴. فازهای روش ISTAT

۲-۱۳-۹- روش AMEQ (فعالیت بر اساس سنجش و تغییر کیفیت اطلاعات محصول)

توضیحات کلی: هدف اصلی روش اندازه گیری مبتنی بر فعالیت و ارزیابی کیفیت اطلاعات محصول ارائه یک ارزیابی دقیق برای کیفیت اطلاعات محصول در بخش ارزیابی و بهسازی (PLQ) به وسیله اهداف منسجم و تعیین شده می باشد (Su and Jin 2004).

این روش مخصوص ارزیابی کیفیت داده ها در شرکت های تولیدی است و سازمان هایی که اطلاعات محصول جز اصلی از پایگاه های عملیاتی داده می باشد. در شرکت های تولیدی رابطه ی بین اطلاعات محصول و فرآیندهای تولید به صورت یک استاندارد ساده و جامع در سراسر شرکت ها مطرح می باشد. همچنین پایگاه داده محصول، معمولاً در شرکت های مختلف مشابه است. این روش یک دستورالعمل شناختی برای مدل های اطلاعات و فرآیندهای تولید مرتبط ارائه می کند.

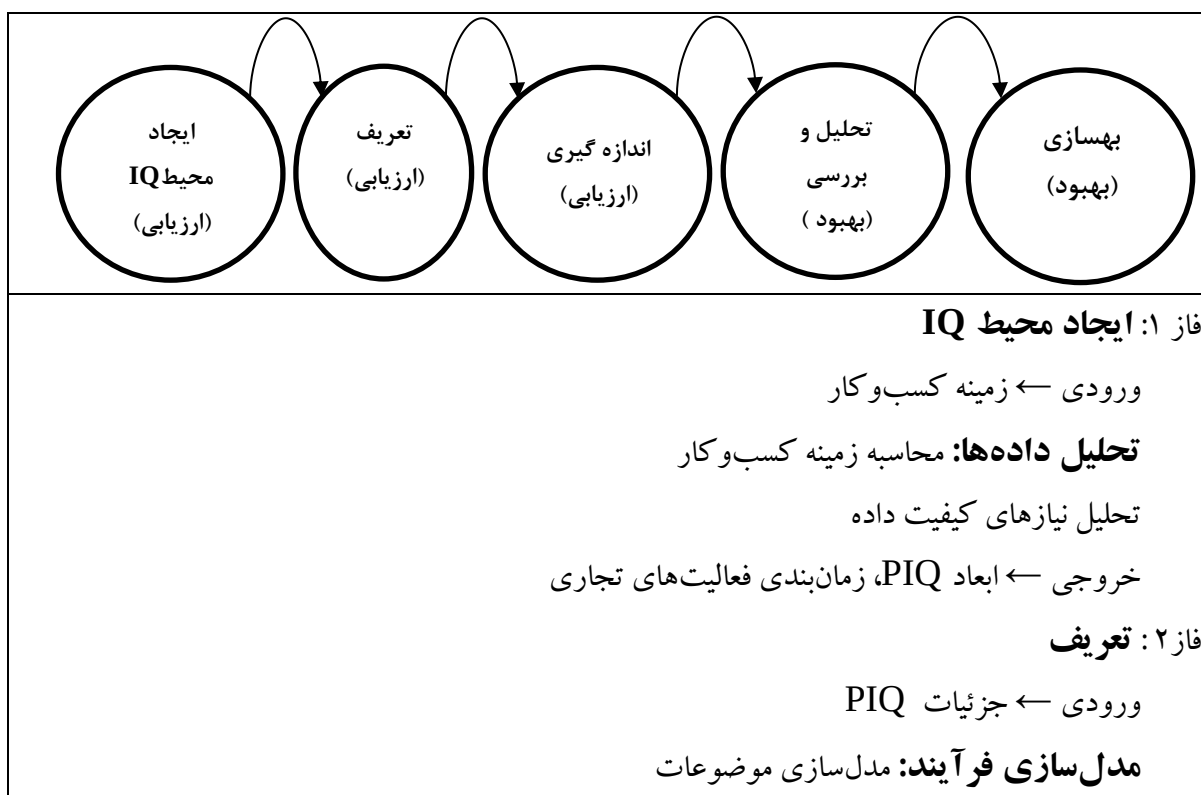
توضیحات تفصیلی: روش اندازه گیری مبتنی بر فعالیت و ارزیابی کیفیت اطلاعات محصول شامل ۵ فاز برای محاسبه و بهسازی کیفیت اطلاعات محصول می باشد. (شکل ۱۵ را ببینید). فاز اول ارزیابی آمادگی فرهنگی یک سازمان می باشد که با استفاده از شبکه مدیریت کیفیت اطلاعات پیاده سازی شده و نقش مهمی در مدیریت دارد. در این فاز ابعاد PLQ بر اساس فعالیت های مختلف کسب و کار تعریف و طبقه بندی می شوند. مرحله دوم، اطلاعات مربوط به محصول را مشخص می سازد. محصول اطلاعات با یک فرآیند کسب و کار مرتبط است و به وسیله تکنیک (ODA) مدل می شود.

در روش اندازه گیری مبتنی بر فعالیت و ارزیابی کیفیت اطلاعات محصول هشت نوع موضوع مدل می شوند که عبارت اند از:

- منابع انسانی
- منابع اطلاعاتی
- فعالیت های شرکت
- ورودی منابع

- منابع فرآیندها
- خروجی منابع
- اندازه گیری عملکرد
- اهداف شرکت.

در این مرحله، یک مدل از روش های اندازه گیری نیز تولید می گردد. مرحله ی سوم بر روی فعالیت های اندازه گیری تمرکز دارد. در مرحله چهارم تیم IQ باید به وسیله تحلیل و بررسی ابعاد کیفیت که نمره پائینتری کسب کرده اند به بررسی علل مشکلات ریشه ای و بالقوه PLQ پردازند. در نهایت، فاز بهبود PLQ می تواند آغاز شود. برای مراحل چهارم و پنجم، روش اندازه گیری مبتنی بر فعالیت و ارزیابی کیفیت اطلاعات محصول ابزارهای دقیقی ارائه نمی کند و تنها دستورالعمل کلی ارائه می دهد.



مشخص نمودن محیط‌های بحرانی: مشخص کردن ابعاد PIQ برای فعالیت ارزش سنجی

خروجی ← نتایج PIQ

فاز ۳: اندازه‌گیری

ورودی ← نتایج و ابعاد PIQ

اندازه‌گیری کیفیت: مشخص کردن بازده اندازه‌گیری PIQ برای هر فعالیت، انتخاب یک پروژه

مدیریت‌پذیر کوچک، تعریف معیارها و سؤالات درباره PIQ و فرآیند اندازه‌گیری

خروجی ← نتایج اندازه‌گیری

فاز ۴: تحلیل و بررسی

ورودی ← نتایج اندازه‌گیری

بررسی عوامل بروز اشکال: بررسی ابعاد کیفیت پائین داده‌ها، تحلیل و بررسی مبانی و محاسبه

بازدهی تاثیرگذار با توجه به اهداف مالی

خروجی ← احتمالات برای ابعاد کیفیت ضعیف

فاز ۵: بهسازی

ورودی ← احتمالات برای ابعاد کیفیت ضعیف

تعیین مسئولیت هر فرآیند داده: استفاده از مدیریت منابع اطلاعات (IRM) برای بهبود بخشی

بازده PLQ در فعالیت‌های تجاری.

طراحی فرآیند و طراحی مجدد: کاربرد مدیریت خود ترتیبی (SOM) در مهندسی مجدد

جریان‌های اطلاعاتی.

کنترل بهسازی: بهسازی مدیریت اهداف مالی (EOM) برای سازگارسازی مجدد ارزش‌های

سازمانی

خروجی ← بهبود طولانی مدت PIQ

شکل ۲-۱۵. فازهای اندازه‌گیری مبتنی بر فعالیت و ارزیابی کیفیت اطلاعات محصول

۲-۱۳-۱۰- روش COLDQ (تأثیر کیفیت پائین داده بر هزینه‌ها)

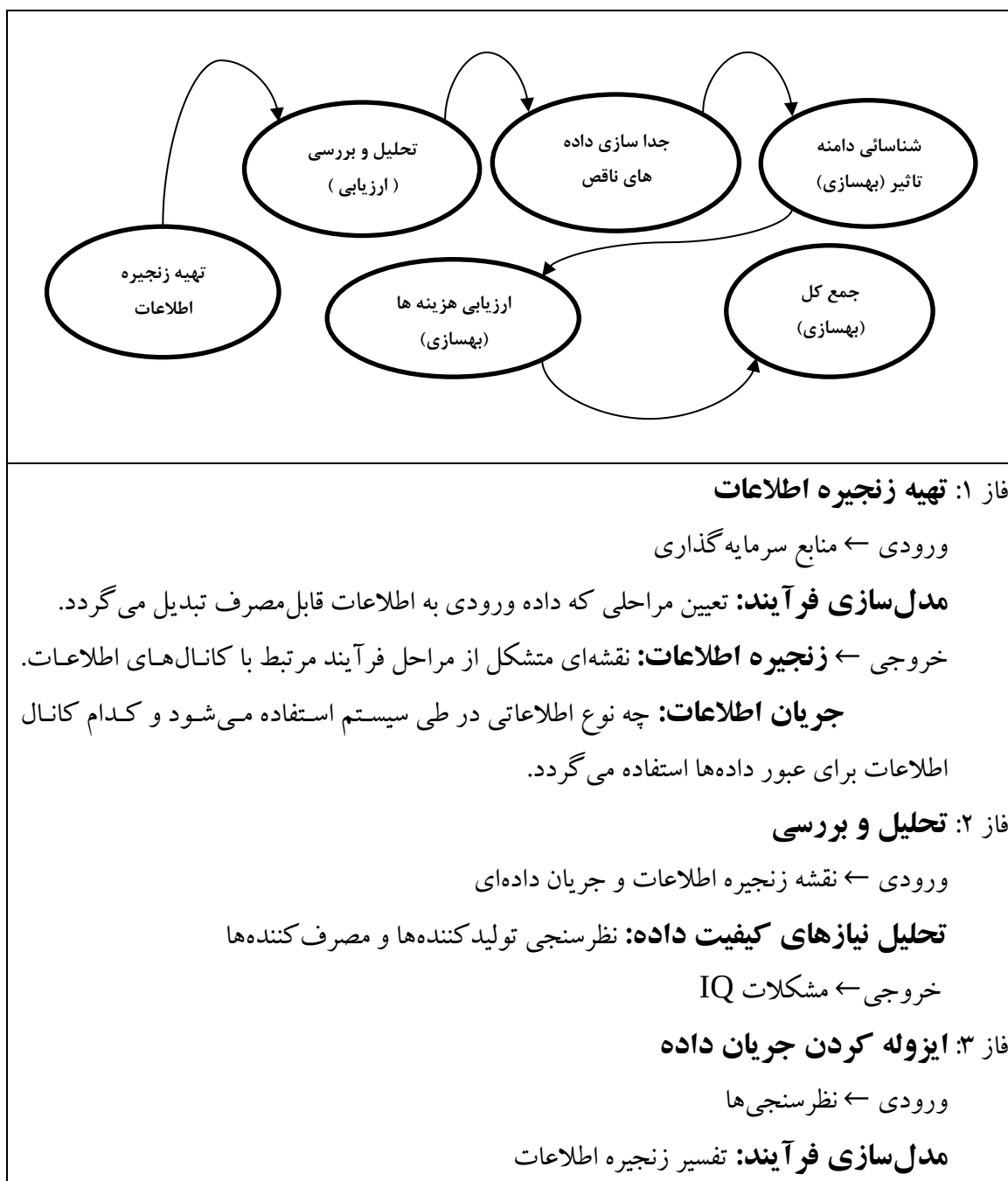
توضیحات کلی: هدف اصلی از روش ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها (Loshin 2004) ارائه کارت امتیاز کیفیت داده جهت حمایت از ارزیابی تأثیرات کیفیت پائین داده بر هزینه‌ها می‌باشد. مانند روش مدیریت کیفیت اطلاعات کلی، این روش نیز به طبقه‌بندی هزینه‌ها، منابع و شرح آن‌ها می‌پردازد. این روش مستقیماً از تمرکز بر هزینه‌های کیفیت پائین به دلیل استفاده از تکنیک‌های بهسازی حاصل می‌گردد.

هدف این روش ارزیابی فرآیندهای کسب و کار با اطلاعات تحت تأثیر قرار گرفته است.

توضیحات تفصیلی: شش مرحله جهت ارزیابی هزینه‌های کیفیت پایین داده طراحی شده است (شکل ۱۶ را مشاهده نمایید). در مرحله اول این روش، تجارت به وسیله دو مدل جریان داده‌ای مدل‌سازی می‌شود. جریان داده‌های استراتژیک که برای تصمیم‌گیری استفاده می‌شوند و جریان داده‌های عملیاتی که برای پردازش داده‌ها استفاده می‌شوند.

هر دو مدل نشان‌دهنده مجموعه‌ای از مراحل پردازش است که جریان اطلاعات را از منبع داده تا موارد مصرف داده توصیف می‌نماید. تجزیه و تحلیل عینی و ذهنی فضای کسب و کار بر اساس این مدل‌ها انجام می‌شود. کاربران داخلی و خارجی، کارکنان و مشتریان به منظور شناسایی داده‌های ناقص بررسی می‌شوند. سپس خطاها به فعالیت‌های معیوب در مدل‌های استراتژیک و عملیاتی کسب و کار نسب داده می‌شوند. این رابطه بین خطاها و فعالیت‌های پایه‌ای بر اساس ارزیابی هزینه ایجاد می‌شود. روش ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها یک طبقه‌بندی کامل و ارزشمند عملیاتی و تاکتیکی فراهم کرده و اثرات استراتژیک و اقتصادی را در نظر می‌گیرد. هر طبقه از هزینه‌ها ارزش اقتصادی خاصی دارند که بر اساس دانش تعیین می‌گردد. هزینه‌ها به عنوان ورودی در مرحله بهبود نهائی استفاده می‌گردند.

در نهایت روش ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها از تجزیه و تحلیل هزینه‌ها، ارزیابی و جمع‌بندی هزینه‌های پروژه‌های بهبود کیفیت را پشتیبانی می‌کند. رهنمودهای این روش نیز از محاسبه بازگشت سرمایه‌گذاری (ROI) و نقاط شکست بهبود پشتیبانی می‌کند.



خروجی ← تعیین فعالیت‌های قابل استفاده جریان داده

فاز ۴: شناسایی داده‌های تأیید شده

ورودی ← تعیین فعالیت‌های قابل استناد جریان داده.

انتخاب استراتژی و تکنیک‌ها: تعیین تأثیر استراتژیک، فنی و تأثیرگذار

خروجی ← دسته‌بندی فعالیت‌ها بر اساس کیفیت پایین داده

فاز ۵: ارزیابی هزینه‌ها

ورودی ← دسته‌بندی فعالیت‌ها بر اساس کیفیت پایین داده

ارزیابی هزینه‌ها: اندازه‌گیری هزینه برای هر یک از تأثیرات استراتژیک، فنی و تأثیرگذار

خروجی ← کارت کیفیت اطلاعات

فاز ۶: جمع‌بندی

ورودی ← ماتریس کارت کیفیت اطلاعات

ارزیابی هزینه‌ها

خروجی ← هزینه کیفیت پایین

شکل ۲-۱۶. مراحل ارزیابی هزینه مؤثر برای کیفیت کم داده‌ها

۲-۱۳-۱۱- روش DaQuinCIS (کیفیت داده در سیستم‌های اطلاعاتی تعاونی)

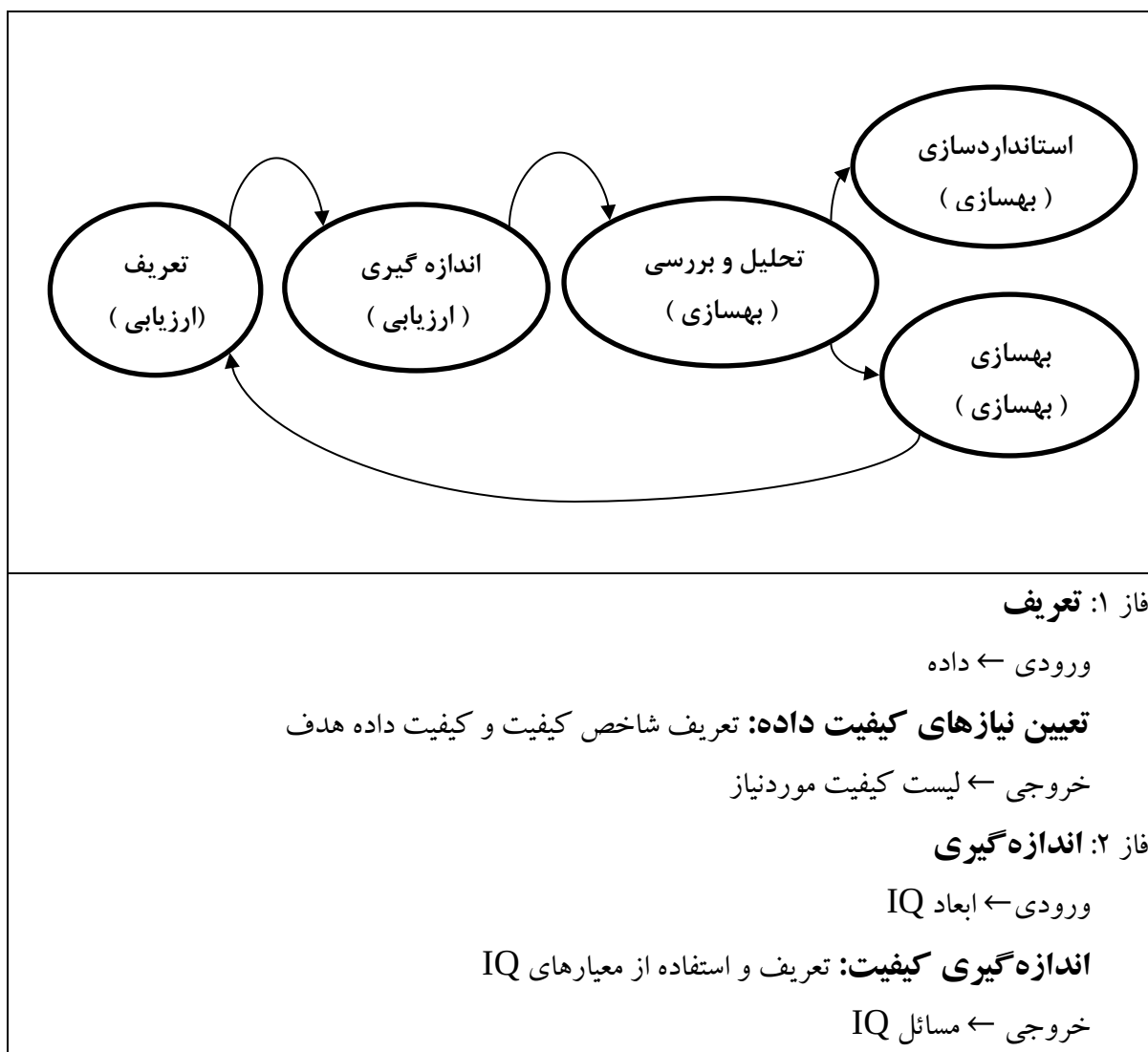
توضیحات کلی: روش ارزیابی کیفیت در سیستم‌های اطلاعات تعاونی (Scannapieco et al. 2004) مربوط به کیفیت داده‌ها در سیستم‌های اطلاعاتی تعاونی است. این روش بر دو مسئله مربوط به کیفیت داده‌ها متمرکز می‌شود. اول از همه، کیفیت داده‌ها را بر اساس اعتماد سازمانی تعیین می‌کند. در این روش کیفیت پایین داده‌ها می‌تواند مانع هماهنگی شده و عواقب زیادی در پی داشته باشد. برای پرداختن به مسئله‌ی اول در این روش به معرفی مفهوم صدور گواهینامه کیفیت داده می‌پردازیم که همراه داده‌ها در هنگام ردوبدل شدن بین سازمان‌ها استفاده می‌شود. مسئله دوم به مکانیسم انتخاب داده‌های مبتنی بر کیفیت مربوط می‌شود. مکانیسم این انتخاب‌ها بر اساس بالاترین کیفیت داده در بین داده‌های متعلق به سازمان‌های مختلف می‌باشد. در این روش، هماهنگی اهرمی برای بهبود کیفیت می‌باشد. توضیحات تفصیلی: این روش یک مدل ابتکاری جهت نشان دادن کیفیت داده‌ها به نام داده و کیفیت داده‌ها ارائه می‌کند که شامل موارد زیر است:

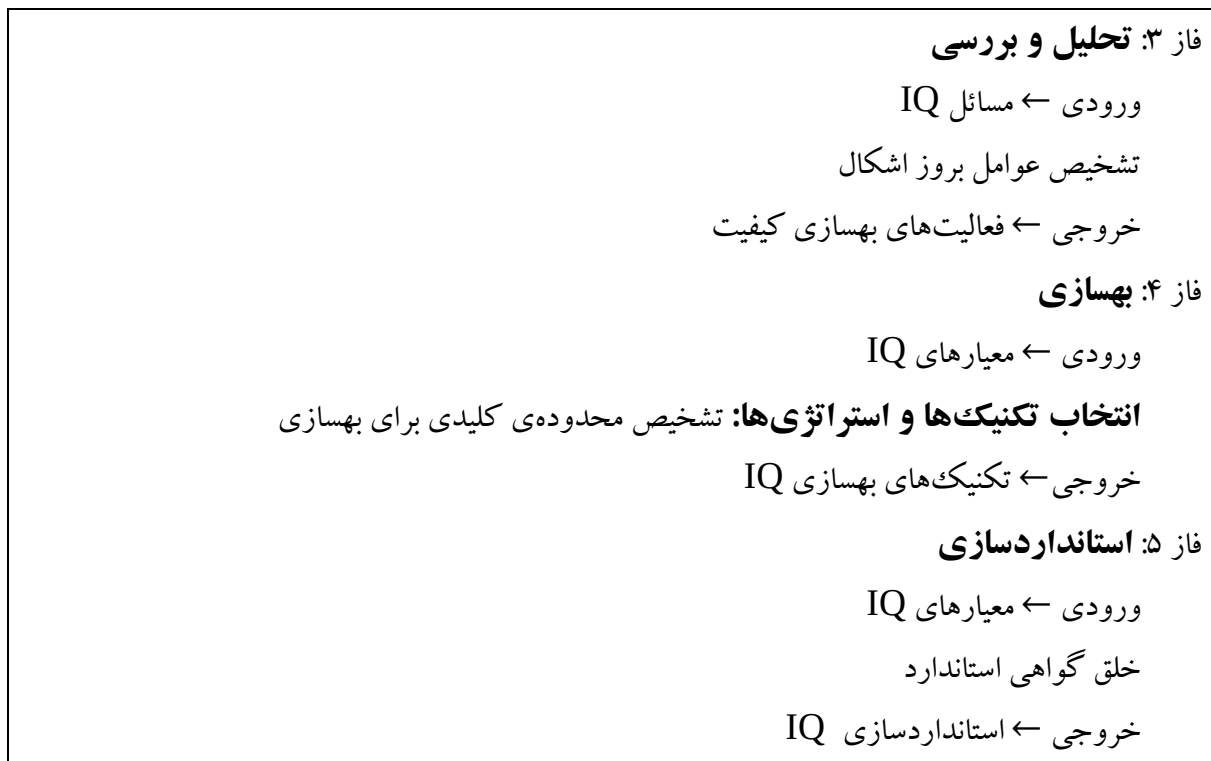
۱. ساختاری برای نشان دادن داده
۲. مجموعه‌ای از خواص کیفی داده
۳. ساختار ارائه‌دهنده خواص و کیفیت
۴. ارتباط بین داده‌ها و کیفیت ابر داده‌ها

منبع معتمد شامل خواص کیفیت می‌باشد. ارزش مناسب با این بعد توسط یک سازمان شخص ثالث بر اساس پارامترهای مختلف از مجله شکایت‌های ساخته شده توسط دیگر سازمان‌ها و درخواست‌های صادر شده تعیین می‌گردد. شکل ۱۷ نشان‌دهنده‌ی مراحل اساسی از این روش به شرح زیر است:

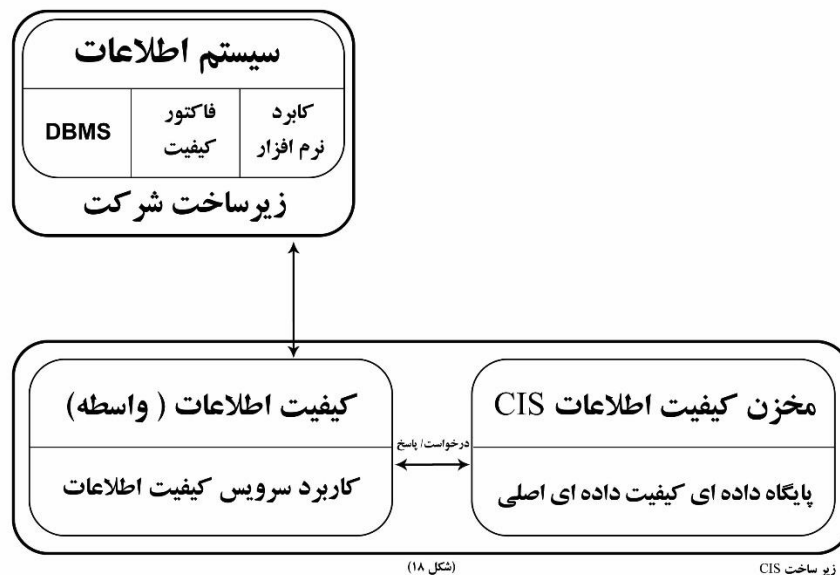
- تجزیه و تحلیل کیفیت
- ارزیابی کیفیت
- صدور گواهینامه کیفیت
- بهبود کیفیت

این روش توسط یک سامانه شامل زیرساخت‌های داخلی و زیرساخت‌های خارجی پشتیبانی می‌گردد. مراحل این روش توسط ماژول‌های مربوط به شاخص کیفیت در هر سازمان متعلق به کشورهای مشترک‌المنافع اجرا می‌شود. ارتباط میان سازمان‌های کشورهای مشترک‌المنافع توسط خدمات اطلاع‌رسانی با کیفیت می‌باشد که از یک موتور مشترک در انتشار به عنوان پیام‌رسان اصلی بین اجزای مختلف استفاده می‌گردد. در معیار کیفیت، درخواست کاربران خارجی کیفیت مورد نیاز لحاظ می‌شود و در نهایت ماژول مناسب ارزیابی کیفیت داده‌ها و ارزش کیفی توسط کاربران تعیین می‌گردد.





شکل ۲-۱۷. فازهای ارزیابی کیفیت در سیستم‌های اطلاعات تعاونی



شکل ۲-۱۸. زیر ساخت‌های ارزیابی کیفیت در سیستم‌های اطلاعات تعاونی

اگر داده‌ها از نظر کیفیت راضی‌کننده نباشند، یک هشدار به کاربر فرستاده می‌شود. در مقابل اگر کیفیت رضایت‌بخش باشد گواهی کیفیت مرتبط با داده به کاربر ارسال می‌گردد. بهسازی کیفیت توسط واسطه‌ی کیفیت اطلاعات انجام می‌گردد. این مازول به وسیله کالیبره شدن با مخزن اطلاعات و بر اساس طرح جامع اطلاعات را تفسیر می‌نماید و داده‌های با کیفیت بالا را انتخاب می‌کند. یکی از تردیدهایی که شرکت‌های خاص دارند این است که کیفیت بر اساس اطلاعات درخواستیشان باشد.

نسخه‌های مختلفی از داده‌های مشابه به عنوان پاسخ درخواست بررسی می‌شوند و در بین آن‌ها بهترین کیفیت انتخاب شده و به شرکت درخواست‌کننده ارائه می‌گردد. با توجه به تکنیک‌های مورد استفاده در فاز بهسازی، روش DiQuinCIS یک الگوریتم جدید را برای تطبیق رکورد (اطلاعات ثبت شده) ارائه می‌دهد. تطبیق‌دهنده‌ی رکورد هم جزئی از واسطه (بررسی‌کننده) کیفیت اطلاعات است. مکانیزم تطبیق‌دهنده رکورد برای بررسی تطبیق کیفیت داده‌های صادر شده از شرکت‌های همکار می‌باشد.

۲-۱۳-۱۲- روش QAFD (ارزیابی کیفیت داده‌های مالی)

توضیحات کلی: روش ارزیابی کیفیت داده‌های مالی (De Amicis and Batini 2004) برای تعریف استانداردهای کیفیت داده‌های عملیاتی مالی و به حداقل رساندن هزینه‌های ابزارهای سنجش کیفیت طراحی شده است. این روش ترکیبی از کیفیت و کمیت برای تعریف مسائل مربوط به کیفیت و انتخاب اقدامات مناسب جهت بهبود کیفیت می‌باشد. در این روش شاخص‌های مربوط به قوانین کیفیت داده‌ها، اندازه‌گیری و استراتژی‌هایی برای ارزیابی کمی و کیفی تعریف می‌شوند. به‌طور کلی این روش برای کیفیت اطلاعات مالی مناسب می‌باشد.

توضیحات تفصیلی: مراحل اصلی این روش در شکل ۱۹ ترسیم شده است. اول، این روش مناسب‌ترین متغیرهای مالی را انتخاب می‌کند که این انتخاب معمولاً بر اساس دانش به‌دست آمده از ارزیابی‌های قبلی، با توجه به میزان اثربخشی آن‌ها انجام می‌شود. متغیرها در دسته مسائل مرتبط که تحت تأثیر ریسک‌های سرمایه‌گذاران، مصرف‌کنندگان و عوامل دیگر فضای کسب‌وکار قرار دارند، گروه‌بندی می‌شوند.

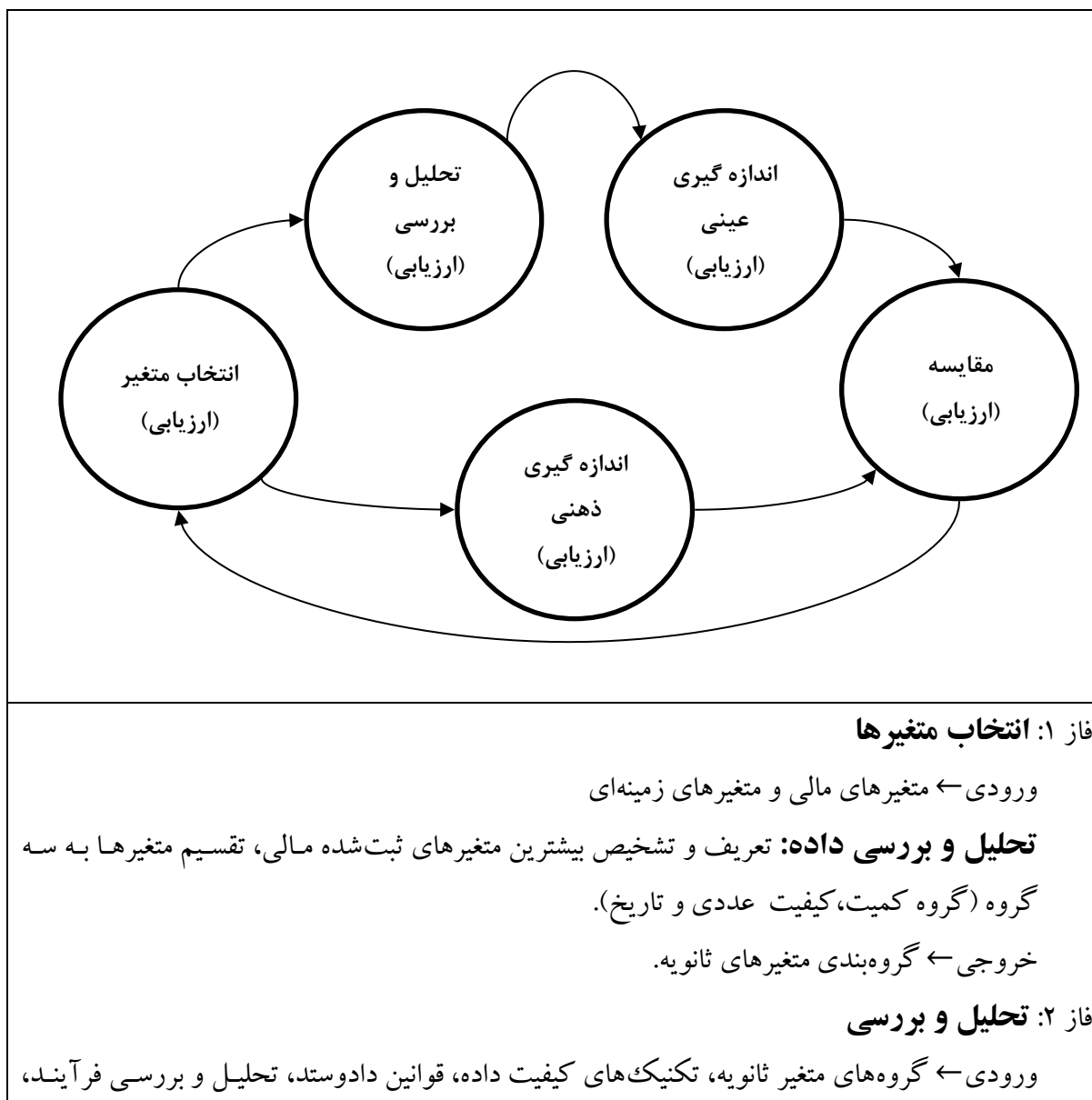
فاز دوم این روش باهدف کشف علل اصلی خطاهاست. مناسب‌ترین ابعاد کیفیت داده‌ها در این مرحله تعیین شده و قوانین کیفیت داده به وجود می‌آیند. قوانین کیفیت داده‌ها نشان‌دهنده‌ی خواص دینامیک متغیرهاست که می‌تواند در ابعاد کیفی اندازه‌گیری شود.

در مرحله سوم، ارزیابی عینی بر اساس شاخص‌های کمی انجام می‌شود. مؤلفین این روش یک مدل ریاضی برای ارزیابی عینی پیشنهاد داده‌اند که در زمان رتبه‌بندی کلی داده‌ها با هر یک از ابعاد کیفیت ارائه می‌شود. ارزیابی ذهنی در مرحله چهارم از سه دیدگاه مختلف انجام می‌شود.

- کارشناسان کسب‌وکار
- مشتریان
- کارشناسان کیفیت داده

که در این سطوح هر شخص برای ارزیابی سطح کیفیت، یک بعد کیفی را بررسی می‌نماید.

ارزیابی کلی به دست آمده به عنوان میانگین ارزیابی ذهنی از هر طبقه‌ی کارشناسان ثبت می‌گردد. در پایان، ارزیابی ذهنی و عینی در فاز پنجم صورت می‌گیرد و برای هر یک از ابعاد، تفاوت ارزیابی ذهنی و عینی محاسبه می‌گردد. اگر تفاوت مثبت باشد، ارزیابی عینی باید بر اساس نقطه نظر کارشناسان کیفیت تجدیدنظر گردد.



مشاهده

تحلیل نیازهای کیفیت داده: تحلیل ابعاد کیفیت داده، درستی نحوی و معنایی، ثبات داخلی و خارجی، تکمیل، واحد پول، جداول زمانی، یکتایی - **تعریف محدوده بحرانی:** کاربرد قوانین کیفیت داده

خروجی ← بررسی کمی

فاز ۳: **اندازه‌گیری اهداف**

ورودی ← تعریف خطاها

ارزیابی کیفی: تعریف شاخص‌هایی برای ارزیابی و تحلیل اهداف سطح‌بندی داده‌ها

خروجی ← ارزیابی اهداف کمی

فاز ۴: **اندازه‌گیری اهداف کیفی**

ورودی ← تخصص دادوستد و کیفیت اطلاعات، اندازه‌گیری، دسته‌بندی متغیرهای ثانویه، بررسی

تخصصی کیفیت داده، مقایسه ارزیابی عینی خروجی ← بررسی کیفی

فاز ۵: **مقایسه**

ورودی ← ابعاد کیفی هدف، ابعاد کیفی موضوع **اندازه‌گیری کیفیت:** محاسبه اختلافات

خروجی ← ارزیابی و اندازه‌گیری

شکل ۲-۱۹. فازهای ارزیابی کیفیت داده‌های مالی

۲-۱۳-۱۳- روش CDQ (کیفیت داده‌های کامل)

توضیحات کلی: روش جامع برای مدیریت کیفیت داده (Batini and Scannapieco 2006; Batini et al. 2008) به طور یکجا، کامل، انعطاف پذیر و دارای استفاده ساده است. این روش با در نظر گرفتن تکنیک‌ها و ابزارهای موجود و یکپارچه سازی آن‌ها در یک چهارچوب می تواند در هر زمینه درونی و برونی سازمانی کار کند و همه نوع داده‌ی ساختاریافته، نیمه ساختاریافته و بدون ساختار را در بر گیرد.

این روش از کاربران در انتخاب مناسب ترین تکنیک‌ها و ابزارها در هر فاز و در هر زمینه پشتیبانی می کند لذا دارای انعطاف پذیری زیادی می باشد. در مجموع می توان به جمع بندی رسید که روش CDR به علت اینکه تکنیک‌های اعمال شده را بر اساس اهداف مشخص شده در فازها انتخاب می کند، کاربرد ساده‌ای دارد. این روش نوآورانه است چرا که انتخاب بهینه‌ای برای فرآیند بهبود کیفیت با حداکثر سازی مزایای بودجه‌ای را در شرایط مختلف فراهم می کند.

این روش بر نیازهای ابتدایی فاز استخراج تأکید دارد. در واقع روش‌های دیگر با این فرض که اطلاعات به دست آمده قبلاً جمع آوری و مدل شده‌اند کار می کنند و تمرکز آن‌ها بر چگونگی رسیدن به کیفیت کلی داده‌ها بدون در نظر گرفتن حالت استفاده از دانش قبلی است. اما هدف اصلی روش جامع برای مدیریت کیفیت داده تشخیص این است که یک ارزیابی کمی تا چه حد توسط فرآیند کسب و کاری با اطلاعات بد تحت تأثیر قرار می گیرد.

توضیحات تفصیلی: سه مرحله اصلی توصیف این روش عبارت‌اند از بازسازی، ارزیابی و انتخاب فرآیند بهبود بهینه (شکل ۲۰ را مشاهده نمایید). در مرحله اول این روش، روابط بین واحدهای سازمانی، فرآیندها، خدمات و داده‌ها بازسازی می شوند.

این روابط با استفاده از ماتریس‌هایی که واحدهای سازمانی از داده‌ها و نقش آن‌ها در فرآیندهای کسب و کار مختلف می سازند، تعریف می شود. همچنین در این مرحله فرآیندها بر اساس سهمشان در تولید محصولات، قوانین حقوقی و سازمانی شرح داده می شوند. مرحله دوم یک هدف کیفی جدید را تعریف می کند که

مورد نیاز فرآیند بهبود کیفیت، ارزیابی هزینه‌ها و مزایا می‌باشد. این مرحله به متغیرهای بحرانی تحت تأثیر کیفیت پایین می‌پردازد. از آنجا که فعالیت‌های بهسازی پیچیده و هزینه‌بر هستند، بهتر است تمرکز روی بخشی از پایگاه داده‌ای و جریان داده باشد که باعث بروز ایرادات بزرگ می‌شود. در نهایت، فاز سوم شامل پنج مرحله است و هدف آن شناسایی روند بهبود می‌باشد. در این مرحله دنباله‌ای از فعالیت‌ها که دارای بالاترین نسبت هزینه اثربخشی می‌باشند شناسایی می‌شوند. هدف کیفی جدید با در نظر گرفتن هزینه‌ها و منافع تنظیم می‌گردد. فعالیت‌های بهسازی می‌توانند برای دستیابی به هدف کیفی جدید انجام گیرند. این روش تکنیک‌های بهسازی داده محور و فرآیند محور را برای پایگاه‌های تحت تأثیر کیفیت پایین پیشنهاد می‌کند.

در نهایت مجموعه‌ای از تکنیک‌های بهسازی سازگار به صورت یک فرآیند بهسازی مناسب، تحلیل هزینه - فایده برتر را انتخاب می‌کنند.



شکل ۲-۲۰. فازهای روش جامع برای مدیریت کیفیت داده

۲-۱۴- خلاصه‌ای از روش‌های ذکر شده

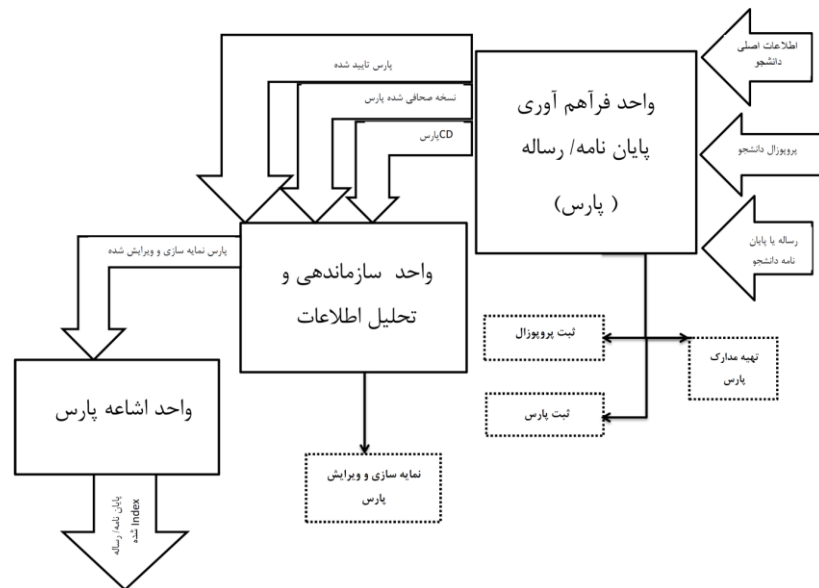
جدول ۲-۱۵. جدولی از جمع‌بندی روش‌های مطالعه شده

مخفف روش	نام اصلی	تحلیل داده‌ها	تحلیل الزامات DQ	شناسایی حوزه بحرانی	مدل‌سازی فرآیند	اندازه‌گیری کیفیت	قابل توسعه به دیگر ابعاد و معیارها
TDQM	مدیریت کیفیت داده	+		+	+	+	ثابت
DWQ	روش‌شناسی کیفیت انبار داده	+	+	+		+	باز
TIQM	مدیریت کیفیت اطلاعات کلی	+	+	+	+	+	ثابت
AIMQ	روشی برای ارزیابی کیفیت اطلاعات	+		+		+	ثابت
CIHI	روش موسسه کانادایی برای روش اطلاعات سلامت	+		+			ثابت
DQA	ارزیابی کیفیت داده	+		+		+	باز
IQM	اندازه‌گیری کیفیت اطلاعات	+				+	باز
ISTAT	روش ISTAT	+				+	ثابت
AMEQ	روش اندازه‌گیری مبتنی بر فعالیت و ارزیابی کیفیت اطلاعات محصول	+		+	+	+	باز
DaQuinCIS Data	کیفیت در سیستم‌های اطلاعات تعاونی	+		+	+	+	باز
QAFD	روشی برای ارزیابی کیفیت داده‌های مالی	+	+	+		+	ثابت
CDQ	روشی جامع برای مدیریت کیفیت داده	+	+	+	+	+	باز

فصل سوم – شناسایی وضعیت موجود

۳-۱ معرفی مدل فرایندی سامانه‌های گردآوری و ثبت، سازماندهی و اشاعه

در این بخش به معرفی مدل فرایندی این نوع سامانه‌های اطلاعاتی و معرفی مشکلات کیفی آنها خواهیم پرداخت. نمودار جریان ورودی‌ها و خروجی‌های در سامانه‌های اطلاعاتی گردآوری و ثبت، سازماندهی و اشاعه اطلاعات پارس در شکل زیر درج گشته است.



شماره بازیابی: ۰۴

شکل ۳-۱. ورودی و خروجی سامانه‌های گردآوری و ثبت، سازماندهی، و اشاعه اطلاعات پارس

• فرایند فراهم آوری اطلاعات

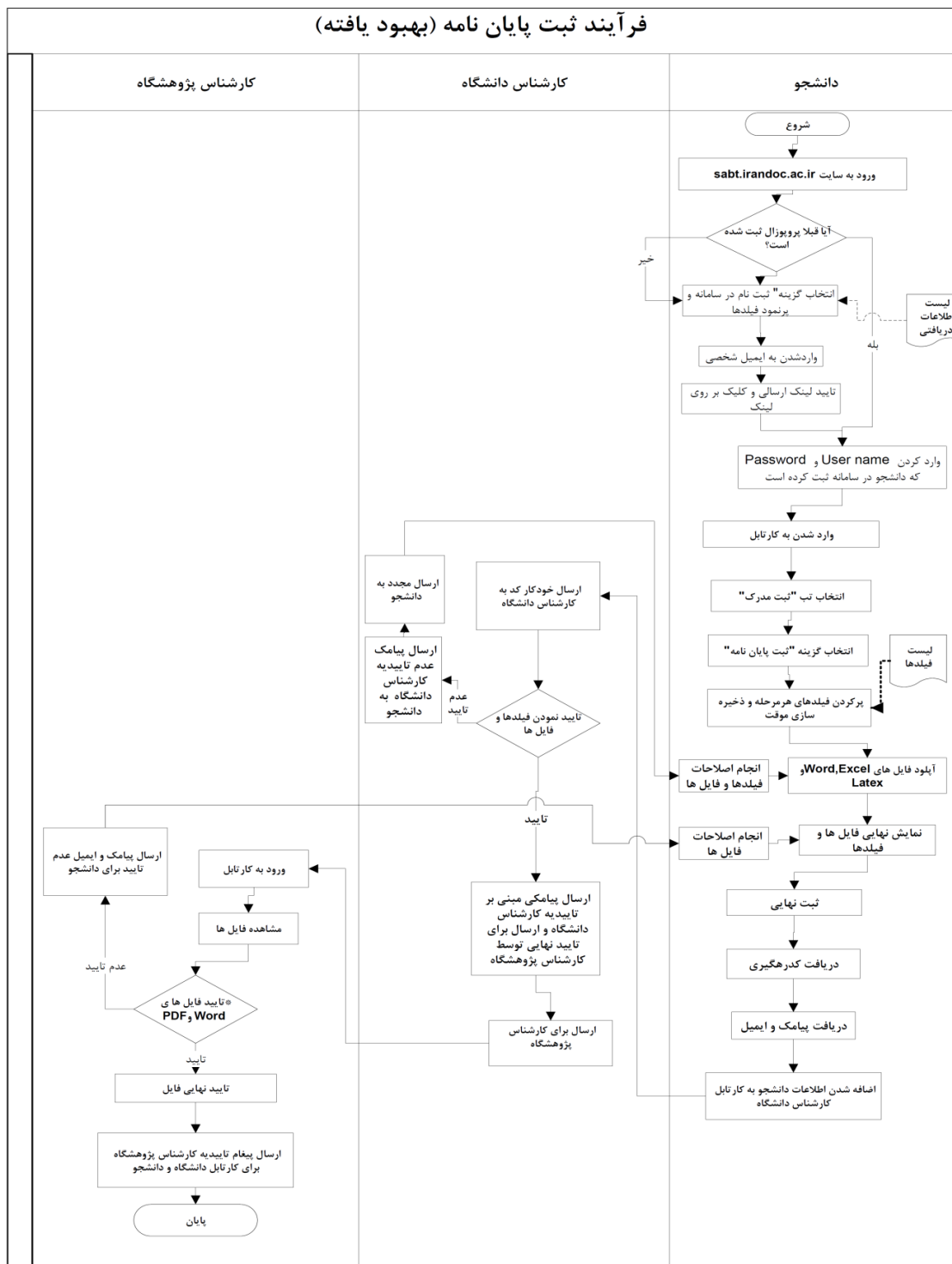
فرایند فراهم آوری اطلاعات شامل بخش‌های تهیه مدارک و ثبت اطلاعات می‌باشد. در تمامی این زیر بخش‌ها و در تمامی فرآیندهای ذکر شده عملیات کنترل کیفی اجرا خواهد شد تا از صحت و بی‌مشکل بودن اطلاعات در سیستم اطمینان خاطر کسب شود.

• فرایند تجزیه و تحلیل اطلاعات

وظایف فرایند تجزیه و تحلیل اطلاعات شامل ویرایش اطلاعات دریافتی و نمایه سازی اطلاعات است. باتوجه به اینکه اطلاعات وارد شده در این بخش قبلاً توسط کاربران ثبت گردیده است لازم است کلیه نواقص در این مرحله مرتفع گردد. همچنین کلیدواژه‌های مولف بررسی و کنترل می‌گردد و در صورت لزوم اصلاحات مقتضی صورت خواهد پذیرفت.

• فرایند اشاعه اطلاعات تحقیقاتی

با توجه به این که در این بخش از اطلاعات خروجی بخش‌های قبلی استفاده می‌شود در صورت وجود و شناسایی مشکلات کیفی این اطلاعات مجدداً به واحد تجزیه و تحلیل اطلاعات عودت داده می‌شود. از آنجاکه پایان‌نامه‌ها و رساله‌های اشاعه داده شده در سامانه گنج مبنای پژوهش خیل کثیری از پژوهشگران در مقطع کارشناسی ارشد و دکتری قرار خواهد گرفت لذا ارتقاء کیفیت این سامانه تأثیری انکارناپذیر در افزایش کیفیت پژوهش‌های صورت گرفته از سوی پژوهشگران خواهد داشت و ضرورت انجام یک تحقیق جامع بر روی دسته‌بندی مشکلات کیفی این سامانه را دوچندان خواهد نمود. مراحل فرایند ثبت پایان‌نامه/رساله بشرح در شکل ۳-۲ نمایش داده شده است. (ارشادی و احترامی ۱۳۹۵) همانگونه که در این فرایند مشخص است کارشناس دانشگاه و کارشناس پژوهشگاه (پژوهشگاه علوم و فناوری اطلاعات) در این فرایند اقدامات کنترل کیفی را در راستای کنترل اقلام اطلاعاتی فراداده‌های پایان‌نامه/رساله‌ها صورت خواهند داد. لازم است این اقدامات در قالب مدل کیفیت فراداده تبیین و طراحی گردد.



شماره بازبینی: ۵۰

۴۸ تا ۷۲ ساعت بازه زمانی فرصت برای تاییدی نهایی کارشناس پژوهشگاه از زمانی که تاییدیه کارشناس دانشگاه را دریافت می کند

شکل ۳-۲. فرایند ثبت پایان نامه / رساله در سامانه ثبت

۳-۲ شناسایی وضعیت سامانه ثبت

از سال ۱۳۸۷ بر اساس ابلاغیه وزیر علوم، تحقیقات و فناوری وقت برای فراهم کردن امکان دسترسی پژوهشگران و علاقه مندان به پایان نامه ها، جلوگیری از تکرار مطالعات انجام شده و تکمیل پایگاه اطلاعاتی پایان نامه های دانشجویی کشور در پژوهشگاه، دانشجویان موظف هستند یک نسخه از پایان نامه دوره های کارشناسی ارشد و دکتری خود را به پژوهشگاه ارسال کنند و فارغ التحصیل شدن آن ها هم منوط به انجام این کار است. در این راستا پایگاه ثبت پیشنهاد و اطلاعات پایان نامه ها و رساله های تحصیلات تکمیلی طراحی شد. که در حال حاضر با آدرس www.thesis.irandoc.ac.ir مورد استفاده دانشجویان و پژوهشگران قرار گرفته است.

با گذشت زمان تلاش برای روزآمدسازی و توسعه نسخه جدیدتر این سامانه، با نیازمندیهای جدید و ابزارها و ساختار جدید آغاز گشت. در این راستا جلسات تحلیلی با حضور اعضای کارگروه بهبود پایگاههای اطلاعاتی و دیگر صاحب نظران برگزار شد. سیستم ثبت جدید دارای سه نوع ثبت مدارک علمی پیش بینی گردید.

- ثبت فردی
- ثبت دستی
- ثبت تجمیعی (بالک)

مستند زیر تنها به تحلیل بخش ابتدایی ثبت، ثبت فردی پرداخته است.

در این سامانه امکان تعریف گروه کاربران متفاوت با دسترسیهای متفاوت توسط کاربر اصلی یا ادمین سیستم وجود دارد از آنجاییکه در این سند به معرفی ثبت فردی پرداختیم، تنها کاربران در ارتباط با این فرآیند را معرفی میکنیم.

دانشجو: در اینجا منظور از دانشجو، کاربری است که از طریق وب، با عنوان دانشجوی تحصیلات تکمیلی جهت ثبت اطلاعات رساله خود به پایگاه مراجعه مینماید .

ادمین اصلی (پژوهشگاه): منظور از ادمین اصلی، کاربری است که به عنوان ادمین در پژوهشگاه علوم و فناوری اطلاعات ایران حضور دارد و کلیه دسترسیها در سیستم را در اختیار دارد.

مدیر دانشگاه (دانشکده): در اینجا منظور مدیریت دانشگاه و یا واحد مربوطه میباشد که در سازمان یا دانشگاه مربوطه قرار دارد و توسط مدیر اصلی سیستم که در پژوهشگاه حضور دارد به صورت زیر تعریف میشود:

- نام و نام خانوادگی ریاست دانشگاه
- شماره تماس دانشگاه
- نام کاربری (به صورت دستی ساخته میشود)
- کلمه عبور (به صورت دستی ساخته میشود)
- لیست اسامی دانشگاهها
- جهت انتخاب نام دانشگاه از لیست مستند اسامی سازمانها
- انتخاب نوع سازمان دانشگاه (آزاد - پیام نور - وزارت علوم و)
- تعیین سطح دسترسی کاربر شامل
 - تایید مدارک
 - امکان پیشنهاد سازمان

در صورتیکه ادمین اصلی سیستم به کاربری این سطح دسترسی را بدهد کاربر در صورتیکه سازمان یا دانشکده خود را در لیست اسامی سازمانهای آمده از پایگاههای اطلاعات پایه سازمانها نیابد میتواند از این طریق نام سازمان خود را پیشنهاد دهد تا به صحت پیشنهاد آن در سیستم ویرایش پرداخته شود و در صورتیکه نام سازمان در سیستم ویرایش تأیید شود به بانک اطلاعات پایه سازمانها اضافه شده و در لیست سازمانها در سیستم ثبت قابل مشاهده خواهد بود.

○ مدیریت کاربران سازمان

○ ارسال فایل xml

○ ارسال فایل‌های تمام‌متن

○ تعیین سطح اشاعه

○ ثبت پایان‌نامه فیزیکی

○ گزارشات کاربران

○ و...

● ثبت و ذخیره

همچنین ادمین اصلی سیستم امکان تعریف گروه‌های کاربری متفاوت را دارد که این گروه‌های کاربری می‌توانند برای هر دانشگاه یا سازمان متفاوت از دانشگاه یا سازمان دیگر باشند، و می‌تواند به هر یک از گروه‌های کاربری تعریف شده دسترسی‌های متفاوتی را بدهد (لیست دسترسی‌ها در سیستم تعیین شده است). تعدادی از این گروه‌های کاربری می‌تواند به شرح زیر باشد:

کارشناس دانشگاه (دانشکده): در اینجا منظور کارشناسی می‌باشد که در سازمان یا واحد مربوطه قرار دارد و می‌تواند توسط مدیر دانشگاه تعریف شده و دارای اختیارات بازیابی، تایید و ویرایش اطلاعات ثبت شده توسط دانشجو و ... باشد.

کارشناس پژوهشگاه (ثبت دستی) یکی دیگر از گروه کاربرانی که توسط ادمین اصلی پژوهشگاه تعریف میشود کارشناس ثبت دستی میباشد که به اختیارات و فرآیند کاری آن در فاز دوم مستند پرداخته میشود.

۳-۳ شناسایی وضعیت داده‌های ارائه شده از سامانه CRM

این گزارش به عملکرد گروه پشتیبانی سامانه های ایرانداک در بررسی و پاسخگویی به درخواست های کاربران سامانه های ایرانداک در بهار ۱۳۹۶ می پردازد. در این مدت گروه پشتیبانی سامانه های ایرانداک پاسخگوی ۱۱ هزار و ۷ درخواست بوده، که نسبت به مدت زمان مشابه در سال گذشته با رشد ۱۷۲ درصدی، توانسته است پاسخگوی تعداد بیشتری از کاربران باشد. بیشترین درخواست ثبت شده کاربران در این دوره زمانی به ترتیب، مربوط به سامانه ملی ثبت، سامانه پیشینه پژوهش و پایگاه اطلاعات علمی ایران (گنج) بوده است.

از ابتدا تا پایان خرداد سال ۱۳۹۶ گروه پشتیبانی سامانه های ایرانداک پاسخگوی ۱۱ هزار و ۷ درخواست بوده است

در ۵۴ روز کاری یاد شده، درخواست های ثبت شده کاربران به شرح زیر بوده است:

جدول ۳- ۱ آمار وضعیت درخواستها در سامانه CRM

سامانه های ایراندک	سامانه ملی ثبت پایان نامه، رساله و پیشنهاد	سامانه پیشینه پژوهش	پایگاه اطلاعات علمی ایران (کنج)	دیگر سامانه ها
درخواست ثبت شده	۵۴۹۰	۴۳۸۸	۱۱۱۸	۱۱

همانطور که مشاهده می شود بیشترین درخواست ثبت شده کاربران در این دوره زمانی به ترتیب، مربوط به سامانه ملی ثبت، پیشینه پژوهش و پایگاه گنج بوده است.

از ابتدای رونمایی نسخه تازه سامانه ملی ثبت در آبان ۱۳۹۵، گروه پشتیبانی تماس های زیادی را از سوی کاربران داشت که

درخواست تغییر رایانامه یا شماره ملی خود را در سامانه داشتند. در بهار ۱۳۹۶ کارشناسان فنی اساء، کاربرگه ای را با هدف ساختارمند کردن این درخواست ها و رسیدگی بهتر و سریع تر به آنها طراحی کردند. از تاریخ ۸ خرداد ۱۳۹۶ که این کاربرگه در دسترس کاربران قرار گرفت تا پایان ماه جاری، در مدت ۲۴ روز، در حدود ۵۰۰ درخواست از کاربران به ثبت رسید و پاسخ داده شد.

موضوعات اصلی سامانه ملی ثبت عبارتند از:

- پرسش درباره فرایندهای کاربری نسخه تازه؛
- تغییر در حساب کاربری گذشته کاربران (درخواست تغییر رایانامه و شماره ملی)؛
- پرسش در مورد چگونگی پذیرش دانشگاه.

۳-۴ شناسایی نحوه کنترل فراداده‌ها در سامانه ثبت به کمک مصاحبه با خبرگان

بمنظور شناسایی دقیق وضعیت موجود سامانه ثبت و نحوه کنترل فراداده‌های مربوط به پایان‌نامه / رساله دانشجویان تحصیلات تکمیلی پس از برگزاری جلسات هماهنگی با مدیر واحد فراهم‌آوری جناب آقای محمودیان با برخی از کارشناسان کلیدی این واحد که مستقیماً در امر کنترل کیفیت فراداده‌ها مبادرت دارند مصاحبه حضوری به عمل آمده و از نزدیک با عملکرد این همکاران آشنایی صورت گرفت. سرکار خانم نبوی و جناب آقای ترابی و جناب آقای سطوت از افرادی بودند که با آنها بصورت حضوری در خصوص نحوه کار با سامانه ثبت مصاحبه به عمل آمد. نکته مهمی که در این مصاحبه‌ها وجود داشت این بود که هریک از این همکاران در خصوص نحوه کنترل فراداده‌ها و عکس‌العملی که در ارائه نتیجه به دانشجویان داشتند رویکرد مخصوص به خود داشته و روش یکسانی در این خصوص نداشتند. بعنوان مثال موارد زیر توسط یکی از همکاران بعنوان چک‌لیست کنترل مشاهده گردید.

ضمن تقدیر و تشکر از زحمات و دقت نظر همه همکاران محترم، لطفاً موارد زیر حتماً در کنترل فایلها مورد بررسی قرار گیرد و هر یک از موارد زیر در فایل های پایان نامه موجود نیست پذیرش نگردد:

- عنوان فارسی
- چکیده فارسی
- فهرست مطالب
- کلیه فصلها به همراه اشکال و جداول
- منابع و ماخذ
- پیوست ها (در صورت اشاره در فهرست مطالب)
- چکیده لاتین
- عنوان لاتین

نکات مهم:

- پایان نامه های دانشگاه تربیت مدرس صفحات حقوقی دانشگاه در ابتدای مدرک کنترل شود.
- پایان نامه هایی که فاقد فایل ورد هستند (فقط فهرست مطالب و منابع) و با نرم افزارهای دیگر نگاشته شده است، در صورتیکه فایل پی دی اف کامل را دارد، پذیرش نمایند.
- پایان نامه های دانشگاه جامعه المصطفی (ص) حتما دارای عنوان فارسی و چکیده فارسی باشد.
- اشکالات نمایش فونتها در فایل ورد حتما در سیستم دیگری بررسی و در صورت داشتن مشکل برگشت داده شود.
- فایل پایان نامه ها بدون پس نقش (واترمارک) باشد.
- صفحات سفید در فایل پی دی اف حذف گردد.
- فایل پایان نامه ها بدون قفل باشد و به راحتی قابل ویرایش باشد.
- فایل پایان نامه بدون آدرس و مشخصات شخصی دانشجو و استاد راهنما باشد.
- اشتباهات موجود در ورود اطلاعات کاربران تا حد ممکن اصلاح شود.

با تشکر - ۱۱ دیماه ۹۵

نمونه دیگری از چک لیستی که مورد استفاده قرار گرفته و خیلی کاملتر از چک لیست قبلی است به شرح زیر است.

لطفا فایل های ورد و پی دی اف را هر کدام بطور مجزا و کامل از صفحه ابتدا (بسم اله) تا آخرین صفحه بارگذاری نمائید : شامل صفحه عنوان فارسی و انگلیسی، چکیده فارسی و انگلیسی، فهرست مطالب با شماره صفحه و فهرست منابع و کل فصل های پایان نامه و پیوست ها است.

لطفا فایل ورد را بطور کامل از صفحه ابتدا (بسم اله) تا آخرین صفحه بارگذاری نمائید : شامل صفحه عنوان فارسی و انگلیسی، چکیده فارسی و انگلیسی، فهرست مطالب و فهرست منابع و کل فصل های پایان نامه و پیوست ها است. (همانند فایل پی دی اف)

در صورتی که قصد دارید فایل پروپوزال یا پیشنهاد خود را ثبت کنید، لطفا در بخش ثبت پیشنهاد اینکار

را انجام دهید. در غیر این صورت فایل‌های ورد و پی دی اف پایان نامه را هر کدام بطور مجزا و کامل بارگذاری نمائید: شامل صفحه عنوان فارسی و انگلیسی، چکیده فارسی و انگلیسی، فهرست مطالب و فهرست منابع و کل فصل‌های پایان نامه و پیوست‌ها است.

لطفا در فایل‌های ورد و پی دی اف شماره صفحات فهرست مطالب را کامل نموده و مجدداً بارگذاری نمائید.

لطفاً صفحه عنوان فارسی (اسم و لوگوی دانشگاه - دانشکده - گروه و گرایش و رشته - عنوان پایان نامه - مولف - استاد راهنما و مشاور - تاریخ دفاع) و صفحه چکیده فارسی را در فایل ورد و پی دی اف قرار داده و مجدداً بارگذاری نمائید.

لطفاً صفحه عنوان فارسی (اسم و لوگوی دانشگاه - دانشکده - گروه و گرایش و رشته - عنوان پایان نامه - مولف - استاد راهنما و مشاور - تاریخ دفاع) را در فایل ورد و پی دی اف قرار داده و مجدداً بارگذاری نمائید.

لطفاً صفحه عنوان انگلیسی (اسم و لوگوی دانشگاه - دانشکده - گروه و گرایش و رشته - عنوان پایان نامه - مولف - استاد راهنما و مشاور - تاریخ دفاع) و چکیده انگلیسی را در فایل ورد و پی دی اف قرار داده و مجدداً بارگذاری نمائید.

لطفاً صفحه عنوان انگلیسی (اسم و لوگوی دانشگاه - دانشکده - گروه و گرایش و رشته - عنوان پایان نامه - مولف - استاد راهنما و مشاور - تاریخ دفاع) را در فایل ورد و پی دی اف قرار داده و مجدداً بارگذاری نمائید.

ثبت پروپوزال یا پیشنهاد در بخش دیگر همین سامانه به نام ثبت پیشنهاد می‌باشد که می‌توانید فایل‌های خود را آنجا بارگذاری نمائید. در صورتی که پایان نامه شما دفاع شده، می‌توانید در این قسمت فایل‌های آن را بارگذاری نمائید.

در صورتی که در تهیه فایل پایان نامه خود از نرم‌افزار دیگری **غیر از Word** استفاده نموده‌اید، فایل‌های آن را به شکل زیپ شده در **بخش دیگر فایل‌ها** قرار دهید.

لطفاً در فایل‌های ورد و پی دی اف صفحه عنوان انگلیسی، چکیده فارسی و انگلیسی را قرار داده و مجدداً بارگذاری نمائید.

لطفاً در فایل‌های ورد و پی دی اف صفحه عنوان انگلیسی و چکیده انگلیسی را قرار داده و مجدداً بارگذاری نمائید.

لطفا در فایل‌های ورد و پی‌دی‌اف صفحه عنوان انگلیسی را قرار داده و مجدداً بارگذاری نمایید.
لطفا در فایل‌های ورد و پی‌دی‌اف صفحه چکیده انگلیسی را قرار داده و مجدداً بارگذاری نمایید.
لطفا پیوست‌ها و ضمائم ذکر شده در فهرست مطالب را در فایل‌های ورد و پی‌دی‌اف قرار داده و مجدداً بارگذاری نمایید.

لطفا در فایل‌های ورد و پی‌دی‌اف صفحات تایید داوران، آئین نامه حق مالکیت، آئین نامه حق چاپ را بصورت امضا شده، قرار داده و مجدداً بارگذاری نمایید. (ترتیب مدرس)
لطفا در فایل‌های ورد و پی‌دی‌اف صفحه عنوان فارسی و چکیده فارسی را قرار داده و مجدداً بارگذاری نمایید.

فایل پی‌دی‌اف خراب است و باز نمی‌شود و لطفاً در فایل ورد فهرست منابع فارسی و غیر فارسی را قرار داده و مجدداً بارگذاری نمایید.

لطفا در فایل‌های ورد و پی‌دی‌اف چکیده فارسی را قرار داده و مجدداً بارگذاری نمایید.
لطفاً کلیدواژه‌های فارسی و انگلیسی درست و مناسب را قرار داده و مجدداً بارگذاری نمایید.
لطفا در فایل‌های ورد و پی‌دی‌اف فهرست منابع را قرار داده و مجدداً بارگذاری نمایید.
لطفاً از کلیدواژه‌های مناسب استفاده نمایید.

لطفاً در فایل پی‌دی‌اف در قسمت فرم مشخصات چاپ چکیده در پژوهشنامه آدرس شخصی را حذف نمایید و مجدداً بارگذاری نمایید.

شماره صفحات مندرج در فهرست مطالب با شماره صفحات مندرج در پایان نامه منطبق نیست. لطفاً اصلاح شود و سپس فایل‌های ورد و پی‌دی‌اف مجدداً بارگذاری نمایید.

پیوست پلان و نقشه‌ها در فایل ورد و پی‌دی‌اف موجود نیست. لذا لطفاً موارد در فایل ورد و پی‌دی‌اف قرار داده و مجدداً بارگذاری نمایید.

در فایل‌های ورد و پی‌دی‌اف فونت‌های متن فهرست مطالب قابل خواندن نیست و کلمات به هم چسبیده است، لطفاً آن را اصلاح نموده و فایل‌ها را مجدداً بارگذاری نمایید.

در فایل ورد فونت‌های متن فهرست مطالب قابل خواندن نیست لطفاً از فونت مناسب استفاده کرده و فایل ورد را مجدداً بارگذاری نمایید.

در فایل پی‌دی‌اف فونت‌های متن فهرست مطالب قابل خواندن نیست لطفاً از فونت مناسب استفاده کرده

و در فایل‌های ورد و پی دی اف صفحه عنوان انگلیسی، چکیده فارسی را قرار داده و مجدداً بارگذاری نمائید.

فهرست مطالب دارای شماره صفحه نمی باشد. طبق سرعنوان‌ها شماره صفحه در فهرست مطالب لطفاً اصلاح شود و سپس فایل‌های ورد و پی دی اف را مجدداً بارگذاری نمائید.

شماره صفحات و سرعنوان‌ها مندرج در فهرست مطالب به هم ریخته و نامشخص می‌باشد، لطفاً اصلاح شود و سپس فایل‌های ورد و پی دی اف را مجدداً بارگذاری نمائید.

در فایل‌های ورد و پی دی اف فهرست مطالب شماره صفحه ندارد، لطفاً شماره صفحات را مطابق با سرعنوان قرار دهید. همچنین صفحه در فایل‌های ورد و پی دی اف عنوان انگلیسی و چکیده انگلیسی را قرار داده و مجدداً بارگذاری نمائید.

در فایل‌های ورد و پی دی اف فهرست مطالب دارای شماره صفحه نیست، لطفاً شماره هر سرفصل و سر عنوان را در فهرست مطالب مشخص کرده، مجدداً بارگذاری نمائید.

در فایل‌های ورد و پی دی اف شماره صفحات مندرج در فهرست مطالب با شماره صفحات پایان نامه منطبق نیست. لطفاً این مورد تصحیح شود. در ضمن صفحه عنوان انگلیسی و چکیده انگلیسی را نیز در فایل‌های ورد و پی دی اف قرار داده و مجدداً بارگذاری نمائید.

صفحات فهرست مطالب ناقص است. لطفاً در فایل‌های ورد و پی دی اف صفحات فهرست مطالب را اصلاح نموده و مجدداً بارگذاری نمائید.

Error! Bookmark not defined از قسمتهای فهرست مطالب دارای شماره صفحه نمی باشد و را نشان می دهد، لطفاً اصلاح شود، سپس فایل‌های ورد و پی دی اف را مجدداً بارگذاری نمائید.

در فایل ورد کلمات فهرست مطالب بهم چسبیده است و قابل خواندن نیست لطفاً تصحیح شود و مجدداً بارگذاری گردد.

در صفحه فهرست مطالب، پیوست ذکر شده است اما در بخش پیوست هیچ مطلبی وجود ندارد، لذا بخش پیوست را کامل نموده و فایل‌های ورد و پی دی اف را مجدداً بارگذاری نمائید.

عنوان پایان نامه مندرج در فایل پی دی اف با عنوان مندرج در فایل ورد مطابقت ندارد. لطفاً عنوان تصحیح شود و مجدداً فایل ورد و پی دی اف بارگذاری گردد.

در طرح پژوهشی حاضر بنا داریم تا با تکیه بر مدل‌های شناخته شده کیفیت داده هریک از فیلدهای فراداده در سامانه ثبت را مورد بررسی قرار داده و طرح یکپارچه‌ای برای کنترل‌های کیفیت فراداده‌ها در سامانه ثبت پایان‌نامه / رساله‌های دانشجویان تحصیلات تکمیلی ایجاد نماییم.

۳-۵ کارگروه بهبود و نقش آن در شناسایی مشکلات سامانه گنج

در این بخش به معرفی مشکلات کیفی مشاهده شده در سامانه گنج و نقش کارگروه تشکیل شده در این حوزه خواهیم پرداخت. (ارشادی و احترامی ۱۳۹۵) کارگروه ویرایش اطلاعات سامانه گنج نه نشست پیوسته برگزار کرده است. اعضای کارگروه ۱۹ ساعت و ۳۰ دقیقه به بحث پیرامون مسائل و مشکلات سامانه گنج پرداخته‌اند. بر پایه بحث‌های کارشناسی مشکلات به ۴ دسته کلی محتوا، نمایش، نرم افزار، و سیاستگذاری تقسیم شده‌اند:

الف. محتوا: مشکلاتی مانند اشتباهات تایپی، اشتباه در ورود اطلاعات، و عدم تناظر بین فیلدها در دسته محتوا گنجانده می‌شوند. مشکلاتی که در دسته محتوا قرار می‌گیرند خود در گروه‌های کوچکتری قرار خواهند گرفت.

الف - ۱. داده‌های پایه: داده‌های پایه به آن دسته از اطلاعاتی اطلاق می‌شود که بر اساس دسته‌بندی منسجم در پشت پایگاه به کار رفته است. کاربرد این اطلاعات هنگام ورود داده است و بازیابی اطلاعات را تسهیل می‌نماید. برخی از مشکلات موجود در سامانه گنج به علت عدم وجود داده‌های پایه است. مانند عدم وجود داده‌های پایه استاندارد برای رشته‌ها و گرایش‌ها، عدم انتشار صحیح مستندات اسامی سازمان‌ها، عدم وجود داده‌های مستندات اسامی اشخاص، جغرافیایی، اصطلاحنامه‌های روزآمد و از این قبیل.

الف-۲. انتقال اطلاعات (تبدیل محتوا): عدم انتقال اطلاعات به صورت کامل و صحیح از پایگاه اطلاعاتی قدیم CDS/ISIS به سامانه گنج. این مشکلات خود به سه دسته جزئی‌تر تقسیم می‌شوند:

الف-۲-۱. مشکلاتی که در روند انتقال و تبدیل اطلاعات (تبدیل محتوا) از پایگاه قدیم به سامانه گنج رخ داده است، مانند اضافه شدن فیلد مجری به اطلاعات پایان نامه.

الف-۲-۲. آمیخته شدن اطلاعات نمایه سازی شده و اطلاعاتی که هنوز ویرایش بر روی آنها صورت نگرفته است. در پایگاه اطلاعاتی قدیم، قسمتی برای ورود اطلاعات توسط کاربران وجود داشته که ویرایشی نیز روی آنها صورت نگرفته است. پس از انتقال اطلاعات به سامانه گنج اطلاعات ویرایش نشده نیز به همراه اطلاعات صحیح و ویرایش شده منتقل شده اند و شماره بازیابی آنها نیز با یکدیگر ادغام شده است. مانند وجود عناوین تست و شسیب در پایگاه. الف-۲-۳. حذف شماره بازیابی مدارک (شماره TH) و شماره سریال مدارک در پایگاه قدیم و ادغام آنها با یکدیگر. شماره سریال مدارکی که در پایگاه اطلاعات قدیم ویرایش شده اند با شماره سریال اطلاعاتی که توسط کاربران ورود اطلاعات شده است، تفاوت داشته و اکنون شماره سریالها با یکدیگر ادغام شده اند. از طرفی شماره بازیابی مدارک (شماره TH) از پایگاه اطلاعاتی حذف شده است.

الف-۳. وجود اطلاعات ویرایش نشده در سامانه گنج: اطلاعاتی که در پایگاه ثبت، (sabt.irandoc.ac.ir) ثبت شده اند بدون ویرایش به سامانه گنج منتقل می شوند. برخی از مشکلات سامانه گنج مربوط به همین دسته از اطلاعات است. امکان ورود اطلاعات اشتباه توسط کاربران بسیار زیاد است. اشتباهاتی مانند رعایت نکردن اصول نگارشی، اشتباهات تایپی، ورود اطلاعات اشتباه در فیلدهای گوناگون.

ب. نمایش: این گروه از مشکلات مربوط به اطلاعاتی هستند که به صورت صحیح و کامل در سامانه گنج وجود دارند، اما به علت فرمت نمایش، هنگام جستجوی اطلاعات نمایش داده نمی شوند.

ج. نرم افزار: آن دسته از مشکلاتی هستند که به علت ساختار نرم افزار و یا موتور جستجو به چشم خورده اند. بررسی های اولیه نشان می دهد که اساساً در نرم افزار سامانه گنج قابلیت انتقال برخی از اطلاعات مدنظر قرار نگرفته است. مواردی مانند:

- عدم جستجوی فرمولها مانند فرمول های ریاضی و شیمی.
- عدم یکسانی کد کاراکترها با یکدیگر و نمایش نتایج جستجوی متفاوت.

- عدم انتقال اطلاعات ویرایش شده از سامانه ویرایش به سامانه گنج.
- عدم وجود برنامه و فرآیند مدون برای روزآمدسازی پایگاه.
- عدم وجود کنترل‌های لازم پس از روزآمدسازی پایگاه.
- عدم وجود استاندارد نمایش اطلاعات در پایگاه.

د.سیاست‌گذاری: آن دسته از مشکلاتی که بر اساس تصمیم‌گیری‌های گذشته در سامانه گنج مشاهده می‌شوند. بدین معنی که اطلاعات آنها در دیگر پایگاه‌ها به صورت کامل و صحیح وجود دارد و نرم‌افزار نیز امکان نمایش آنها را دارد اما به علت تصمیمات گذشته، نمایش داده نمی‌شوند. مشکلاتی مانند عدم نمایش کلیدواژه‌های لاتین.

بر اساس نظرات اعضای کارگروه برای بررسی مشکلات سامانه گنج، فهرستی از مشکلات تهیه شد و هر کدام از آنها به یکی از اعضای کارگروه برای بررسی محول شد. اعضای کارگروه مشکلات را بررسی نموده و گزارش ریشه‌های احتمالی مشکلات را به کارگروه ارائه کردند (جدول ۳-۱).

جدول ۳-۲. نتایج بررسی برخی از مشکلات موجود در سامانه گنج

شرح مشکل	ریشه مشکل
• یک رکورد دیده شده است که در آن عبارت "چکیده ندارد" برای رکوردی که چکیده دارد، به کار رفته است.	• فرمت نمایش وجود دو چکیده تمام نما و راهنما.
• عدم نمایش کامل کلید واژه موجود در کاربرگه ویرایش شده، در فیلد مرتبط.	• فرمت نمایش • انتقال دیتا (تبدیل محتوا) • اطلاعات اولیه • خطای نیروی انسانی • خطا در روش تبدیل محتوا • خطا در نرم افزار تبدیل محتوا • مدیریت و فرایند اطلاعات

• خطای نمایه سازی • خطای در تبدیل محتوا	عدم وجود عنوان برای برخی از مدارک در سامانه گنج (نمونه)
	وجود کاراکترهای علامت سوال در محتوای ارسالی از بخش ویرایش نهایی اطلاعات.
• سیستم ورود اطلاعات و نمایش اطلاعات بر می گردد.	وجود کاراکترهای نامرتبط مانند علامت سوال به جای کاراکترهایی که در کاربرگه ویرایش شده بدرستی درج شده است.
• خطا در ورود اطلاعات کاربر بیرونی • در ابتدای کار با نرم افزار جدید در بازه زمانی کوتاهی خطا در نمایه سازی رخ داده است.	وجود کلیدواژه در انتهای چکیده ها در برخی از رکوردها.
• انتقال اطلاعاتی که احتمالا به صورت آزمایشی در سی دی اس یا ثبت وارد شده و عینا به سامانه گنج منتقل شده است. • در حال حاضر برخی از اطلاعات ثبتی به صورت بدون کنترل وارد سامانه گنج می شوند. • ادغام اطلاعات کنترل شده و کنترل نشده.	وجود رکوردهایی که با جستجوی کلمات کلیدی نامفهوم و تصادفی مانند "شسیب" یافت می شوند.

بمنظور حل ریشه‌ای مشکلات مذکور، رویکردهای مبتنی بر تکنیک‌های تحلیل ریشه می‌تواند مورد استفاده قرار گرفته تا به تدریج و گام به گام به کمک این تکنیک‌ها مشکلات کیفی مذکور تحلیل شده و مرتفع گردند. فرایندهای معرفی شده در ابتدای این فصل در قالب تضمین کیفیت در صورتی که به درستی اجرا شوند می‌توانند به بهبود مستمر و در نتیجه حل ریشه‌ای مشکلات کیفی منجر شوند. در ادامه روش تحلیل ریشه که در قالب کارگروه بهبود مورد استفاده قرار گرفته است مورد اشاره قرار می‌گیرد.

بطور کلی اگر سه دسته اصلی مشکلات کیفی مطرح شده در ابتدای این بخش را در نظر بگیریم مشکلات مربوط به بخش الف یعنی محتوا با توجه به این که به تدریج شناسایی می‌شوند بایستی به کمک فرایند اقدامات اصلاحی و ممیزی سیستمی توانایی سازمان در شناسایی آنها را افزایش داده و به کمک

رویکردهای تحلیل ریشه و مراحل فرایند اقدامات اصلاحی به رفع ریشه‌ای آنها پرداخت. در خصوص مشکلات مربوط بخش‌های ب و ج یعنی نمایش و نرم افزار مدل ۷ مرحله‌ای روش اجرایی بازننگری طراحی و توسعه سامانه گنج می‌تواند اقدامات ریشه‌ای بمنظور پیشگیری از بروز مشکلات در آینده ارائه دهد. (ارشادی و احترامی ۱۳۹۵)

۳-۶ شناسایی و دسته‌بندی مشکلات سامانه گنج

برخی از اشکالات موجود در سامانه گنج بر اساس بازخوردهای جمع آوری شده از بخشهای مختلف فراهم آوری، سازماندهی، و اشاعه و نیز برخی از نظرات کاربران را می‌توان به صورت جدول زیر دسته بندی نمود. بر اساس این فهرست ۴۱ گونه مشکل شناسایی شده، مورد توافق اعضای کارگروه بوده است.

جدول ۳-۳. فهرست مشکلات موجود در سامانه گنج

ردیف	شرح مشکل
۱	عبارت "چکیده ندارد" برای رکوردی که چکیده دارد، به کار رفته است.
۲	عدم نمایش کامل کلید واژه موجود در کاربرگه ویرایش شده، در فیلد مرتبط.
۳	عدم وجود عناوین برخی از مدارک ویرایش شده در سامانه گنج و عدم اطلاع از تخمین تعداد چنین مواردی.
۴	وجود کاراکترهای علامت سوال در محتوای ارسالی از بخش ویرایش نهایی اطلاعات.
۵	وجود کاراکترهای نامرتب مانند علامت سوال به جای کاراکترهایی که در کاربرگه ویرایش شده بدرستی درج شده است.
۶	وجود کلیدواژه در انتهای چکیده‌ها در برخی از رکوردها.
۷	وجود رکوردهایی که با جستجوی کلمات کلیدی نامفهوم و تصادفی مانند "شسیب" یافت می‌شوند.
۸	وجود عبارتهای نادرست مانند "این عنوان فرضی است" در فیلد عنوان برخی از رکوردهایی که در سایر فیلدهایش اطلاعات مدرک شناختی وجود دارد.
۹	وجود اشکال در تاریخ انتشار مدرک. (برای مثال نمایش تاریخ صفر برای بیش از دویست هزار رکورد).
۱۰	درج کارکترهای نا مرتبط و بعضا ناشناس مانند صفر، خط تیره و ... در برخی از فیلدها، مانند فیلد تاریخ در بسیاری از رکوردها.
۱۱	وجود کلید واژه‌هایی که معادل انگلیسی ندارند، (از تاریخ ۱۳۷۴ به بعد).
۱۲	وجود رکوردهایی که علیرغم داشتن کلیدواژه‌های لاتین در کاربرگه قدیم، نمایش داده نمی‌شود.
۱۳	وجود رکوردهایی با اطلاعات کتابشناختی نادرست (مثال عدم تناظر بین رشته اصلی و فرعی، بین دانشگاه و

	موضوع).
۱۴	درج عبارت "ندارد" در فیلدهای مرتبط با اطلاعات کتابشناختی برخی از رکوردهای پایان نامه‌ها.
۱۵	عدم درج نام و یا درج نام ناصحیح برخی از موسسات، پژوهشکده‌ها، دانشکده‌ها و دانشگاه‌ها در فیلدهای مرتبط.
۱۶	عدم امکان درج و نمایش صحیح کاراکترهای فرمولی در عناوین، مانند فرمولهای ریاضی، شیمی و حروف یونانی مانند $\eta, \epsilon, \theta, \beta, \mu, \lambda, \phi$ ، اعداد و ...
۱۷	تکراری بودن شمار قابل توجهی از رکوردها.
۱۸	نبود برخی از صفحات تمام متن.
۱۹	لینک اشتباه به صفحات تمام متن.
ردیف	شرح مشکل
۲۰	عدم نمایش صفحات تمام متن (نمایش پیغام ۴۰۴)
۲۱	ناخوانا بودن برخی از عناوین پایان نامه‌ها در فیلد عنوان.
۲۲	نمایه سازی بر اساس عنوان مدرک که بازنمایی مناسبی از اطلاعات مدرک را بدست نمی‌دهد.
۲۳	نبود عنوان در برخی از رکوردهای پایان نامه‌های لاتین
۲۴	جستجو نشدن کاراکترهای مانند :، ؛ ، "" «» ؟ ! * و....
۲۵	نتایج مرتب سازی بر اساس عنوان به ترتیب الفبایی نمایش داده نمی شود.
۲۶	نام برخی از نویسندگان مثلا در پایان نامه‌هایی که دارای بیش از یک نویسنده هستند نمایش داده نمی شود.
۲۷	امکان ثبت بیش از یک نویسنده برای پایان نامه‌ها وجود ندارد.
۲۸	درج اسامی ترکیبات شیمیایی بصورت ناقص و نادرست به عنوان کلید واژه.
۲۹	ناهمسانی جمع و مفردها و جمع یا مفرد به کار بردن عام و خاص در بسیاری از کلید واژه‌ها.
۳۰	ناهمسانی نگارشی و رسم الخطی در موارد بسیاری مانند: ئیدروکربن و ییدروکربن و هیدروکربن، شیمیایی و شیمیائی، آلومینیم و آلمینیم، متانول و متانل، متیل و میتل، کینون و کوینون، برمید و برماید، برومو و برمو و
۳۱	تمایز قائل نشدن بین حروف بزرگ و کوچک لاتین که متضمن معانی متفاوت‌اند در برخی از رکوردها مانند: CdS (کادمیم سولفید) و CDS (نام یک نرم افزار و یا یک سرواژه)، CO (کربن مونوکسید) و Co (کبالت) و ...
۳۲	جایجایی ممیزها در برخی از گزارشات کمی، مانند گزارش میزان Ph. 14.3 به جای 3.14.
۳۳	یکسان نبودن استفاده از فاصله و نیم فاصله در بسیاری از کلید واژه های مرکب.
۳۴	یکسان نبودن در نگارش کوتاه‌نوشته‌ها و سرواژه‌ها در هر دو حالت فارسی و لاتین.
۳۵	یکسان نبودن اسامی اشخاص و اعلام و نگارش اسامی یک فرد به صور مختلف.
۳۶	در مرتب سازی بر اساس ربط، واژه‌های ترکیبی مدنظر قرار نمی گیرد. (مثال حسن کریمی)
۳۷	برای کاربرانی که یک بار جستجو نموده و نیاز به بازبایی مجدد دارند، کد بازبایی بهترین گزینه می‌باشد، در حالی که هم اکنون درج و جستجو بر اساس شماره بازبایی در لیست عناوین امکان پذیر نمی باشد

۳۸	چیزی که به عنوان کد نوشتار نوشته می شود منحصر به فرد(یونیک) نیست.
۳۹	تطبیق دادن شماره یو.آر.ال (شماره عنوان) با شماره متن مشکل ساز است، بازگشت روند کد واحد شماره بازیابی th پیشنهاد می شود.
۴۰	اگر حروف فارسی را با کدهای مختلف جستجو کنیم نتایج مختلفی می گیریم.
۴۱	جستجوی لاتین ضعیف عمل می کند.

سرمنشا و ریشه مشکلات سامانه گنج دلایل گوناگونی دارد که می توان آنها را به صورت زیر دسته بندی نمود:

۱. اولین و مهم ترین علت بروز چنین مشکلاتی را می توان، وارد شدن اطلاعات پایگاه ثبت به سامانه گنج بدون کنترل و ویرایش دانست. در حال حاضر پس از ثبت اطلاعات توسط دانشجو در پایگاه ثبت، کارشناسان دانشگاه می بایست اطلاعات وی را بر اساس کد رهگیری بازیابی نمایند و در صورت صحیح بودن اطلاعات را تایید کنند. البته کارشناسان دانشگاه به علت حجم زیاد امور، اطلاعات را به صورت کامل بررسی نکرده و اکثراً با اشتباهات واضح اطلاعات را تایید می کنند. از طرفی مرکز اطلاعات بصورت دوره ای از اطلاعات تایید شده توسط دانشگاه ها خروجی تهیه کرده و بدون ویرایش و نمایه سازی آنها را به سامانه گنج اضافه می کند. عدم کنترل اطلاعات باعث شده است که مشکلاتی که در دسته محتوا بیان شده اند، بروز کند. مشکلاتی مانند اشتباهات تایپی، وارد نکردن کلیدواژه ها بصورت صحیح، وارد شدن کلیدواژه در فیلد چکیده، وارد شدن چکیده در فیلد عنوان، عدم تناظر فیلد رشته و گرایش، دانشگاه و دانشکده و از این قبیل مشکلات.

۲. دومین دلیل مهم برای بروز اشکالات در سامانه گنج، تبدیل اطلاعات در گذشته بوده است. اطلاعات موجود در پایگاه قدیم به علت وجود چندین مرحله کنترل، کامل و صحیح است. همچنین این اطلاعات بصورت نسخه چاپی موجود هستند. اما در گذشته به دلیل یکسان نبودن ساختار سامانه گنج و پایگاه های اطلاعاتی قدیم کارشناسان مجبور به تبدیل اطلاعات شده اند، از اینرو در جابه جایی ها امکان بروز اشتباه و یا از بین رفتن اطلاعات بسیار زیاد است.

۳. سومین دلیل اشکالات در سامانه گنج، وجود سیاست هایی بوده است که در گذشته تصمیم گیری شده اند و اکنون نیز اسنادی برای چرایی آنها وجود ندارد. سیاست هایی مانند عدم نمایش کلیدواژه های لاتین، عدم نمایش مشخصات بیش از یک نویسنده برای پایان نامه هایی که بیش از یک نویسنده دارند. نمایش دو فیلد برای چکیده

شامل چکیده راهنما و چکیده ترویجی که هر دو فیلد بصورت یک فیلد با عنوان چکیده در سامانه گنج نمایش داده می شوند.

۴. چهارمین مشکل در سامانه گنج به علت عدم وجود محتوای ثابت و لیست های یکپارچه از اطلاعات است. بدین معنی که فهرست منسجم و مستند از اسامی سازمان ها، رشته ها و گرایش های، شهرها و استان ها، اسامی افراد و... وجود ندارد. کاربران در پایگاه ثبت این دسته از اطلاعات را تایپ می کنند در نتیجه یک رشته خاص به چندین صورت گوناگون در پایگاه ذخیره شده و به همان صورت نیز در سامانه گنج نمایش داده می شود. برای نمونه می توان واژه رضایت شغلی را در سامانه گنج جستجو کرد. در نتایج جستجو مشاهده می کنیم چندین نوع رضایت شغلی نمایش داده می شود.

۵. پنجمین مشکل در سامانه گنج، نبود استاندارد برای مقادیر اطلاعاتی است. به عنوان مثال برای فیلد تاریخ هیچ گونه استاندارد در مرکز اطلاعات و مدارک علمی ایران تدوین نشده است. و یا استاندارد برای نمایش اطلاعات اساتید راهنما و مشاور وجود ندارد. کلیه فیلدهای اطلاعاتی سامانه گنج بدون داشتن استاندارد و بر اساس تصمیمات و سیاست های اتخاذ شده در گذشته و بدون داشتن مستندات طراحی شده اند.

چند نکته کلیدی و بسیار مهم را نباید از نظر دور داشت:

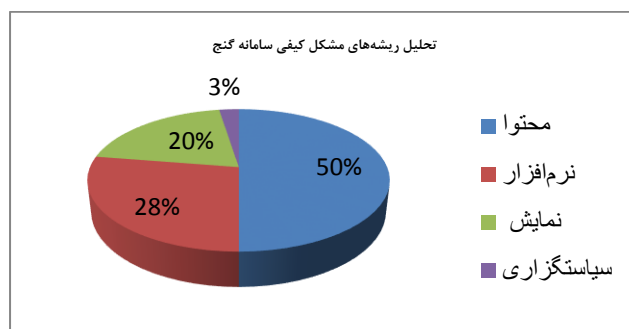
الف - نمایه سازی تا حد ممکن باید توسط نمایه سازان تحصیل کرده در رشته های تحصیلی مرتبط و نزدیک به موضوع مدارک انجام شود تا علاوه بر کاهش خطاهای صوری، مدارک نمایه شده از استانداردهای لازم محتوایی نیز برخوردار باشند.

ب - علاوه بر ویراستاری، سیستم کنترل کیفیت برای فرایند نمایه سازی در بخش سازماندهی اطلاعات باید طراحی و اجرا شود.

ج - لازم است منابع مرجع مناسب و روز آمد در اختیار نمایه سازان قرار گیرد.

د - فرایندهای موجود در بخش تحقیق و توسعه در ارتباط با سامانه گنج و توسعه آن می بایست مورد ارزیابی دقیق علمی و کارشناسی قرار گیرند. در این خصوص لازم است به کمک روش اجرایی ممیزی سیستم کیفیت چارچوبها و گردش کارهای ارائه شده در حوزه سامانه گنج را به صورت مداوم مورد ممیزی قرار داد و مغایرتهای مشاهده شده را در قالب فرایند اقدامات اصلاحی رفع نمود.

نتایج تجزیه تحلیل مشکلات کیفی سامانه گنج در شکل ۳-۳ قابل مشاهده است. همانطور که در این شکل مشخص است



شکل ۳-۳. تحلیل مشکلات کیفی سامانه گنج

همانطور که در این شکل ملاحظه می گردد مشکل محتوای سامانه بیشترین درصد را به خود اختصاص داده است. نرم افزار و نمایش به ترتیب با ۲۸ و ۲۰ درصد ریشه‌های دیگر مشکلات کیفی را تشکیل می دهند. بمنظور ایجاد تناظر میان اطلاعات جدول ۳-۳ و شکل ۳-۳ در ادامه شماره ردیف‌های مشکلات کیفی درج شده در جدول ۳-۳ به ازای هر یک از چهار دسته ذکر شده در شکل ۳-۳ اشاره شده‌اند.

شماره ردیف‌های مربوط به مشکل محتوا عبارتند از ۱، ۳، ۶، ۷، ۸، ۹، ۱۰، ۱۱، ۱۳، ۱۴، ۱۵، ۱۸، ۲۹، ۳۰، ۳۱، ۳۳ و ۳۴، ۳۹، ۴۰ و ۴۱.

شماره ردیف‌های مربوط به مشکل نرم‌افزار عبارتند از ۱۷، ۱۹، ۲۰، ۲۱، ۲۳، ۲۴، ۲۵، ۲۷، ۲۶، ۳۷ و ۳۸.

شماره ردیف‌های مربوط به مشکل نمایش عبارتند از ۲، ۴، ۵، ۱۲، ۱۶، ۲۸ و ۳۲.

شماره ردیف‌های مربوط به مشکل سیاست‌گذاری عبارتند از ۲۲ و ۲۶.

لازم بذکراست، در شکل ۳-۳ یک دسته‌بندی اولیه از نوع مشکلات کیفی ارائه شده است. تعیین فراوانی هر مشکل کار بسیار سختی است که به زیرساختارهای نرم افزاری بهتری نیاز دارد. در فاز دوم پروژه حاضر پیشنهاداتی در خصوص بهبود زیرساختارهای نرم افزاری بگونه ای که بتوان فراوانی مشکلات کیفی و ریشه های آن را دقیقتر تحلیل کرد ارائه خواهد شد.

فصل چهارم – تعیین وضعیت مطلوب اقلام اطلاعاتی سامانه ثبت

۴-۱- مقدمه

در این بخش به شناسایی و معرفی فیلدهای مختلف سامانه ثبت خواهیم پرداخت. بدین منظور اقلام اطلاعاتی این سامانه به صورت جداگانه معرفی خواهد گردید. همچنین مدل کیفیت فراداده انتخاب شده ارائه خواهد شد. از سوی دیگر هریک از اقلام داده‌ای سامانه ثبت مطابق ابعاد کیفی فراداده‌ها در مدل انتخابی مورد بررسی قرار خواهد گرفت. در فاز دوم پروژه (بخش پنجم) به کمک کارگروه بهبود سامانه ثبت ابعاد کیفی معرفی شده تصدیق خواهد شد و چک‌لیست کنترل کیفی سامانه ثبت توسعه داده خواهد شد.

۴-۲- معرفی فیلدهای سامانه ثبت

همانگونه که در انتهای فصل سوم تبیین گردید در سامانه ثبت پایان نامه/ رساله دانشجویان تحصیلات تکمیلی اقلام اطلاعاتی متفاوتی از دانشجویان اخذ می گردد. در این فصل بنا داریم ابعاد مدل کیفیت داده را روی اقلام مختلف فراداده پایان نامه ها و رساله ها تصویر نماییم.

به این منظور در ابتدا به معرفی اقلام اطلاعاتی قابل ثبت در سامانه ثبت خواهیم پرداخت. در ادامه و در جدول زیر به معرفی این اقلام در قالب شناسنامه فیلدهای این سامانه خواهیم پرداخت.

جدول ۴- ۱. شناسنامه فیلدهای فراداده های ثبت پایان نامه ها/ رساله ها

شناسنامه فیلدهای فراداده ی ثبت پایان نامه ها/ رساله ها			
نام فیلد	نام فیلد	ویژگی فیلد	توضیحات
نام	نام	صرفا حروف	مقداری ثابت و تنها قابل تغییر با مراجعه به پروفایل کاربری
نام خانوادگی	نام خانوادگی	صرفا حروف	مقداری ثابت و تنها قابل تغییر با مراجعه به پروفایل کاربری
رایانامه	رایانامه		مقداری ثابت و غیرقابل تغییر
شماره ملی	شماره ملی	صرفا عدد	مقداری ثابت و غیرقابل تغییر
شماره دانشجویی	شماره دانشجویی	صرفا عدد	نباید کمتر از ۴ و بیشتر از ۱۰ رقم باشد
مقطع تحصیلی	مقطع تحصیلی	وجود combo box	دارای قابلیت جستجو
وابستگی سازمانی	وابستگی سازمانی	وجود combo box	دارای ارتباط Intra-relation با فیلد بعدی خود با قابلیت جستجو

نام فیلد	نام فیلد	ویژگی فیلد	توضیحات
موسسه آموزشی/پژوهشی	موسسه آموزشی/پژوهشی	وجود combo box	Intra- دارای ارتباط relation با فیلد بعدی خود با قابلیت جستجو
دانشکده/ پژوهشکده	دانشکده/ پژوهشکده	وجود combo box	
زبان اصلی پایان نامه/ رساله	زبان اصلی پایان نامه/ رساله	وجود combo box	با قابلیت جستجو
عنوان (فارسی)	عنوان (فارسی)		نباید کمتر از ۱۰ و بیشتر از ۵۰۰ حرف باشد
عنوان (انگلیسی)	عنوان (انگلیسی)		نباید کمتر از ۱۰ و بیشتر از ۵۰۰ حرف باشد
گروه تحصیلی	گروه تحصیلی	وجود combo box	Intra- دارای ارتباط relation با فیلد بعدی خود با قابلیت جستجو
رشته	رشته	وجود combo box	Intra- دارای ارتباط relation با فیلد بعدی خود با قابلیت جستجو
گرایش	گرایش	وجود combo box	با قابلیت جستجو
تاریخ دفاع	تاریخ دفاع	وجود combo box	به فرمت تاریخ شمسی با قابلیت جستجو
نام	نام خانوادگی نام خانوادگی نام خانوادگی (انگلیسی) نام خانوادگی (انگلیسی) رایانامه شماره تلفن همراه شماره ملی	*	صرفا حروف
نام خانوادگی		*	صرفا حروف
نام خانوادگی (انگلیسی)			صرفا حروف انگلیسی
نام خانوادگی (انگلیسی)			صرفا حروف انگلیسی
رایانامه		*	
شماره تلفن همراه			صرفا عدد
شماره ملی			صرفا عدد
نام فیلد	نام فیلد	ویژگی فیلد	توضیحات
چکیده (فارسی)	چکیده (فارسی)		نباید کمتر از ۲۰ و بیشتر

از ۳۰۰۰ حرف باشد			
نباید کمتر از ۲۰ و بیشتر از ۳۰۰۰ حرف باشد		چکیده(انگلیسی)	چکیده(انگلیسی)
قابل انتخاب از بین گزینه های موجود		نوع پایان نامه/ رساله	نوع پایان نامه/ رساله
نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد		فهرست مطالب	فهرست مطالب
نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد		فهرست منابع فارسی	فهرست منابع فارسی
نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد		فهرست منابع غیر فارسی	فهرست منابع غیر فارسی
حداقل تعداد کلید واژه ها به فارسی و انگلیسی برای هر زبان دو مورد است با حداکثر سه مورد		کلید واژه ها	کلید واژه ها
قالب فایل pdf مجاز است		فایل pdf پایان نامه/ رساله	فایل pdf پایان نامه/ رساله
قالب فایل docx, doc مجاز است		فایل word پایان نامه/ رساله	فایل word پایان نامه/ رساله
		دیگر فایل های پایان نامه/ رساله	دیگر فایل های پایان نامه/ رساله

در ویرایش جدید سامانه ثبت فیلهای سامانه به خوبی طراحی شده‌اند و مشکلات کیفی قبلی تا حد زیادی با روش‌های مختلف مرتفع شده‌اند. با این وجود می‌توان برخی نکات زیر را در طراحی‌های آتی مورد توجه قرار داد.

- در قسمت پروفایل کاربری برای ورود در فیلهای نام و نام خانوادگی و نام پدر وارد کردن کاراکترهای ؟، /، (،)، % و ... مجاز است.
- در قسمت نام "فیلد شماره دانشجویی" وارد کردن مقادیر حروف و کاراکترهایی مانند ؟، /، (،)، %، \$ و ... مجاز است.

- در قسمت استاد راهنما / مشاور، فیلدهای نام و نام خانوادگی به صورت فارسی و انگلیسی وارد کردن کارکترهای ؟، /، % و ... مجاز است.
- در فیلد چکیده (فارسی) قابلیت ورود حروف انگلیسی وجود دارد.
- در فیلد چکیده (انگلیسی) قابلیت ورود حروف فارسی وجود ندارد.
- برای انتخاب فایل پایان نامه در فیلد "فایل word پایان نامه/ رساله" این امکان برای دانشجو وجود دارد که اگر پایان نامه/رساله خود را با نرم افزار دیگری تهیه کرده باشد، میتواند فهرست مطالب و همچنین فهرست منابع خود را در این قسمت بارگذاری کرده و دیگر بخش ها را در فیلد "دیگر فایل های پایان نامه/ رساله" بارگذاری کند.

لازم بذکرست موارد بالا برخی نکات است که در بررسی اولیه بدست آمده است. در انتهای طرح پژوهشی حاضر و بررسی جایگاه ابعاد مختلف فراداده‌ها بر روی فیلدهای سامانه ثبت می‌توان به الگوی جامعی در خصوص نحوه کنترل کیفی فراداده‌های ثبت شده در سامانه دست یافت. در عین حال که به روش‌ها و اصولی که بتوان از بروز مشکلات کیفی پیشگیری نمود نیز می‌توان تا حد مطلوبی نائل آمد.

در ادامه به معرفی مدل کیفیت فراداده انتخاب شده و ابعاد مختلف آن بمنظور نگاشت روی اقلام داده‌ای پایان‌نامه/ رساله‌ها خواهیم پرداخت.

۴-۳- معرفی مدل کیفیت فراداده انتخابی

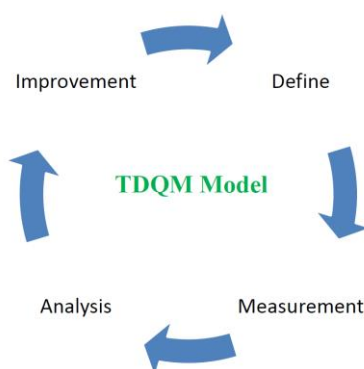
همانگونه که در فصل اول این گزارش به تفصیل اشاره گردید مدل‌ها و چارچوب‌های مختلفی برای کیفیت داده در متون علمی وجود دارد. یکی از مهم‌ترین این چارچوب‌ها که در مقالات مختلفی مورد توجه قرار گرفته و در فصل اول نیز به آن اشاره گردید مدل مدیریت کیفیت داده است.

نخستین پژوهش‌ها در حوزه مدیریت کیفیت جامع داده‌ها^۱ به دهه ۱۹۸۰ میلادی برمی‌گردد. تحقیقات اولیه در حوزه کیفیت داده‌ها به طور عمده بر توسعه تکنیک‌ها برای پرس و جو از منابع اطلاعاتی متعدد و ایجاد مخازن اطلاعاتی بزرگ متمرکز بود. تحقیق وانگ و مادنیک (۱۹۸۹) رویکرد سیستماتیک را برای بررسی کیفیت داده‌های مرتبط بکار می‌گیرد (Wang, Madnick 1989). این کار بیشتر هنگام تفکیک هویت اطلاعات و یکپارچه سازی اطلاعات از منابع مختلف و با حضور رکوردهایی که تداخل و همپوشانی دارند، شناخته شده و برجسته است. این محققان برای تعیین این که آیا رکوردهای جدا شده واقعا هویت‌های مرتبطی دارند، روشهایی را یافتند. این مسئله با واژه‌هایی مانند اتصال رکورد، غیرتکراری کردن رکورد و تطبیق رکورد، شناخته می‌شود. پس از آن، وانگ و مادنیک (۱۹۹۰) یک مدل پلی‌ژن^۲ (پلی برای متعدد و ژن برای منبع) را توسعه دادند (Wang, Madnick 1990) تا پردازش برجسبهای منبع داده را در پردازنده پرس و جو زیر نظر بگیرند و به این طریق جوابی برای پرسشهای مرتبط با کیفیت داده‌ها بیابند، پرسشهایی مانند: "این داده‌ها از کجا آمده‌اند؟" و "کدام منابع داده‌ای میانجی برای رسیدن به این داده‌ها مورد استفاده قرار گرفته است؟" در ادامه پژوهش، توسعه یک روش مدل سازی (مدل رابطه هویت کیفیت) مد نظر بوده است تا در مرحله طراحی پایگاه داده مفهومی (Wang et al. 1993) به جمع آوری روشمند معیارهای جامع کیفیت داده‌ها به عنوان فراداده پردازد. (Storey and Wang 1998) در این روش پردازنده پرس و جو با استفاده از جبر رابطه‌ای بسط یافته، فراداده کیفیت داده‌های سلسله مراتبی را پردازش می‌کند (Wang et al. 1993) این جریان پژوهشی، بر روی تحقیق و طراحی پایگاه داده مدرن، مانند منشا و تبار داده‌ها که بانمن و همکاران در سال ۲۰۰۱ ارائه نمودند (Buneman et al. 2001) و همچنین بر روی سایر زمینه‌های مرتبط با جبر رابطه‌ای برای امنیت و مدیریت حفظ حریم خصوصی داده‌ها، اثرگذار بوده است. تلاشهای تحقیقاتی اولیه، محققان را برای آغاز تحقیق روشمند در تمام گستره مسائل کیفیت داده‌ها تشویق کرد که به نوبه خود منجر به افتتاح برنامه مدیریت کیفیت داده جامع در اوایل ۱۹۹۰ در دانشگاه MIT شد. تحقیق اولیه در برنامه توسعه چارچوب مدیریت کیفیت داده‌بهبود کیفیت مستمر داده‌ها را با دنبال کردن مراحل تعریف،

^۱ Total Data Quality Management

^۲ Polygen

اندازه‌گیری، تجزیه و تحلیل و بهبود حمایت می‌کند (Madnick, Wang 1992). این مدل، چهارچوب مدیریت کیفیت جامع^۱ (TQM) را برای بهبود کیفیت در حوزه تولید به حوزه داده توسعه می‌دهد. (شکل ۴-۱)



شکل ۴-۱. چهار مرحله اصلی در مدیریت کیفیت جامع داده

دمینگ در سال ۱۹۸۲ (Wang, Strong 1996) و جوران و جوفری در سال ۱۹۹۹ از مهمترین پایه‌گذاران و توسعه‌دهندگان حوزه مدیریت کیفیت جامع بودند. بینش کلیدی این بود که اگر چه داده، در واقع، یک محصول تولید شده (یا محصول جانبی) توسط بسیاری از سازمان‌ها است، ولی آن طور که باید با آن رفتار نشده و مورد مطالعه قرار نگرفته است. در ادامه به پژوهش‌های صورت گرفته در حوزه نظریه‌ها، روش‌ها، و تکنیک‌ها برای چهار مرحله از چارچوب مدیریت کیفیت داده توسعه خواهیم پرداخت.

• تعریف مراحل

تعریف کیفیت داده‌ها از نقطه نظر مصرف‌کننده، دستیابی به یک موفقیت بزرگ بود. مطابق با این تعریف تناسب بین استفاده و شناسایی ابعاد کیفیت داده‌ها از طریق یک مطالعه پیمایشی چند مرحله‌ای روشمند عملی شد (Wang, Strong 1996). ابعاد کلیدی کیفیت داده‌ها با استفاده از تحلیل عاملی بر روی بیش از ۱۰۰

^۱ Total Quality Management

نشانه مشخص کیفیت داده‌ها از طریق مطالعه پیمایشی کشف شد. این ابعاد به صورت چهار طبقه کیفیت داده یعنی دسترس پذیری، محتوایی، نمایشگری و ذاتی سازمان‌دهی شده‌اند. در ادامه و در زیربخش بعدی به مرحله بعدی از مدل مدیریت کیفیت داده یعنی اندازه‌گیری اشاره خواهیم کرد.

• اندازه‌گیری

لی و همکاران در سال ۲۰۰۲ دومین مرحله از رویکرد مدیریت کیفیت جامع یعنی ارزیابی و اندازه‌گیری کیفیت داده‌ها را مشابه شرایط عملیاتی به منظور اندازه‌گیری کیفیت داده‌ها در سازمانها توسعه دادند. (Lee et al. 2002) این ابزار هر یک از ابعاد معرفی شده در انتهای بخش قبل را به چهار پنج شاخص قابل اندازه‌گیری تقسیم می‌دهد و صورتهای کاربردی مناسب برای این آیتم‌ها بکار گرفته می‌شوند تا هر یک از ابعاد را ارزیابی کنند. برای مطالعه بیشتر تحقیق صورت گرفته توسط پپینو و همکاران در سال ۲۰۰۲ را ملاحظه کنید. (Pipino, Yang and Wang 2002) این ابزار همچنین این قابلیت را دارد که با نیازهای خاص هر سازمانی منطبق گردد.

• تجزیه و تحلیل

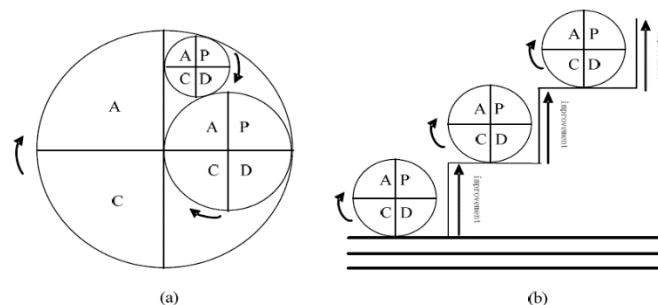
نتایج اندازه‌گیری در این مرحله تفسیر می‌گردد. تکنیک‌های تجزیه و تحلیل شکاف (Lee et al. 2002) تفاوت‌های ادراکی بین ابعاد داده‌ها و نقش‌های از قبل تعیین شده در حوزه مربوط به کیفیت داده را نشان می‌دهد (Strong, Lee and Wang 1997). لی و استرانگ در سال ۲۰۰۴ سه نقش عمده جمع‌آوری کننده‌های اطلاعات، متولیان داده‌ها، و مصرف کنندگان داده را در این حوزه معرفی نمودند (Lee, Strong 2004). تجزیه و تحلیل‌های صورت گرفته در این مرحله، ابعادی را که بیشتر نیاز به بهبود دارند

شناسایی می‌کند و به تحلیل ریشه مشکلات کیفیت داده می‌پردازد. در زیربخش بعدی به معرفی مرحله آخر یعنی بهبود خواهیم پرداخت.

• بهبود

در این مرحله، اقدامات صورت گرفته هم در راستای تغییر ارزش داده‌ها و هم بمنظور تغییر فرآیندهایی است که داده تولید می‌کنند. روش دوم همان طور که در مطالعات بالو و همکاران (Ballou et al. 1998) و وانگ و همکاران (Wang et al. 1998) در سال ۱۹۹۸ مورد بحث قرار داده‌اند موثرتر است. در این روش مراحل به سوی مدیریت اطلاعات به عنوان یک محصول فراهم شده است. علاوه بر این، فناوری‌های قبلی مانند پلی‌ژن و مدل رابطه هویت کیفیت، به عنوان بخشی از فرآیند بهبود مستمر اعمال می‌شود.

مدل مدیریت کیفیت داده به لحاظ ساختاری و مدیریتی تشابه زیادی به مدل بهبود مستمر دمنینگ دارد که در شکل زیر بصورت شماتیک به آن اشاره شده است.



شکل ۴-۲. معرفی چرخه دمنینگ

در هر ۴ مرحله بالا یعنی: برنامه ریزی^۱، اجرا^۲، کنترل^۳ و اقدام^۴ لازم است سازمان اقدامات لازم برای تضمین کیفیت خدمات خود و کاهش خطا در فرآیندهای تولید داده و اطلاعات را انجام دهد. همانگونه که در

^۱. Plan

^۲. Do

^۳. Check

^۴. Act

شکل ۳ مشاهده می‌نمایید با کاربرد مداوم چرخه دمیگ بهبود مرحله به مرحله در سامانه یا فرایند مورد کاربرد رخ خواهد داد.

علاوه بر توسعه و غنی سازی چارچوب مدیریت کیفیت داده، برنامه مدیریت کیفیت داده در جهت یکپارچه سازی حوزه کیفیت داده، تلاش های قابل توجهی را سبب شده است. همچنین باعث گسترش تأثیر پژوهش و ترویج همکاری دولت با دانشگاه و صنعت از طریق نشریات، سمینارها، دوره های آموزشی و کنفرانس های بین المللی سالانه در حوزه کیفیت اطلاعات (ICIQ^۱)، از سال ۱۹۹۶ به بعد شده است. دانشگاه آرکانزاس در شهر لیتل راک با کمک برنامه مدیریت کیفیت داده، برای اولین بار از نوع خود در ایالات متحده، دوره کارشناسی ارشد و دکتری کیفیت داده ایجاد کرده است. ایجاد این دوره ها به منظور پاسخگویی به افزایش تقاضا و نیاز به متخصصان کیفیت داده آموزش دیده در سطح بالا و نیز آماده سازی دانشجویان برای تحقیقات پیشرفته در زمینه کیفیت داده، بوده است. در بخش صنعت و در حوزه کیفیت داده ها، فروشندگان اصلی نرم افزار شروع به پیاده سازی فناوری های کیفیت داده در ارائه محصول و خدمات خود کرده اند. در بخش دولتی، کیفیت داده به یک جزء مهم در بسیاری از بخش های دولت الکترونیک و طرح های معماری سازمانی تبدیل شده است. در این میان، تحقیقات کیفیت داده ها با چالش های جدیدی مواجهه است که مربوط است به محیط های کسب و کار همیشه در حال تغییر، الزامات قانونی، انواع افزایش فرم های داده ها / رسانه ها، و فناوری های اینترنت و این که اساسا چگونه اطلاعات تولید، ذخیره، دستکاری، و مصرف می شود. تحقیقات کیفیت داده ها که از دو دهه پیش شروع شده، وارد عصر جدیدی شده است که در آن تعداد فزاینده ای از محققان به طور فعال به منظور افزایش درک از مشکلات کیفیت داده و توسعه راه حل برای مسائل نوظهور در این حوزه تلاش می کنند.

وانگ و همکاران (Wang et al. 1995) چارچوب مدونی بر مبنای الزامات استاندارد ISO 9001 بمنظور دسته بندی تحقیقات صورت گرفته روی مبحث کیفیت داده ارائه دادند. چارچوب دسته بندی تحقیقات توسط آنها شامل گروهبندی زیر است.

^۱ International Conference on Information Quality

۱. مسئولیت مدیریت
۲. هزینه‌های تضمین و عملیات
۳. تحقیق و توسعه
۴. تولید
۵. توزیع
۶. مدیریت پرسنل
۷. عملکرد قانونی

نلسون و همکاران (Nelson et al. 2005) عواملی را که به طور همزمان بر کیفیت سیستم و اطلاعات تاثیرگذار هستند را مورد بررسی قرار داده و نتایج را در قالب مطالعه موردی روی سازمان‌هایی که انباشت داده دارند مورد تحلیل قرار دادند. بیاگی و همکاران (Biagi et al. 2014) در مرکز تحقیقات کشور ایتالیا کنترل کیفیت در سامانه‌های اطلاعاتی تحقیقاتی و عوامل موثر بر آن را مورد بررسی قرار داده‌اند.

علاوه بر ابعاد کیفیت که در این بخش از منابع و مراجع مختلف موارد متفاوتی ارائه شد و نیز مدل سیستماتیک مدیریت کیفیت داده که به اختصار معرفی گردید سایر حوزه‌های مختلف نیز بایستی مورد توجه قرار گرفته و در مدل‌ها مرتبط با حوزه کیفیت داده در نظر گرفته شوند. در شکل ۴-۳ بصورت شماتیک این حوزه‌ها ارائه شده‌اند. (Batini et al. 2009)



شکل ۴-۳. حوزه‌های اصلی مرتبط در زمینه کیفیت داده

مدل مدیریت کیفیت داده‌چهار سطح مختلف از ابعاد را در معیارهای مختلف مورد توجه قرار داده است که به شرح زیر هستند:

۱. کیفیت ذاتی داده‌ها^۱
۲. کیفیت زمینه‌ای داده‌ها^۲
۳. قابلیت بازنمایی داده‌ها^۳
۴. قابلیت دسترسی^۴

• کیفیت ذاتی داده‌ها

^۱ Intrinsic

^۲ Contextual

^۳ Representational

^۴ Accessibility

کیفیت ذاتی داده ها نه تنها دقت و عینیت را در بر دارد بلکه باورپذیری و اعتبار عمومی را نیز شامل می شود. صحت به معنی ارائه مقادیر واقعی مشخصه های مورد نظر یک مفهوم در حوزه عملیاتی مورد نظر است. مشتریان داده ها برخلاف باور سنتی مهندسان IT فقط صحت و عینیت را مورد توجه قرار نمی دهند بلکه باورپذیری و اعتبار عمومی داده ها نیز برایشان مهم است. برای مثال صحت و عینیت داده های آماری زمانی باورپذیرتر می شود که به مرجع ارائه آمار نیز اشاره شود.

• کیفیت زمینه ای داده ها

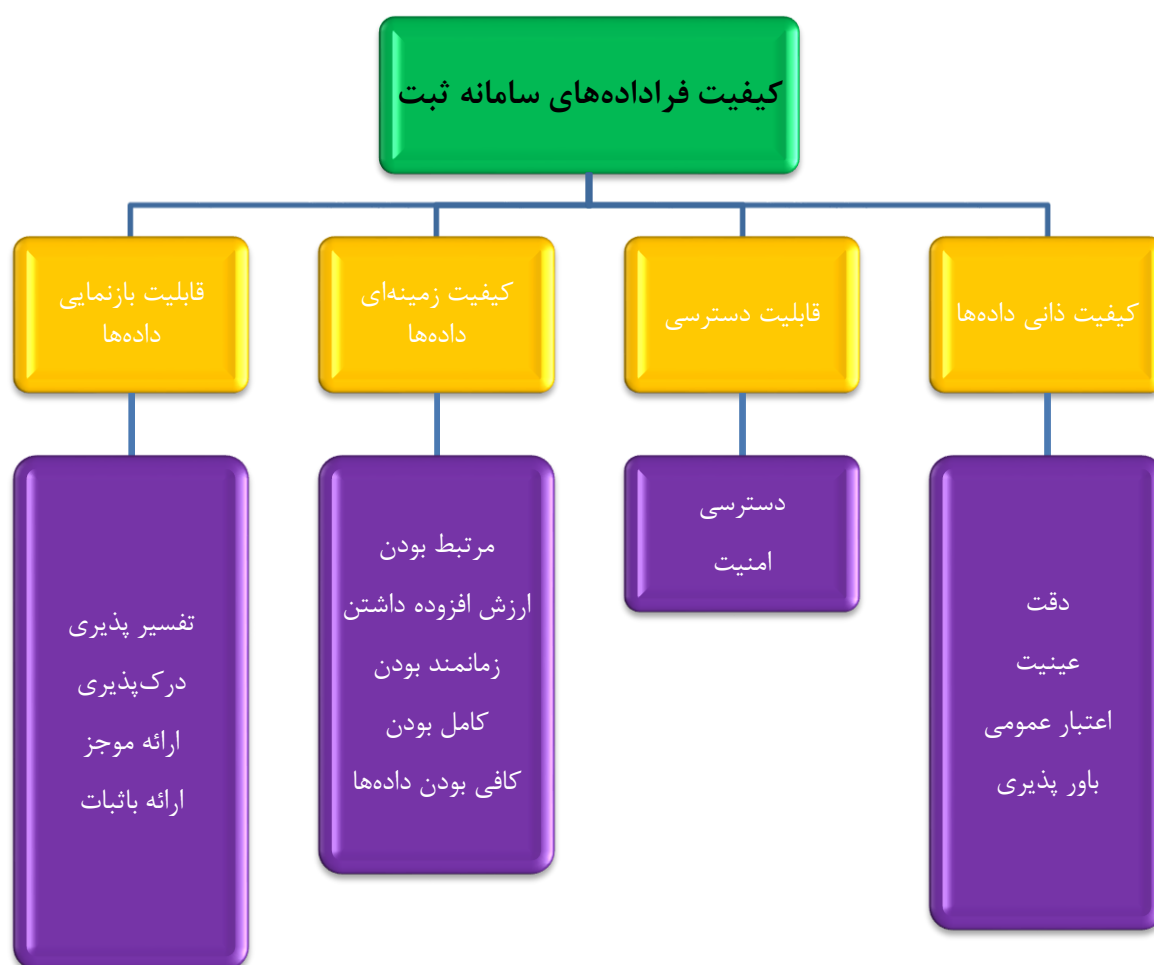
کیفیت داده ها باید در زمینه کاری که داده ها استفاده می شود مورد توجه قرار گیرد. این بعد شامل بر ابعاد فرعی ارزش افزوده داشتن، مرتبط بودن، زمان مند بودن، کامل بودن و کافی بودن داده هاست. از آن جا که کارهایی که با داده ها انجام می گیرد در طول زمان متغیر بوده و نیز کاربران داده ها نیز متفاوت هستند ارزیابی کیفیت زمینه ای داده ها کاری دشوار است. یک روش برای انجام این ارزیابی، تعیین پارامترهای ابعاد ارزیابی هر کار است تا مشتری بتواند بر مبنای آن مشخص کند که چه نوع کاری در حال انجام است و ویژگی داده های مورد استفاده چگونه باید باشد.

• قابلیت بازنمایی داده ها

قابلیت بازنمایی داده ها درباره چارچوب ارائه (ارائه موجز و با ثبات) و معنای داده ها (تفسیرپذیری و درک پذیری) است. این دو جنبه نشان می دهند که داده ها چقدر برای کاربران داده ها خوب ارائه شده اند.

• قابلیت دسترسی

به طور سنتی دسترسی، شاخص جداگانه ای از ابعاد کیفیت شناخته می شد اما تحقیق وانگ (Wang) و همکاران نشان داد که شاخصی کیفی و مورد توجه کاربران داده هاست. دسترسی خود شامل بر دو بعد فرعی دسترسی به داده ها و نیز امنیت دسترسی به داده هاست.



شکل ۴-۴. ابعاد کیفیت فراداده‌ها در سامانه ثبت

در ادامه این بخش به نحوه ترسیم هریک از ابعاد مذکور بر روی هریک از اقلام فراداده پایان‌نامه/رساله دانشجویان خواهیم پرداخت.

۴-۴- نگاشت اولیه ابعاد کیفی فراداده روی اقلام اطلاعاتی سامانه ثبت

در این بخش به ازای هریک از ابعاد کیفی مدل مدیریت کیفیت داده که در بخش قبلی مورد شناسایی قرار گرفت ارتباط میان اقلام اطلاعاتی سامانه ثبت پایان نامه/ رساله با آن ابعاد مورد بررسی قرار می گیرد. جداول ۴-۲ تا ۴-۵ به ازای هر یک از چهار گروه شکل ۴-۴ این ارتباط را نمایان می سازد.

در خصوص جدول فوق لازم بذکرست در برخی از ابعاد به ازای هر فیلد از پایان نامه/ رساله آن بعد قابل تبیین و تفسیر خواهد بود. بعنوان مثال در بعد دقت در فیلد "نام" مواردی بایستی مورد توجه قرار گیرد که با فیلد "کد ملی" متفاوت است. از سوی دیگر برخی ابعاد مانند باورپذیری یا اعتبار عمومی به کل فراداده های پایان نامه/ رساله ها تفسیر شده اند.

جدول ۴-۲. طرح کیفی فراداده‌ها در سامانه ثبت (کیفیت ذاتی)

طرح کیفی فراداده‌ی ثبت پایان نامه‌ها				
کیفیت ذاتی داده‌ها ^۱				نام فیلد
اعتبار عمومی ^۵	باورپذیری ^۴	عینیت ^۳	دقت ^۲	بعد
معتبر بودن در سامانه ثبت پایان نامه: تمامی فیلدهای سامانه دارای اعتبار هستند و هر کاربری با مشاهده ی آنها می‌توانند به سوالات پاسخ ده	در اطلاعات مربوط به متن پایان‌نامه/ رساله قواعد نگارشی به درستی رعایت شده باشد و متن از گویایی کامل برخوردار باشد تا کاربران سامانه گنج به راحتی بتوانند از این اطلاعات استفاده نمایند. بعلاوه در انتخاب واژگان کلیدی و عنوان مناسب توسط نویسنده این موضوع رعایت شده باشد. همچنین در مرحله بعدی از فرایند اشاعه یعنی نمایه‌سازی و ویرایش این موضوع می‌تواند مورد توجه قرار گیرد.	مطابقت با مستندات هویتی (مانند کارت ملی و ...)	مطابقت با متن پایان‌نامه/ رساله	نام
		مطابقت با مستندات هویتی (مانند کارت ملی و ...)	مطابقت با متن پایان‌نامه/ رساله	نام خانوادگی
			در زمان ایجاد پروفایل کاربری برای ادامه ی ثبت نام به رایانامه داده شده یک ایمیل حاوی لینک ورود به قسمت ثبت پایان نامه‌ها فرستاده میشود. حتما باید فرمت زیر را رعایت کند: xxx@xxx.com	رایانامه
			مطابقت با الگوریتم شماره ملی مطابقت با وب سرویس تطبیق کد ملی و نام و نام خانوادگی با ثبت احوال	شماره ملی

^۱ intrinsic^۲ accuracy^۳ objectivity^۴ believability^۵ reputation

شماره دانشجویی	نباید کمتر از ۴ و بیشتر از ۱۰ رقم باشد	مطابقت با اطلاعات مندرج در پایگاه علمی محل تحصیل		
نام فیلد	کیفیت ذاتی داده‌ها ^۱			
	بعد	دقت ^۲	عینیت ^۳	باورپذیری ^۴
مقطع تحصیلی	مطابقت با متن پایان‌نامه/رساله	مطابقت با اطلاعات مندرج در پایگاه علمی محل تحصیل	در اطلاعات مربوط به متن پایان‌نامه/رساله قواعد نگارشی به درستی رعایت شده باشد و متن از گویایی کامل برخوردار باشد تا کاربران سامانه گنج به راحتی بتوانند از این اطلاعات استفاده نمایند.	معتبر بودن در سامانه ثبت پایان‌نامه: تمامی فیلدهای سامانه دارای اعتبار هستند و هر کاربری با مشاهده‌ی آنها می‌توانند به سوالات پاسخ دهد.
وابستگی سازمانی	مطابقت با متن پایان‌نامه/رساله	مطابقت با اطلاعات مندرج در پایگاه علمی محل تحصیل	باشد و متن از گویایی کامل برخوردار باشد تا کاربران سامانه گنج به راحتی بتوانند از این اطلاعات استفاده نمایند.	
موسسه آموزشی/پژوهشی	مطابقت با متن پایان‌نامه/رساله		اطلاعات استفاده نمایند.	
دانشکده/پژوهشکده	مطابقت با متن پایان‌نامه/رساله		بعلاوه در انتخاب واژگان کلیدی و عنوان مناسب توسط نویسنده این موضوع رعایت شده باشد.	
زبان اصلی پایان‌نامه	مطابقت با متن پایان‌نامه/رساله		همچنین در مرحله بعدی از فرایند اشاعه یعنی نمایه‌سازی و	
عنوان فارسی	نباید کمتر از ۱۰ و بیشتر از ۵۰۰ حرف باشد	مطابقت با متن پایان‌نامه/رساله		
عنوان انگلیسی	نباید کمتر از ۱۰ و بیشتر از ۵۰۰ حرف باشد	مطابقت با متن پایان‌نامه/رساله		

^۱ intrinsic

^۲ accuracy

^۳ objectivity

^۴ believability

^۵ reputation

گروه تحصیلی		مطابقت با متن پایان نامه/ رساله		مطابقت با اطلاعات مندرج در پایگاه علمی محل تحصیل		ویرایش این موضوع می تواند مورد توجه قرار گیرد.	
رشته		مطابقت با متن پایان نامه/ رساله		مطابقت با اطلاعات مندرج در پایگاه علمی محل تحصیل			
نام فیلد		کیفیت ذاتی داده ها ^۱					
		دقت ^۲		عینیت ^۳		باورپذیری ^۴	
استاد راهنما/ استاد مشاور		گرایش		مطابقت با متن پایان نامه/ رساله		در اطلاعات مربوط به متن پایان نامه/ رساله قواعد نگارشی به	
		تاریخ دفاع		مطابقت با متن پایان نامه/ رساله		درستی رعایت شده باشد و متن از گویایی کامل برخوردار باشد تا کاربران سامانه گنج به راحتی بتوانند از این	
		نام		نباید کمتر از ۲ و بیشتر از ۲۵ حرف باشد مطابقت با متن پایان نامه/ رساله		اطلاعات استفاده نمایند.	
		نام خانوادگی		نباید کمتر از ۲ و بیشتر از ۵۰ حرف باشد مطابقت با متن پایان نامه/ رساله		بعلاوه در انتخاب واژگان کلیدی و عنوان مناسب توسط نویسنده این	
		نام (انگلیسی)		نباید کمتر از ۲ و بیشتر از ۲۵ حرف باشد مطابقت با متن پایان نامه/ رساله		موضوع رعایت شده باشد.	
		نام خانوادگی (انگلیسی)		نباید کمتر از ۲ و بیشتر از ۵۰ حرف باشد مطابقت با متن پایان نامه/ رساله		همچنین در مرحله بعدی از فرایند اشاعه یعنی نمایه سازی و	
		رایانامه		حتما باید فرمت زیر را رعایت کند: xxx@xxx.com		ویرایش این موضوع	

^۱ intrinsic^۲ accuracy^۳ objectivity^۴ believability^۵ reputation

	شماره تلفن همراه	باید به صورت عدد و با فرمت صحیح باشد، یعنی ۰۹ شروع شود.		می تواند مورد توجه قرار گیرد.	
	شماره ملی	مطابقت با الگوریتم شماره ملی			
کیفیت ذاتی داده‌ها ^۱					
نام فیلد	بعد	دقت ^۲	عینیت ^۳	باورپذیری ^۴	اعتبار عمومی ^۵
	چکیده(فارسی)	نباید کمتر از ۲۰ و بیشتر از ۳۰۰۰ حرف باشد	مطابقت با متن پایان‌نامه/ رساله	در اطلاعات مربوط به متن پایان‌نامه/ رساله	معتبر بودن در سامانه ثبت پایان نامه: تمامی فیلدهای سامانه دارای اعتبار هستند و هر کاربری با مشاهده ی آنها میتوانند به سوالات پاسخ دهد.
	چکیده(انگلیسی)	نباید کمتر از ۲۰ و بیشتر از ۳۰۰۰ حرف باشد	مطابقت با متن پایان‌نامه/ رساله	قواعد نگارشی به درستی رعایت شده باشد و متن از گویایی کامل برخوردار باشد تا کاربران سامانه گنج به راحتی بتوانند از این اطلاعات استفاده نمایند.	
	نوع پایان نامه/ رساله			باشد و متن از گویایی کامل برخوردار باشد تا کاربران سامانه گنج به راحتی بتوانند از این اطلاعات استفاده نمایند.	
	فهرست مطالب	نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد	مطابقت با متن پایان‌نامه/ رساله	توسط نویسنده این موضوع رعایت شده باشد.	
	فهرست منابع فارسی	نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد	مطابقت با متن پایان‌نامه/ رساله	بعلاوه در انتخاب واژگان کلیدی و عنوان مناسب	
	فهرست منابع غیر فارسی	نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد	مطابقت با متن پایان‌نامه/ رساله	توسط نویسنده این موضوع رعایت شده باشد.	
	کلید واژه ها	کامل بودن فیلدها	مطابقت با متن پایان‌نامه/ رساله	همچنین در مرحله بعدی از فرایند اشاعه یعنی نمایه‌سازی و ویرایش این موضوع می تواند مورد توجه قرار گیرد.	
	فایل pdf پایان نامه/ رساله	عدم پذیرش فایل با فرمتی دیگر	کامل بودن محتوا		
	فایل word پایان نامه/ رساله	عدم پذیرش فایل با فرمتی دیگر	کامل بودن محتوا		
	دیگر فایل های پایان نامه/ رساله				

^۱ intrinsic^۲ accuracy^۳ objectivity^۴ believability^۵ reputation

جدول ۴-۳. طرح کیفی فراداده‌ها در سامانه ثبت (کیفیت زمینه‌ای)

طرح کیفی فیلدهای فراداده‌ی ثبت پایان نامه‌ها					
نام فیلد	بعد	کیفیت زمینه‌ای داده‌ها ^۱			
		مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کامل بودن ^۵
نام	تمامی اطلاعاتی که باید در اختیار سامانه ثبت پایان نامه قرار گرفته شود به صورت فیلدهایی تعبیه شده است. به منظور ارزیابی میزان مرتبط بودن داده‌های هر رکورد کلیه نیازمندی‌ها در سامانه‌های آتی را شناسایی نمود تا در صورت نیاز فیلدی اضافه یا کم نمود.	اطلاعات دریافتی در حال حاضر ارزش افزوده کافی را از منظر کاربران سامانه (گنج) دارد و هیچ اطلاعات اضافی دریافت نمی‌شود. در خصوص اضافه شدن فیلدهای دیگر به سامانه ثبت جهت افزایش ارزش افزوده بایستی در جلسات کارشناسی تصمیم‌گیری شود	به روز بودن در سامانه ثبت پایان نامه: تمامی فیلدها به جز فیلدهای نام، نام خانوادگی، رایانامه و شماره ملی که در مرحله ورود اطلاعات شخصی و تحصیلی به صورت از پیش تعیین شده و با توجه با اطلاعات پروفایل کاربری کامل شده اند، میتوانند قبل از تایید نهایی کاربر به روز شوند و در آنها تغییرات ایجاد شود. البته فیلدهای شماره ملی و رایانامه از این قاعده مستثنی هستند و کاربر در آنها نمیتواند تغییری را اعمال کند.	جامعیت در سامانه ثبت پایان نامه: می توان اشاره کرد که با ثبت نام هر دانشجو و اختصاص پروفایل کاربری به هر فرد، هر فردی که دارای پروفایل کاربری است میتواند وارد مراحل ثبت پایان نامه شود و اطلاعات خود را در فیلدهای مشخص شده وارد نماید. وارد کردن اطلاعات در هر فیلدی که با ستاره مشخص شده است برای فرد اجباری است و فیلدهایی که ستاره دار نمی باشند، اختیاری است. به این ترتیب برای هر فرد که در سیستم پروفایل کاربری دارد یکسری اطلاعات قبل از شروع به مراحل ثبت پایان نامه د اختیار سیستم قرار می گیرد.	صرفاً حروف را میپذیرد
نام خانوادگی	نام خانوادگی	مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کامل بودن ^۵
رایانامه	رایانامه	مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کامل بودن ^۵
شماره ملی	شماره ملی	مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کامل بودن ^۵
شماره دانشجویی	شماره دانشجویی	مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کامل بودن ^۵
مقطع تحصیلی	مقطع تحصیلی	مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کامل بودن ^۵
وابستگی سازمانی	وابستگی سازمانی	مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کامل بودن ^۵
موسسه آموزشی/پژوهشی	موسسه آموزشی/پژوهشی	مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کامل بودن ^۵
دانشکده/پژوهشکده	دانشکده/پژوهشکده	مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کامل بودن ^۵
زبان اصلی پایان نامه	زبان اصلی پایان نامه	مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کامل بودن ^۵

^۱ contextual^۲ relevancy^۳ Value added^۴ timeliness^۵ completeness^۶ Amount of data

عنوان فارسی	فیلد اضافه شود.			پذیرفتن حروف و عدد
نام فیلد	کیفیت زمینه‌ای داده‌ها ^۱			
	مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کامل بودن ^۵
استاد راهنما/ استاد مشاور	عنوان انگلیسی	اطلاعات دریافتی در حال حاضر ارزش افزوده کافی را از منظر کاربران سامانه (گنج) دارد و هیچ اطلاعات اضافی دریافت نمی‌شود. در خصوص فیلدهای دیگر به سامانه ثبت جهت افزایش ارزش افزوده بایستی در جلسات کارشناسی تصمیم‌گیری شود	به روز بودن در سامانه ثبت پایان نامه: تمامی فیلدها به جز فیلدهای نام، نام خانوادگی، رایانامه و شماره ملی که در مرحله ی ورود اطلاعات شخصی و تحصیلی به صورت از پیش تعیین شده و با توجه با اطلاعات پروفایل کاربری کامل شده اند، میتوانند قبل از تایید نهایی کاربر به روز شوند و در آنها تغییرات ایجاد شود. البته فیلدهای شماره ملی و رایانامه از این قاعده مستثنی هستند و	جامعیت در سامانه ثبت پایان نامه: می توان اشاره کرد که با ثبت نام هر دانشجو و اختصاص پروفایل کاربری به هر فرد، هر فردی که دارای پروفایل کاربری است میتواند وارد مراحل ثبت پایان نامه شود و اطلاعات خود را در فیلدهای مشخص شده وارد نماید. وارد کردن اطلاعات در هر فیلدی که با ستاره مشخص شده است برای فرد اجباری است و فیلدهایی که ستاره دار نمی باشند، اختیاری است. به این ترتیب برای هر فرد که در سیستم پروفایل کاربری دارد یکسری اطلاعات قبل از شروع به مراحل ثبت پایان نامه د اختیار سیستم قرار می گیرد.
	گروه تحصیلی	تمامی اطلاعاتی که باید در اختیار سامانه ثبت پایان نامه قرار گرفته شود به صورت فیلدهایی تعبیه شده است. به منظور ارزیابی میزان مرتبط بودن داده‌های هر رکورد کلیه نیازمندی‌ها در سامانه‌های آتی را شناسایی نمود تا در صورت نیاز فیلدی اضافه یا کم نمود. بعنوان مثال اگر در آینده قرار است بانک اطلاعاتی پژوهشگران تقویت و تکمیل شود آدرس Google Scholar هر نویسنده یا استاد وی اخذ شود (یک فیلد اضافه شود).		
	رشته			
	گرایش			
	تاریخ دفاع			
	نام			
	نام خانوادگی			
	نام (انگلیسی)			
	نام خانوادگی (انگلیسی)			
	رایانامه			
	شماره تلفن همراه			

^۱ contextual^۲ relevancy^۳ Value added^۴ timeliness^۵ completeness^۶ Amount of data

		شماره ملی			کاربر در آنها نمی‌تواند تغییری را اعمال کند.	
عدد	پذیرفتن حروف و عدد					
کیفیت زمینه‌ای داده‌ها ^۱						
نام فیلد		بعد	مرتبط بودن ^۲	ارزش افزوده داشتن ^۳	زمانمند بودن ^۴	کافی بودن داده ها ^۶
		چکیده(فارسی)	تمامی اطلاعاتی که باید در اختیار سامانه ثبت پایان نامه قرار گرفته شود به صورت فیلدهایی تعبیه شده است. به منظور ارزیابی میزان مرتبط بودن داده‌های هر رکورد کلیه نیازمندی‌ها در سامانه‌های آتی را شناسایی نمود تا در صورت نیاز فیلدی اضافه یا کم نمود. بعنوان مثال اگر در آینده قرار است بانک اطلاعاتی پژوهشگران تقویت و تکمیل شود آدرس Google Scholar هر نویسنده یا استاد وی اخذ شود (یک فیلد اضافه شود).	اطلاعات دریافتی در حال حاضر ارزش افزوده کافی را از منظر کاربران سامانه (گنج) دارد و هیچ اطلاعات اضافی دریافت نمی‌شود. در خصوص اضافه شدن فیلدهای دیگر به سامانه ثبت جهت افزایش ارزش افزوده بایستی در جلسات	به روز بودن در سامانه ثبت پایان نامه: تمامی فیلدها به جز فیلدهای نام، نام خانوادگی، رایانامه و شماره ملی که در مرحله ی ورود اطلاعات شخصی و تحصیلی به صورت از پیش تعیین شده و با توجه با اطلاعات پروفایل کاربری کامل شده اند، می‌توانند قبل از تایید نهایی کاربر به روز شوند و در آنها تغییرات ایجاد شود. البته فیلدهای شماره ملی و رایانامه از این قاعده مستثنی هستند و کاربر در آنها نمی‌تواند تغییری را اعمال کند.	جامعیت در سامانه ثبت پایان نامه: می توان اشاره کرد که با ثبت نام هر دانشجو و اختصاص پروفایل کاربری به هر فرد، هر فردی که دارای پروفایل کاربری است میتواند وارد مراحل ثبت پایان نامه شود و اطلاعات خود را در فیلدهای مشخص شده وارد نماید. وارد کردن اطلاعات در هر فیلدی که با ستاره مشخص شده است برای فرد اجباری است و فیلدهایی که ستاره دار نمی باشند، اختیاری است. به این ترتیب برای هر فرد که در سیستم پروفایل کاربری دارد یکسری اطلاعات قبل از شروع به مراحل ثبت پایان نامه د اختیار سیستم قرار می گیرد.
پذیرش صرفا حروف انگلیسی		چکیده(انگلیسی)				
		نوع پایان نامه/ رساله				
پذیرفتن حروف و عدد		فهرست مطالب				
پذیرفتن حروف و عدد		فهرست منابع فارسی				
پذیرش صرفا حروف و اعداد انگلیسی		فهرست منابع غیر فارسی				
پذیرفتن حروف و عدد		کلید واژه ها				
پذیرش		فایل pdf پایان نامه/ رساله				

^۱ contextual^۲ relevancy^۳ Value added^۴ timeliness^۵ completeness^۶ Amount of data

فایل			کارشناسی		
پذیرش فایل			تصمیم‌گیری شود		فایل word پایان نامه/ رساله
پذیرش فایل					دیگر فایل های پایان نامه/ رساله

جدول ۴- ۴. طرح کیفی فراداده‌ها در سامانه ثبت (کیفیت بازنمایی)

طرح کیفی فیلدهای فراداده‌ی ثبت پایان نامه‌ها					
قابلیت بازنمایی داده‌ها ^۱				نام فیلد	
ارائه بایستای	ارائه موجز	درک پذیری	تفسیرپذیری	بعد	
طراحی سامانه ثبت بایستی به گونه‌ای باشد که داده‌ها و فراداده‌های ثبت شده در آن با تغییرات آتی در سخت افزارها و به روز شدن نرم‌افزارها ارائه و نمایششان با مشکل مواجه نشود.	فرمت کاراکترهای تایپ شده بایستی به گونه‌ای باشد که در نمایش در سامانه‌های بعدی مانند نمایه‌سازی و ویرایش و همچنین گنج فرمت مطلوب کاربران را داشته باشد.	سهولت درک در سامانه ثبت پایان نامه: تمامی فیلدهای موجود در سامانه به راحتی برای کاربر قابل تشخیص و قابل فهم می باشد.	فرمت و ساختار کلیه فیلدهای سامانه ثبت به گونه‌ای بایستی باشد که به راحتی و بدون اشکال در سامانه گنج انتقال داده شود و برای کاربران سامانه گنج که از مرورگرها و سیستم عاملهای مختلف استفاده میکنند قابل نمایش باشد.	نام نام خانوادگی رایانامه شماره ملی شماره دانشجویی مقطع تحصیلی وابستگی سازمانی موسسه آموزشی/پژوهشی دانشکده/پژوهشکده زبان اصلی پایان نامه عنوان فارسی عنوان انگلیسی گروه تحصیلی رشته گرایش تاریخ دفاع	استاد راهنما/ استاد مشاور
	قابلیت انتخاب از میان گزینه های موجود			نام	
	قابلیت انتخاب از میان گزینه های موجود			نام خانوادگی	
	قابلیت انتخاب از میان گزینه های موجود			رایانامه	
	قابلیت انتخاب از میان گزینه های موجود			شماره ملی	
				شماره دانشجویی	
	قابلیت انتخاب از میان گزینه های موجود			مقطع تحصیلی	
	قابلیت انتخاب از میان گزینه های موجود			وابستگی سازمانی	
				موسسه آموزشی/پژوهشی	استاد راهنما/ استاد مشاور
				دانشکده/پژوهشکده	
				زبان اصلی پایان نامه	
				عنوان فارسی	
	قابلیت انتخاب از میان گزینه های موجود			عنوان انگلیسی	
	قابلیت انتخاب از میان گزینه های موجود			گروه تحصیلی	
	قابلیت انتخاب از میان گزینه های موجود			رشته	
	قابلیت انتخاب از میان گزینه های موجود			گرایش	
				تاریخ دفاع	
				نام	استاد راهنما/ استاد مشاور
				نام خانوادگی	
				نام (انگلیسی)	
				نام خانوادگی (انگلیسی)	
				رایانامه	
				شماره تلفن همراه	

^۱ contextual

				شماره ملی	
				چکیده(فارسی)	
				چکیده(انگلیسی)	
				نوع پایان نامه/ رساله	
				فهرست مطالب	
قابلیت بازنمایی داده‌ها ^۱				نام فیلد	بعد
ارائه بایستای	ارائه موجز	درک پذیری	تفسیر پذیری		
طراحی سامانه ثبت بایستی به گونه‌ای باشد که داده‌ها و فراداده‌های ثبت شده در آن با تغییرات آتی در سخت افزارها و به روز شدن نرم‌افزارها ارائه و نمایششان با مشکل مواجه نشود.		سهولت درک در سامانه ثبت پایان نامه: تمامی فیلدهای موجود در سامانه به راحتی برای کاربر قابل تشخیص و می باشد.	فرمت و ساختار کلیه فیلدهای سامانه ثبت به گونه ای بایستی باشد که به راحتی و بدون اشکال در سامانه گنج انتقال داده شود و برای کاربران سامانه گنج که از مرورگرها و سیستم عاملهای مختلف استفاده میکنند قابل نمایش باشد.	فهرست منابع فارسی	
			فهرست منابع غیر فارسی		
			کلید واژه ها		
			فایل pdf پایان نامه/ رساله		
			فایل word پایان نامه/ رساله		
			دیگر فایل های پایان نامه/ رساله		

جدول ۴-۵. طرح کیفی فراداده‌ها در سامانه ثبت (قابلیت دسترسی)

طرح کیفی فیلدهای فراداده‌ی ثبت پایان نامه‌ها			
قابلیت دسترسی ^۱		نام فیلد	
امنیت ^۳	دسترسی ^۲	بعد	
امنیت در سامانه ثبت پایان نامه: داشتن رمز عبور و رایانامه که هر کاربر تنها برای پروفایل کاربری برای دسترسی به سامانه ثبت پایان نامه خود استفاده میکند، نمونه ای از داشتن امنیت است.	طراحی لایه‌های امنیتی در سامانه ثبت و همچنین سامانه گنج که به طور مستقیم با کاربران نهایی در ارتباط هستند بایستی به گونه‌ای باشد که امکان دسترسی به اطلاعات شخصی دانشجویان تحصیلات تکمیلی برای افراد غیرمجاز مقدور نباشد.	نام	
		نام خانوادگی	
		رایانامه	
		شماره ملی	
		شماره دانشجویی	
		مقطع تحصیلی	
		وابستگی سازمانی	
		موسسه آموزشی/پژوهشی	
		دانشکده/پژوهشکده	
		زبان اصلی پایان نامه	
		عنوان فارسی	
		عنوان انگلیسی	
		گروه تحصیلی	
		رشته	
		گرایش	
		تاریخ دفاع	
		نام	استاد راهنما/ استاد مشاور
		نام خانوادگی	
		نام (انگلیسی)	
		نام خانوادگی (انگلیسی)	
		رایانامه	
		شماره تلفن همراه	
		شماره ملی	

^۱ accessibility

^۲ access

^۳ security

		چکیده(فارسی)
قابلیت دسترسی ^۱		نام فیلد بعد
امنیت ^۲	دسترسی ^۲	
امنیت در سامانه ثبت پایان نامه: داشتن رمز عبور و رایانامه که هر کاربر تنها برای پروفایل کاربری برای دسترسی به سامانه ثبت پایان نامه خود استفاده میکند، نمونه ای از داشتن امنیت است.	طراحی لایه‌های امنیتی در سامانه ثبت و همچنین سامانه گنج که به طور مستقیم با کاربران نهایی در ارتباط هستند بایستی به گونه‌ای باشد که امکان دسترسی به اطلاعات شخصی دانشجویان تحصیلات تکمیلی برای افراد غیرمجاز مقدور نباشد	چکیده(انگلیسی)
		نوع پایان نامه/ رساله
		فهرست مطالب
		فهرست منابع فارسی
		فهرست منابع غیر فارسی
		کلید واژه ها
		فایل pdf پایان نامه/ رساله
		فایل word پایان نامه/ رساله
		دیگر فایل های پایان نامه/ رساله

۴-۵- نتیجه‌گیری و جمع‌بندی مطالب فصل چهارم

در فصل حاضر به معرفی ابعاد مختلف مدیریت کیفیت داده در چهار حوزه کیفیت ذاتی، زمینه‌ای، بازنامایی و دسترسی، و همچنین چگونگی جایگاه آنها بر روی کیفیت نهایی اقلام اطلاعاتی سامانه ثبت مورد بررسی قرار گرفت. در واقع به ازای هر قلم اطلاعاتی سامانه ثبت وضعیت مطلوب آن قلم در این چهار بعد مورد ارزیابی قرار گرفت. همان‌طور که از جداول ۴-۲ تا ۴-۵ مشاهده می‌شود، هریک از اقلام اطلاعاتی در سامانه ثبت برپایه مدل TDQM وضعیت مطلوب توصیف شده در این جداول را خواهد داشت. در این فصل مدل TDQM بعنوان یک مدل مطلوب، پرکاربرد و متناسب برای بومی شدن و نهادینه شدن در سامانه ثبت معرفی گردد. چهار مرحله اصلی این مدل شامل برنامه‌ریزی، اندازه‌گیری، تجزیه‌تحلیل و بهبود بعنوان چارچوب اصلی این مدل مورد اشاره قرار گرفت. در ادامه به کمک گروههای خبرگان هریک از ابعاد کیفی فراداده

^۱ accessibility

^۲ access

^۳ security

که در جداول بالا ارائه شده تعریف و تبیین دقیق‌تری خواهد یافت. عدم انطباق‌های مشاهده شده در هر فراداده تعیین و دسته‌بندی خواهد شد و همچنین چک‌لیست‌های کنترل کیفیت فراداده‌های سامانه ثبت نهایی خواهد شد.

فصل پنجم – تدوین طرح کیفیت فراداده‌ها در سامانه ثبت

۵-۱ تدوین اولیه طرح کیفیت سامانه ثبت بر مبنای مصاحبه با پرسنل کلیدی

همانگونه که در بخش چهارم اشاره شد سامانه ثبت پایان نامه / رساله دانش آموختگان داخل و خارج از کشور پایه و اساس داده‌های اشاعه داده شده در سامانه اشاعه پایان نامه / رساله‌ها (گنج) خواهد بود. لذا در صورتی که بتوان مشکلات کیفی داده‌ها در هریک از اقلام اطلاعاتی (فیلدها) را در این سامانه مرتفع نمود کیفیت داده‌های اشاعه داده شده در گنج به همان نسبت ارتقاء خواهد یافت. در این بخش مستندی با عنوان طرح کیفیت تدوین و توسعه داده خواهد شد که طی آن کیفیت مطلوب و مورد انتظار هریک از اقلام اطلاعاتی سامانه ثبت در آن ارائه شده است. بمنظور ارائه مستند طرح کیفیت مناسب در ابتدا طی جلساتی که با رئیس اداره ثبت و فراهم‌آوری اطلاعات برگزار گردید پرسنل کلیدی معاونت اطلاعات علم و فناوری که متولی کنترل داده‌های ثبت شده در سامانه ثبت هستند شناسایی گردید و جلسات متعددی با این کارشناسان برگزار شده و طرح اولیه کیفیت فراداده‌ها در سامانه ثبت مبتنی بر این جلسات تدوین شد. در حین برگزاری جلسات با کارشناسان کنترل‌های پراکنده‌ای که توسط هریک از همکاران صورت می‌گرفت شناسایی و مستند گردید که در بخش ۵-۱-۱ نحوه این کنترل‌ها و چک‌لیست‌های موجود ارائه شده است. در ادامه بر مبنای مصاحبه‌های صورت گرفته و مطالعه چک‌لیست‌های هریک از افراد طرح اولیه کیفیت فراداده تدوین گردید که این طرح در بخش ۵-۱-۲ ارائه شده است. در بخش ۵-۱-۳ فرایند موجود کنترل کیفیت در سامانه ثبت بصورت مرحله به مرحله گزارش شده است. در بخش ۵-۲ مراحل می‌گرفت شناسایی و مستند گردید که ثبت منتج شده است تبیین شده است. بمنظور رفع نواقص طرح ارائه شده و نیز صحت‌گذاری آن خبرگان پژوهشگاه علوم و فناوری اطلاعات ایران که به فرایند، مشکلات احتمالی و سیاست‌های سازمان در این حوزه مسلط بودند شناسایی شده و جلسات مختلفی برگزار گردید که نتیجه آن طرح کیفی فراداده‌های سامانه ثبت پایان نامه / رساله دانش آموختگان داخل و خارج از کشور بود. در ادامه و در بخش ۵-۳ فرایند کنترل کیفیت سامانه در قالب چک‌لیست‌های کنترل کیفی ارائه شده است. در بخش ۵-۴ پیشنهادات مختلفی در حوزه طراحی ویرایش‌های آتی سامانه ثبت بر مبنای آنچه در جلسات خبرگان سامانه ثبت مورد بررسی قرار گرفت ارائه شده است.

۵-۱-۱ کنترل‌های صورت گرفته توسط کارشناسان اداره ثبت و فراهم‌آوری اطلاعات

در این بخش کنترل‌های صورت گرفته توسط کارشناسان اداره ثبت و فراهم‌آوری اطلاعات ارائه شده است. این کنترل‌ها بدون اشاره به نام فرد در ادامه آورده شده است. در شکل‌های ۵-۱ تا ۵-۳ آورده شده است.

نمونه چک لیست برای رد ایرادناک:

لطفا فایل‌های ورد و پی دی اف را هر کدام بطور مجزا و کامل از صفحه ابتدا(بسم اله) تا آخرین صفحه بارگذاری نمائید : شامل صفحه عنوان فارسی و انگلیسی، چکیده فارسی و انگلیسی، فهرست مطالب با شماره صفحه و فهرست منابع و کل فصلهای پایان نامه و پیوست ها است.

لطفا فایل ورد را بطور کامل از صفحه ابتدا(بسم اله) تا آخرین صفحه بارگذاری نمائید : شامل صفحه عنوان فارسی و انگلیسی، چکیده فارسی و انگلیسی، فهرست مطالب و فهرست منابع و کل فصلهای پایان نامه و پیوست ها است.(همانند فایل پی دی اف)

در صورتی که قصد دارید فایل پروپوزال یا پیشنهاد خود را ثبت کنید، لطفا در بخش ثبت پیشنهاده اینکار را انجام دهید. در غیر این صورت فایل‌های ورد و پی دی اف پایان نامه را هر کدام بطور مجزا و کامل بارگذاری نمائید: شامل صفحه عنوان فارسی و انگلیسی، چکیده فارسی و انگلیسی، فهرست مطالب و فهرست منابع و کل فصلهای پایان نامه و پیوست ها است.

لطفا در فایل‌های ورد و پی دی اف شماره صفحات فهرست مطالب را کامل نموده و مجددا بارگذاری نمائید.

لطفا صفحه عنوان فارسی (اسم و لوگوی دانشگاه - دانشکده - گروه و گرایش و رشته - عنوان پایان نامه - مولف - استاد راهنما و مشاور - تاریخ دفاع) و صفحه چکیده فارسی را در فایل ورد و پی دی اف قرار داده و مجددا بارگذاری نمائید.

لطفا صفحه عنوان فارسی (اسم و لوگوی دانشگاه - دانشکده - گروه و گرایش و رشته - عنوان پایان نامه - مولف - استاد راهنما و مشاور - تاریخ دفاع) را در فایل ورد و پی دی اف قرار داده و مجددا بارگذاری نمائید.

لطفا صفحه عنوان انگلیسی (اسم و لوگوی دانشگاه - دانشکده - گروه و گرایش و رشته - عنوان پایان نامه - مولف - استاد راهنما و مشاور - تاریخ دفاع) و چکیده انگلیسی را در فایل ورد و پی دی اف قرار داده و مجددا بارگذاری نمائید.

لطفا صفحه عنوان انگلیسی (اسم و لوگوی دانشگاه - دانشکده - گروه و گرایش و رشته - عنوان پایان نامه - مولف - استاد راهنما و مشاور - تاریخ دفاع) را در فایل ورد و پی دی اف قرار داده و مجددا بارگذاری نمائید.

ثبت پروپوزال یا پیشنهاد در بخش دیگر همین سامانه به نام ثبت پیشنهاده میباشد که میتوانید فایل‌های خود را آنجا بارگذاری نمائید. در صورتی که پایان نامه شما دفاع شده، میتوانید در این قسمت فایل‌های آن را بارگذاری نمائید.

در صورتی که در تهیه فایل پایان نامه خود از نرم‌افزار دیگری **غیر از Word** استفاده نموده‌اید، فایل‌های آن را به شکل زیر شده در **بخش دیگر فایلها** قرار دهید.

لطفا در فایل‌های ورد و پی دی اف صفحه عنوان انگلیسی، چکیده فارسی و انگلیسی را قرار داده و مجددا بارگذاری نمائید.

لطفا در فایل‌های ورد و پی دی اف صفحه عنوان انگلیسی و چکیده انگلیسی را قرار داده و مجددا بارگذاری نمائید.

لطفا در فایل‌های ورد و پی دی اف صفحه عنوان انگلیسی را قرار داده و مجددا بارگذاری نمائید.

لطفا در فایل‌های ورد و پی دی اف صفحه چکیده انگلیسی را قرار داده و مجددا بارگذاری نمائید.

لطفا پیوست‌ها و ضمايم ذکر شده در فهرست مطالب را در فایل‌های ورد و پی دی اف قرار داده و مجددا بارگذاری نمائید.

لطفا در فایل‌های ورد و پی دی اف صفحات تایید داوران، آئين نامه حق مالکيت، آئين نامه حق چاپ را بصورت امضا شده، قرار داده و مجددا بارگذاری نمائید. (تربيت مدرس)

لطفا در فایل‌های ورد و پی دی اف صفحه عنوان فارسی و چکیده فارسی را قرار داده و مجددا بارگذاری نمائید.

فایل پی دی اف خراب است و باز نمی‌شود و لطفا در فایل ورد فهرست منابع فارسی و غیر فارسی را قرار داده و مجددا بارگذاری نمائید.

لطفا در فایل‌های ورد و پی دی اف چکیده فارسی را قرار داده و مجددا بارگذاری نمائید.

لطفا کلیدواژه‌های فارسی و انگلیسی درست و مناسب را قرار داده و مجددا بارگذاری نمائید.

لطفا در فایل‌های ورد و پی دی اف فهرست منابع را قرار داده و مجددا بارگذاری نمائید.

لطفا از کلیدواژه‌های مناسب استفاده نمائید.

لطفا در فایل پی دی اف در قسمت فرم مشخصات چاپ چکیده در پژوهشنامه آدرس شخصی را حذف نمائید و مجددا بارگذاری نمائید.

شماره صفحات مندرج در فهرست مطالب با شماره صفحات مندرج در پایان نامه منطبق نیست. لطفا اصلاح شود و سپس فایل‌های ورد و پی دی اف مجددا بارگذاری نمائید.

پیوست پلان و نقشه‌ها در فایل ورد و پی دی اف موجود نیست. لذا لطفا موارد در فایل ورد و پی دی اف قرار داده و مجددا بارگذاری نمائید.

در فایل‌های ورد و پی دی اف فونت‌های متن فهرست مطالب قابل خواندن نیست و کلمات به هم چسبیده است، لطفا آن را اصلاح نموده و فایل‌ها را مجددا بارگذاری نمائید.

در فایل ورد فونت‌های متن فهرست مطالب قابل خواندن نیست لطفا از فونت مناسب استفاده کرده و فایل ورد را مجددا بارگذاری نمائید.

در فایل پی دی اف فونت‌های متن فهرست مطالب قابل خواندن نیست لطفا از فونت مناسب استفاده کرده و در فایل‌های ورد و پی دی اف صفحه عنوان انگلیسی، چکیده فارسی را قرار داده و مجددا بارگذاری نمائید.

فهرست مطالب دارای شماره صفحه نمی‌باشد. طبق سرعنوان‌ها شماره صفحه در فهرست مطالب لطفا اصلاح شود و سپس فایل‌های ورد و پی دی اف را مجددا بارگذاری نمائید.

شماره صفحات و سرعنوان‌ها مندرج در فهرست مطالب به هم ریخته و نامشخص می‌باشد، لطفا اصلاح شود و سپس فایل‌های ورد و پی دی اف را مجددا بارگذاری نمائید.

در فایل‌های ورد و پی دی اف فهرست مطالب شماره صفحه ندارد، لطفا شماره صفحات را مطابق با سرعنوان قرار دهید. همچنین صفحه در فایل‌های ورد و پی دی اف عنوان انگلیسی و چکیده انگلیسی را قرار داده و مجددا بارگذاری نمائید.

در فایل‌های ورد و پی دی اف فهرست مطالب دارای شماره صفحه نیست، لطفا شماره هر سرفصل و سر عنوان را در فهرست مطالب مشخص کرده، مجددا بارگذاری نمایید.

در فایل‌های ورد و پی دی اف شماره صفحات مندرج در فهرست مطالب با شماره صفحات پایان نامه منطبق نیست. لطفا این مورد تصحیح شود. در ضمن صفحه عنوان انگلیسی و چکیده انگلیسی را نیز در فایل‌های ورد و پی دی اف قرار داده و مجددا بارگذاری نمایید.

صفحات فهرست مطالب ناقص است. لطفا در فایل‌های ورد و پی دی اف صفحات فهرست مطالب را اصلاح نموده و مجددا بارگذاری نمایید.

بعضی از قسمتهای فهرست مطالب دارای شماره صفحه نمی باشد و Error! Bookmark not defined را نشان می دهد، لطفا اصلاح شود، سپس فایل‌های ورد و پی دی اف را مجددا بارگذاری نمایید.

در فایل ورد کلمات فهرست مطالب بهم چسبیده است و قابل خواندن نیست لطفا تصحیح شود و مجددا بارگذاری گردد. در صفحه فهرست مطالب، پیوست ذکر شده است اما در بخش پیوست هیچ مطلبی وجود ندارد، لذا بخش پیوست را کامل نموده و فایل‌های ورد و پی دی اف را مجددا بارگذاری نمایید.

عنوان پایان نامه مندرج در فایل پی دی اف با عنوان مندرج در فایل ورد مطابقت ندارد. لطفا عنوان تصحیح شود و مجددا فایل ورد و پی دی اف بارگذاری گردد.

لطفا با شماره تلفن ۶۶۹۵۱۴۳۰ داخلی ۲۹۴ در ساعات اداری تماس بگیرید.

شکل ۵- ۱. نمونه چک لیست اول در کنترل فراداده‌ها در سامانه ثبت

در ادامه چک لیست دیگری در این خصوص ارائه شده است.

بسمه تعالی

ضمن تقدیر و تشکر از زحمات و دقت نظر همه همکاران محترم، لطفا موارد زیر حتما در کنترل فایلها مورد بررسی قرار گیرد و هر یک از موارد زیر در فایل های پایان نامه موجود نیست پذیرش نگردد:

- عنوان فارسی
- چکیده فارسی
- فهرست مطالب
- کلیه فصلها به همراه اشکال و جداول
- منابع و مآخذ

- پیوست ها (در صورت اشاره در فهرست مطالب)

- چکیده لاتین

- عنوان لاتین

نکات مهم:

- پایان نامه های دانشگاه تربیت مدرس صفحات حقوقی دانشگاه در ابتدای مدرک کنترل شود.

- پایان نامه هایی که فاقد فایل ورد هستند (فقط فهرست مطالب و منابع) و با نرم افزارهای دیگر نگاشته شده است، در صورتیکه فایل پی دی اف کامل را دارد، پذیرش نماید.

- پایان نامه های دانشگاه جامعه المصطفی (ص) حتما دارای عنوان فارسی و چکیده فارسی باشد.

- اشکالات نمایش فونتها در فایل ورد حتما در سیستم دیگری بررسی و در صورت داشتن مشکل برگشت داده شود.

- فایل پایان نامه ها بدون پس نقش (واترمارک) باشد.

- صفحات سفید در فایل پی دی اف حذف گردد.

- فایل پایان نامه ها بدون قفل باشد و به راحتی قابل ویرایش باشد.

- فایل پایان نامه بدون آدرس و مشخصات شخصی دانشجو و استاد راهنما باشد.

- اشتباهات موجود در ورود اطلاعات کاربران تا حد ممکن اصلاح شود.

با تشکر - ۱۱ دیماه ۹۵

شکل ۵-۲. نمونه چک لیست دوم در کنترل فراداده ها در سامانه ثبت

باسلام

فایل ناقص است

فایل های پایان نامه باید دارای موارد ذیل باشد و بصورت کامل و یکپارچه ارسال گردد.

۱. صفحه عنوان فارسی و انگلیسی

۲. چکیده فارسی و انگلیسی

۳. فهرست مطالب

۴. متن تمام فصلها

۵. منابع و مراجع

۶. ضمائم و پیوست ها در صورت موجود

با سلام

لطفاً موارد ذیل را به فایل‌های پایان‌نامه خود اضافه کنید و دوباره ارسال فرمایید

۱. تاییدیه داوران با امضاء

۲. آیین‌نامه حق چاپ امضا

۳. دستورالعمل حق مالکیت با امضا

با سلام

شماره صفحات با فهرست مطالب همخوانی ندارد

لطفاً اصلاح فرمایید

با تشکر

با سلام

لطفاً کلیدواژه‌ها را در سایت ایرانداک بصورت تک‌تک و جداگانه در فیلدها وارد کنید

با سلام

فایل ارسالی مربوط پروپوزال می باشد

لطفاً فایل‌های خود را در قسمت مربوط به پروپوزال ارسال فرمایید.

با تشکر

با سلام

لطفاً صفحه عنوان انگلیسی و چکیده انگلیسی را به فایل‌های پایان‌نامه اضافه کرده و دوباره ارسال فرمایید

با تشکر

با سلام

لطفاً صفحه عنوان انگلیسی را به فایلها اضافه کنید و دوباره ارسال فرمایید

صفحه عنوان باید دارای موارد ذیل باشد

۱. آرم و نام دانشگاه
۲. دانشکده
۳. مقطع تحصیلی
۴. موضع پایان نامه
۵. نام استاد راهنما
۶. نام دانشجو
۷. تاریخ دفاع

با سلام

لطفاً در فایلهای لاتکس را در قسمت دیگر فایلهای پایان نامه ارسال نمایید

با تشکر

شکل ۵-۳. نمونه چک لیست سوم در کنترل فراداده‌ها در سامانه ثبت

همانگونه که در شکل‌های ۵-۱ تا ۵-۳ مشخص است یکپارچگی و هماهنگی میان چک لیست‌های مورد استفاده قرار گرفته توسط کارشناسان مختلف اداره ثبت و فراهم‌آوری اطلاعات مشاهده نمی‌شود و لزوم تدوین و مستندسازی کنترل‌های صورت گرفته در سامانه ثبت به صورت یکپارچه احساس می‌گردد لذا برپایه مصاحبه با پرسنل کلیدی که در امور کنترل کیفیت داده‌ها در سامانه ثبت همکاری دارند طرح اولیه کیفیت فراداده‌ها در سامانه ثبت ارائه شد که در ادامه به معرفی آن خواهیم پرداخت.

۱,۲-۵ تعیین، دسته‌بندی و درجه‌بندی عدم انطباق‌های بالقوه در سامانه ثبت پایان‌نامه/رساله‌ها

در جداول ۱-۵ تا ۳-۵ طرح اولیه کیفیت فراداده‌ها در سامانه ثبت در سه دسته مختلف ارائه شده است. دسته اول مربوط به اقلام اطلاعاتی (فراداده) هر پایان‌نامه/رساله است که توسط کارشناس مربوطه کنترل می‌گردد. دسته دوم مربوط به کنترل‌های صورت گرفته در محتوای فایل پی دی اف می‌باشد و دسته سوم مربوط به کنترل‌های صورت گرفته در فایل ورد می‌باشد.

همانگونه که در این جداول نیز مشاهده می‌گردد در این جداول به ازای هریک از اقلام اطلاعاتی وضعیت مطلوب و عدم انطباق بالقوه آن قلم اطلاعاتی تبیین شده است. عدم انطباق مربوط به هر قلم اطلاعاتی در دو دسته A و B با تعاریف زیر تقسیم شده است.

عدم انطباق نوع A: این دسته از عدم انطباق‌ها شامل مواردی است که بحرانی بوده و شخص کنترل کننده بایستی در صورت مشاهده یکی از این موارد مدرک را به دانشجو عودت دهد.

عدم انطباق نوع B: این دسته از عدم انطباق‌ها شامل مواردی بوده که غیر بحرانی بوده و در صورت مشاهده این موارد کنترل کننده می‌تواند اصلاحات را در سامانه انجام دهد. وجود چندین عدم انطباق نوع B می‌تواند منجر به عدم انطباق نوع A در مدرک شده و عودت داده شود.

جدول ۵-۱: طرح اولیه کیفیت فراداده‌ها در سامانه ثبت

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام فیلد
نوع B	نوع A		
توضیحات: از آنجاکه امکان اصلاح نام و نام خانوادگی در سامانه وجود ندارد در صورت مشاهده هرگونه عدم انطباقی مدرک به دانشجو عودت داده خواهد شد و لذا عدم انطباقی نوع B در این موارد وجود ندارد.	ناقص بودن نام در سامانه ثبت در موارد زیر - مواردی مانند ال، حروف اضافه و ... به درستی نوشته نشده است. - یک حرف اشتباه تایپ شده است (پ به جای ژ یا ...) - نام بصورت لاتین نوشته شده است - نام با صفحه عنوان پایان‌نامه مطابقت ندارد. - بخشی از یک نام چندبخشی نوشته نشده باشد.	مطابقت با متن پایان‌نامه/ رساله	نام
	ناقص بودن نام خانوادگی در سامانه ثبت - مواردی مانند ال، حروف اضافه و ... به درستی نوشته نشده است. - یک حرف اشتباه تایپ شده است (پ به جای ژ یا ...) - نام خانوادگی بصورت لاتین نوشته شده است - نام خانوادگی با صفحه عنوان پایان‌نامه مطابقت ندارد. - بخشی از یک نام خانوادگی چندبخشی نوشته نشده باشد.	مطابقت با متن پایان‌نامه/ رساله	نام خانوادگی
رایانامه از آنجاکه که مبنای مکاتبات و تعامل با دانشجویان است لذا بصورت غیر مستقیم و در عمل مورد صحه‌گذاری قرار می‌گیرد و کنترلی روی آن انجام نخواهد شد.		مطابقت با فرمت زیر: xxx@xxx.com	رایانامه

دسته‌بندی عدم انطباق‌ها در اقسام اطلاعاتی سامانه ثبت			
نام فیلد	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
شماره ملی	مطابقت با الگوریتم شماره ملی مطابقت با وب سرویس تطبیق کد ملی و نام و نام خانوادگی در سامانه ثبت احوال	در حال حاضر با توجه به عدم امکان استفاده از سرویس ثبت احوال امکان کنترل شماره ملی وجود ندارد و کنترلی روی آن انجام نخواهد شد.	
شماره دانشجویی	مطابقت با مشخصات دانشجو در بانک‌های اطلاعاتی داخلی مرکز آموزشی/ پژوهشی	این فیلد توسط نماینده دانشگاه کنترل خواهد شد.	
مقطع تحصیلی	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله	• وجود هرگونه مغایرت میان صفحه عنوان و فیلد تکمیل شده در سامانه ثبت	-
وابستگی سازمانی	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله	• عدم تطابق با صفحه عنوان پایان‌نامه/ رساله	• عدم تطابق میان وابستگی سازمانی و نام دانشگاه (بعنوان مثال وابستگی سازمانی وزارت علوم انتخاب شده است در حالیکه دانشگاه آزاد بعنوان نام موسسه آموزشی ذکر شده است)
موسسه آموزشی/پژوهشی	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله	• در دانشگاه‌هایی مانند دانشگاه آزاد و پیام نور واحدها مشخص نشده باشد. • مغایرت با صفحه عنوان پایان‌نامه/ رساله	توضیحات: از آنجاکه پایان‌نامه/ رساله پس از تایید ایرانداک به کارتابل نماینده دانشگاه می‌رود و در صورت انتخاب اشتباه موسسه به کارتابل فرد دیگری رفته و مشکلات

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت			
نام فیلد	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
			اجرائی ایجاد مینماید لذا در این آیتم بایستی دقت فراوان صورت گرفته و در صورت هرگونه مغایرت به دانشجو عودت داده شود.
دانشکده/پژوهشکده	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله		مغایرت با نام دانشکده/پژوهشکده مندرج در صفحه عنوان پایان‌نامه/ رساله
زبان اصلی پایان نامه	مطابقت با متن پایان‌نامه/ رساله (بایستی یکی از زبانهای فارسی، انگلیسی، فرانسوی، آلمانی یا عربی باشد).	هرگونه مغایرت میان زبان انتخاب شده و متن نوشته شده در پایان‌نامه/ رساله	-
عنوان فارسی	مطابقت با عنوان مندرج در صفحه اول پایان‌نامه/ رساله	وجود مغایرت اساسی (چند کلمه یا بیشتر) میان فیلد عنوان و متن پایان‌نامه	- وجود کاراکترهای اضافی مانند ^۱space - وجود غلط‌های املایی در متن عنوان - عدم انطباق کامل (بیش از یک کلمه) با متن پایان‌نامه
عنوان انگلیسی	مطابقت با عنوان مندرج در صفحه عنوان انگلیسی پایان‌نامه/ رساله	- عدم انطباق کامل با متن پایان‌نامه - مطابق نبودن با عنوان فارسی	- وجود کاراکترهای اضافی مانند space - وجود غلط‌های املایی در متن عنوان
گروه تحصیلی	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله		مغایرت با نام گروه مندرج در صفحه عنوان پایان‌نامه/ رساله

^۱ توسط کارشناس اصلاح می‌گردد

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت				
نام فیلد	وضعیت مطلوب		عدم انطباق‌های بالقوه	
			نوع A	نوع B
رشته	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله			مغایرت با نام رشته مندرج در صفحه عنوان پایان‌نامه/ رساله
گرایش	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله			مغایرت با نام گرایش مندرج در صفحه عنوان پایان‌نامه/ رساله
تاریخ دفاع	مطابقت با متن پایان‌نامه/ رساله		غیر واقعی بودن (بعد از تاریخ ثبت در سامانه ثبت باشد)	مطابقت نداشتن با اطلاعات صفحه عنوان پایان‌نامه/ رساله
استاد راهنما/ استاد مشاور ^۱	نام	نباید کمتر از ۲ و بیشتر از ۲۵ حرف باشد مطابقت با متن پایان‌نامه/ رساله	- وجود یک نقص اساسی و یا کلیدی در نام/ نام‌خانوادگی - نام/ نام خانوادگی با صفحه عنوان پایان‌نامه مطابقت ندارد.	ناقص بودن نام/ نام‌خانوادگی در سامانه ثبت در موارد زیر - مواردی مانند ال، حروف اضافه و ... به درستی نوشته نشده است. - یک حرف اشتباه تایپ شده است (پ به جای ژ یا ...) - نام/ نام خانوادگی بصورت لاتین نوشته شده است - بخشی از یک نام/ نام خانوادگی چندبخشی نوشته نشده باشد.
	نام خانوادگی	نباید کمتر از ۲ و بیشتر از ۵۰ حرف باشد مطابقت با متن پایان‌نامه/ رساله		
	نام (انگلیسی)	نباید کمتر از ۲ و بیشتر از ۲۵ حرف	وجود یک نقص اساسی و یا کلیدی در نام/	یک حرف اشتباه تایپ شده است.

^۱ آقای ولی زاده چک نمیکنند

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت					
نام فیلد		وضعیت مطلوب		عدم انطباق‌های بالقوه	
چکیده(فارسی)	مطابقت با متن پایان‌نامه/ رساله	باشد		نام‌خانوادگی لاتین	نام/ نام خانوادگی بصورت فارسی نوشته شده است
		مطابقت با متن پایان‌نامه/ رساله			
		نباید کمتر از ۲ و بیشتر از ۵۰ حرف باشد			
	نام خانوادگی(انگلیسی)		رایانامه	نام/ نام خانوادگی با صفحه عنوان لاتین پایان‌نامه مطابقت ندارد.	نام/ نام خانوادگی بصورت فارسی نوشته شده است
	مطابقت با متن پایان‌نامه/ رساله				
	حتما باید فرمت زیر را رعایت کند: xxx@xxx.com				
شماره تلفن همراه		شماره ملی	در حال حاضر کنترل نمی‌شود.	نام/ نام خانوادگی بصورت فارسی نوشته شده است	
باید به صورت عدد و با فرمت صحیح باشد، یعنی ۰۹ شروع شود.					
مطابقت با الگوریتم شماره ملی		مطابقت با متن پایان‌نامه/ رساله		چکیده وارد نشده است.	نام/ نام خانوادگی بصورت فارسی نوشته شده است
چکیده(فارسی)					

دسته‌بندی عدم انطباق‌ها در اقسام اطلاعاتی سامانه ثبت			
نام فیلد	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
	نباید کمتر از ۲۰ و بیشتر از ۳۰۰۰ حرف باشد	- متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص و تفاوت اساسی دارد. - حروف کلمات مختلف به هم چسبیده است - عنوان در چکیده وارد شده است - در چکیده متن عربی یا انگلیسی وارد شده است. - ایمیل شخص در چکیده ذکر شده است.	- چکیده - متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص و تفاوت جزئی دارد.
چکیده(انگلیسی)	مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۳۰۰۰ حرف باشد	- چکیده وارد نشده است. - متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص و تفاوت اساسی دارد. - ایمیل شخص در چکیده ذکر شده است.	- قرار ندادن فرمول‌های شیمیایی و ریاضی در فیلد چکیده - متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص و تفاوت جزئی دارد.
نوع پایان نامه/ رساله	انتخاب فقط یکی از گزینه‌های بنیادی، کاربردی و توسعه‌ای	در حال حاضر کنترل نمی‌شود.	
فهرست مطالب	مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد	عدم همخوانی شماره صفحات با فهرست مطالب (کمتر از ۵ صفحه)	عدم همخوانی شماره صفحات با فهرست مطالب (بیشتر یا مساوی ۵ صفحه)
فهرست منابع فارسی	مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد	در حال حاضر کنترل نمی‌شود.	
فهرست منابع غیر	مطابقت با متن پایان‌نامه/ رساله		

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت			
نام فیلد	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
فارسی	نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد		
کلید واژه‌ها	کامل بودن فیلدها هر کلید واژه در یک فیلد باشد (به ازای هر کلید واژه یک فیلد) ^۱	- با کلیدواژه موجود در متن مطابقت داشته باشد - یک کلید واژه در چند فیلد تکرار شده باشد. - مطابقت نداشتن کلید واژه فارسی و لاتین - جابجا وارد شدن کلید واژه (انگلیسی به جای فارسی یا بالعکس)	در یک فیلد بیش از یک کلید واژه وارد شده باشد.
فایل pdf پایان نامه/ رساله	انطباق کامل فایل پایان‌نامه با قواعد نگارشی و محتوایی مرکز آموزشی آپلود الزامی فایل با فرمت pdf	آپلود فایلی با فرمتی غیر از pdf توضیحات: موارد و آیتم‌های قابل کنترل در فایل pdf در جدول دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل PDF) ذکر شده است.	-
فایل word پایان نامه/ رساله	انطباق کامل فایل پایان‌نامه با قواعد نگارشی و محتوایی مرکز آموزشی آپلود الزامی فایل با فرمت doc	آپلود فایلی با فرمتی غیر از doc توضیحات: موارد و آیتم‌های قابل کنترل در فایل ورد در جدول دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل ورد) ذکر شده است.	-
دیگر فایل‌های پایان نامه/ رساله	انطباق کامل فایل‌های ضمیمه (غیر از فایل اصلی پایان‌نامه) با فایل‌های الزام شده در مرکز آموزشی (بعنوان مثال فایل لاتکس در	آپلود فایلی با فرمتی غیر از txt	-

^۱ در اینگونه موارد فایل توسط کارشناس اصلاح میشود

طراحی مدل کیفیت فراداده در سامانه ثبت ... □ ۲۰۰

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام فیلد
نوع B	نوع A		
		رشته ریاضی و ...)	

جدول ۵-۲: طرح اولیه کیفیت فایل Pdf داده‌ها در سامانه ثبت

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل PDF)			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام محتوا
نوع B	نوع A		
ناقص بودن نام خانوادگی در متن فایل در مقایسه با اطلاعات وارد شده در سامانه ثبت^۱	- کامل نبودن آیتم‌های صفحه عنوان - کامل نبودن نام دانشگاه (در دانشگاه‌های پیام نور/ آزاد نام استان، شهر و ... کامل ذکر نشده است)	- کامل بودن آیتم‌های زیر در صفحه عنوان - لوگوی دانشگاه - نام دانشگاه (نام واحد) - دانشکده - مقطع (ارشد/ دکتری) - رشته و گرایش - عنوان پایان‌نامه - نام و نام خانوادگی نویسنده - نام استادراهنما - تاریخ دفاع	صفحه عنوان فارسی
- قرار گرفتن چکیده در دو صفحه (چینش نامناسب در صفحه) - ایمیل شخص در چکیده ذکر شده است. - مناسب نبودن فرمت در چکیده (مواردی مانند به هم چسبیدن کلمات، خوانا نبودن حروف و ...)	قرار نگرفتن چکیده در پایان‌نامه	- قرار گرفتن یک چکیده بصورت کامل در یک صفحه - چینش مناسب چکیده در وسط صفحه - مناسب بودن فرمت چکیده	چکیده فارسی
به هم چسبیدن کلمات (نامناسب بودن فرمت	نبود شماره صفحه	وجود شماره صفحه	فهرست مطالب

^۱ از آنجاییکه در جستجوهای آتی توسط کاربران، اطلاعات ثبت شده در فیلدهای سامانه مهم است این عدم انطباق میتواند جزئی تلقی گردد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل PDF)			
نام محتوا	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
	• مطابق بودن تعداد فصلها با متن • ترتیب منطقی میان فصلهای پایان‌نامه	• عدم وجود فصولی مانند نتیجه‌گیری • مطابق نبودن شماره صفحه فصل پایانی فایل با تعداد صفحات فایل	نوشتاری فهرست مطالب)^۱
فهرست جداول	• کامل بودن فهرست جداول • مطابقت میان شماره صفحه جداول ذکر شده در فهرست با محل قرارگیری آنها در متن	• عدم تطابق میان شماره صفحه ذکر شده در فهرست جداول و شماره صفحه قرار گرفته شده در متن • قرار ندادن جدول در متن در حالیکه در فهرست ذکر شده باشد	• در فهرست جداول فقط شماره آمده است (عنوان ذکر نشده است)
فهرست اشکال	• کامل بودن فهرست اشکال • مطابقت میان شماره صفحه اشکال ذکر شده در فهرست با محل قرارگیری آنها در متن	• عدم تطابق میان شماره صفحه ذکر شده در فهرست اشکال و شماره صفحه قرار گرفته شده در متن • قرار ندادن شکل در متن در حالیکه در فهرست ذکر شده باشد	• در فهرست اشکال فقط شماره آمده است (عنوان ذکر نشده است)
فهرست نمودار	• کامل بودن فهرست نمودارها • مطابقت میان شماره صفحه نمودارها ذکر شده در فهرست با محل قرارگیری آنها در متن	• عدم تطابق میان شماره صفحه ذکر شده در فهرست نمودار و شماره صفحه قرار گرفته شده در متن • قرار ندادن نمودار در متن در حالیکه در فهرست ذکر شده باشد	• در فهرست نمودار فقط شماره آمده است (عنوان ذکر نشده است)
متن پایان‌نامه	• کامل بودن متن پایان‌نامه • فایل پایان‌نامه بدون پس‌نقش (واترمارک) باشد.	• ناقص بودن پایان‌نامه (فصلها کم است) • قرار ندادن پیوست در پایان‌نامه در حالیکه در فهرست وجود دارد	• وجود صفحات سفید (خصوصاً در سی صفحه اول)

^۱ اگر فایل ورد ارسالی توسط نویسنده مناسب باشد، از آن برای اصلاح فایل PDF استفاده خواهد شد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل PDF)			
نام محتوا	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
	فایل پایان‌نامه بدون آدرس و مشخصات شخصی دانشجو و استاد راهنما باشد.	- نداشتن شماره صفحه - مناسب نبودن فونت پایان‌نامه - قفل شدن فایل پی‌دی‌اف توسط نویسنده^۱	
جداول	جداول مندرج در متن دارای عنوان و شماره باشند		نبود عنوان و شماره برای جداول
اشکال	اشکال مندرج در متن دارای عنوان و شماره باشند		نبود عنوان و شماره برای اشکال
نمودار	نمودارهای مندرج در متن دارای عنوان و شماره باشند		نبود عنوان و شماره برای نمودارها
مراجع	وجود بخش مراجع در متن	نبود بخش مراجع	وجود نقص در فونت و متن مراجع
صفحه عنوان انگلیسی	کامل بودن آیتم‌های زیر در صفحه عنوان - لوگوی دانشگاه - نام دانشگاه (نام واحد) - نام دانشکده - مقطع (ارشد/ دکتری) - رشته و گرایش	- کامل نبودن آیتم‌های صفحه عنوان - کامل نبودن نام دانشگاه (در دانشگاه‌های پیام نور/ آزاد نام استان، شهر و ... کامل ذکر نشده است)	ناقص بودن نام خانوادگی در متن فایل در مقایسه با اطلاعات وارد شده در سامانه ثبت^۲

^۱ در صورت کامل بودن فایل ورد از آن پی‌دی‌اف تهیه شده و آپلود می‌گردد

^۲ از آنجاییکه در جستجوهای آتی توسط کاربران، اطلاعات ثبت شده در فیلدهای سامانه مهم است این عدم انطباق می‌تواند جزئی تلقی گردد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل PDF)			
نام محتوا	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
	- عنوان پایان‌نامه - نام و نام خانوادگی نویسنده - نام استادراهنما - تاریخ دفاع		
چکیده انگلیسی	- قرار گرفتن یک چکیده بصورت کامل در یک صفحه - چینش مناسب چکیده در وسط صفحه	قرار نگرفتن چکیده در پایان‌نامه	- قرار گرفتن چکیده در دو صفحه (چینش نامناسب در صفحه) - ایمیل شخص در چکیده ذکر شده است. - مناسب نبودن فرمت در چکیده (مواردی مانند به هم چسبیدن کلمات، خوانا نبودن حروف و ...)

۵-۱۲ ارائه طرح نهایی کیفیت فراداده

بمنظور ارائه طرح نهایی کیفیت فراداده‌ها در سامانه ثبت پس از مصاحبه تفصیلی با کارشناسان اداره ثبت و فراهم‌آوری اطلاعات طرح اولیه کیفیت فراداده‌ها مطابق آنچه در بخش ۵-۱ ارائه گردید مستند شد. بمنظور بررسی محدودیت‌های اجرایی این طرح و نیز همراستایی آن با سیاست‌های کلان در حوزه ثبت، ویرایش و اشاعه پایان‌نامه/رساله و نیز رفع نواقص احتمالی این طرح کارگروه بررسی روایی طرح تشکیل گردید. در این کارگروه اعضای زیر حضور داشتند.

- معاون محترم اطلاعات علم و فناوری ایران، جناب آقای دکتر جلالی‌منش
- مدیر محترم سازماندهی و تحلیل اطلاعات، سرکار خانم دهرایی
- رئیس محترم اداره ثبت و فراهم‌آوری اطلاعات، جناب آقای محمودیان
- رئیس محترم اداره سامانه‌های اطلاعاتی، سرکار خانم شمس‌ی
- همکار اصلی پروژه طراحی و توسعه سامانه ثبت پایان‌نامه/رساله‌ها، سرکار خانم رضایی

برمبنای مصاحبه‌های انجام گرفته با خبرگان حوزه سامانه ثبت، اصلاحات مختلفی در طرح صورت گرفته و طرح نهایی کیفیت فراداده‌ها مطابق آنچه در بخش بعدی ارائه شده است، تهیه گردید.

۵-۱-۲ طرح نهایی کیفیت فراداده در سامانه ثبت پایان‌نامه/رساله‌های دانش‌آموختگان داخل کشور

جدول ۵-۳: طرح نهایی کیفیت فراداده‌ها در سامانه ثبت

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان داخل کشور			
نام فیلد	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A ^۱	نوع B ^۲
نام	مطابقت با متن پایان‌نامه/ رساله	• ناقص بودن نام در سامانه ثبت در موارد زیر • مواردی مانند ال، حروف اضافه و ... به درستی نوشته نشده است. • یک حرف اشتباه تایپ شده است (پ به جای ژ یا ...) • مواردی مانند سید، سیده، میر در ابتدای اسم آورده نشده است. • مواردی مانند مهندس، دکتر، حجه الاسلام در ابتدای اسم آورده شده است. • نام بصورت لاتین نوشته شده است. • بطور کلی نام با صفحه عنوان پایان‌نامه مطابقت ندارد. • بخشی از یک نام چندبخشی نوشته نشده باشد.	توضیحات: • از آنجاکه امکان اصلاح نام و نام خانوادگی در سامانه وجود ندارد در صورت مشاهده هرگونه عدم انطباقی مدرک به دانشجو عودت داده خواهد شد و لذا عدم انطباق نوع B در این موارد وجود ندارد. • مبنای اصلاحات در بخش نام و نام خانوادگی متن پایان‌نامه خواهد بود.
نام خانوادگی	مطابقت با متن پایان‌نامه/ رساله	• ناقص بودن نام خانوادگی در سامانه ثبت	

^۱ **عدم انطباق نوع A:** این دسته از عدم انطباق‌ها شامل مواردی است که بحرانی بوده و شخص کنترل‌کننده بایستی در صورت مشاهده یکی از این موارد مدرک را به دانشجو عودت دهد.

^۲ **عدم انطباق نوع B:** این دسته از عدم انطباق‌ها شامل مواردی بوده که غیر بحرانی بوده و در صورت مشاهده این موارد کنترل‌کننده می‌تواند اصلاحات را در سامانه انجام دهد. وجود چندین عدم انطباق نوع B می‌تواند منجر به عدم انطباق نوع A در مدرک شده و عودت داده شود.

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان داخل کشور

عدم انطباق‌های بالقوه		وضعیت مطلوب	نام فیلد
نوع A ^۱	نوع B ^۲		
• مواردی مانند ال، حروف اضافه و ... به درستی نوشته نشده است. • یک حرف اشتباه تایپ شده است (پ به جای ژ یا ...) • نام خانوادگی بصورت لاتین نوشته شده است • بطور کلی نام خانوادگی با صفحه عنوان پایان‌نامه مطابقت ندارد. • بخشی از یک نام خانوادگی چندبخشی نوشته نشده باشد. • در صورتیکه پسوند اگر معنادار نباشد (مانند پور، زاده و ...) نیم‌فاصله و اگر معنادار باشد فاصله داشته باشد.			
• رایانامه از آنجاکه که مبنای مکاتبات و تعامل با دانشجویان است لذا بصورت غیر مستقیم و در عمل مورد صحه‌گذاری قرار می‌گیرد و کنترلی روی آن انجام نخواهد شد.		مطابقت با فرمت زیر: xxx@xxx.com	رایانامه
• درحال حاضر با توجه به عدم امکان استفاده از سرویس ثبت احوال امکان کنترل شماره ملی وجود ندارد و کنترلی روی آن انجام نخواهد شد. • الگوریتم و رابطه میان اعداد کدملی در حال حاضر در این بخش چک خواهد شد.		مطابقت با الگوریتم شماره ملی مطابقت با وب سرویس تطبیق کد ملی و نام و نام خانوادگی در سامانه ثبت احوال	شماره ملی
• این فیلد توسط نماینده دانشگاه کنترل خواهد شد.		مطابقت با مشخصات دانشجو دربانک‌های اطلاعاتی داخلی مرکز آموزشی/ پژوهشی	شماره دانشجویی

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان داخل کشور

نام فیلد	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A ^۱	نوع B ^۲
مقطع تحصیلی	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله	برگرداندن مدرک در صورت عدم امکان تماس تلفنی با کاربر و عدم اطمینان از مقطع تحصیلی درج شده در سیستم	وجود هرگونه مغایرت میان صفحه عنوان و فیلد تکمیل شده در سامانه ثبت در صورت اطمینان از مغایرت مدرک با اطلاعات وارد شده در سیستم (در این حالت کاربر سیستم در صورت امکان با دانشجوی مربوطه تماس می‌گیرد و اصلاح مربوطه را انجام می‌دهد)
وابستگی سازمانی	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله	عدم تطابق با صفحه عنوان پایان‌نامه/ رساله	عدم تطابق میان وابستگی سازمانی و نام دانشگاه (بعنوان مثال وابستگی سازمانی وزارت علوم انتخاب شده است در حالیکه دانشگاه آزاد بعنوان نام موسسه آموزشی ذکر شده است)
موسسه آموزشی/پژوهشی	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله	- در دانشگاه‌هایی مانند دانشگاه آزاد و پیام نور واحدها مشخص نشده باشد. - مغایرت با صفحه عنوان پایان‌نامه/ رساله	توضیحات: - از آنجاکه پایان‌نامه/ رساله پس از تایید ایرانداک به کارتابل نماینده دانشگاه می‌رود و در صورت انتخاب اشتباه موسسه به کارتابل فرد دیگری رفته و مشکلات اجرایی ایجاد مینماید لذا در این آیتم بایستی دقت فراوان صورت گرفته و در صورت هرگونه مغایرت به دانشجو عودت داده شود. - در حوزه علمیه چون لیست دانشگاه وجود ندارد، بصورت عنوان "دانشگاه‌های وابسته به حوزه علمیه" انتخاب می‌شود.

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان داخل کشور

نام فیلد	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A ^۱	نوع B ^۲
دانشکده/پژوهشکده	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/رساله	نام دانشکده/پژوهشکده بصورت دستی توسط دانشجو در سامانه ثبت وارد شده است در حالیکه در لیست کرکره‌ای موجود است.	- مغایرت با نام دانشکده/پژوهشکده مندرج در صفحه عنوان پایان‌نامه/رساله - در صورتیکه توسط دانشجو نام دانشکده/پژوهشکده بصورت دستی انتخاب شده باشد بایستی اطمینان از صحت آن توسط نماینده مربوطه در واحد سازماندهی صورت گیرد.
زبان اصلی پایان نامه	مطابقت با متن پایان‌نامه/رساله (بایستی یکی از زبانهای فارسی، انگلیسی، فرانسوی، آلمانی یا عربی باشد).	هرگونه مغایرت میان زبان انتخاب شده و متن نوشته شده در پایان‌نامه/رساله	-
عنوان فارسی	مطابقت با عنوان مندرج در صفحه اول پایان‌نامه/رساله	وجود مغایرت اساسی (چند کلمه یا بیشتر) میان فیلد عنوان و متن پایان‌نامه	- وجود کاراکترهای اضافی مانند space - وجود غلط‌های املایی در متن عنوان - عدم انطباق کامل (بیش از یک کلمه) با متن پایان‌نامه
عنوان انگلیسی	مطابقت با عنوان مندرج در صفحه عنوان انگلیسی پایان‌نامه/رساله	- عدم انطباق کامل با متن صفحه عنوان انگلیسی پایان‌نامه - مطابق نبودن با عنوان فارسی	- وجود کاراکترهای اضافی مانند space - وجود غلط‌های املایی در متن عنوان
گروه تحصیلی	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/رساله		مغایرت با نام گروه مندرج در صفحه عنوان پایان‌نامه/رساله

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان داخل کشور

نام فیلد	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A ^۱	نوع B ^۲
رشته	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله		مغایرت با نام رشته مندرج در صفحه عنوان پایان‌نامه/ رساله
گرایش	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله		مغایرت با نام گرایش مندرج در صفحه عنوان پایان‌نامه/ رساله
تاریخ دفاع	مطابقت با متن پایان‌نامه/ رساله	غیر واقعی بودن (تاریخ دفاع دانشجو بعد از تاریخ ثبت در سامانه ثبت باشد)	مطابقت نداشتن با اطلاعات صفحه عنوان پایان‌نامه/ رساله
استاد راهنما/ استاد مشاور	نام	- ناقص بودن نام در سامانه ثبت در موارد زیر - مواردی مانند ال، حروف اضافه و ... به درستی نوشته نشده است. - یک حرف اشتباه تایپ شده است (پ به جای ژ یا ...) - مواردی مانند سید، سیده، میر در ابتدای اسم آورده نشده است.	- ناقص بودن نام/ نام خانوادگی در سامانه ثبت در موارد زیر - مواردی مانند ال، حروف اضافه و ... به درستی نوشته نشده است.
	نام خانوادگی	- مواردی مانند مهندس، دکتر، حجه الاسلام در ابتدای اسم آورده شده است. - نام بصورت لاتین نوشته شده است. - بطور کلی نام با صفحه عنوان پایان‌نامه مطابقت ندارد. - بخشی از یک نام چندبخشی نوشته نشده باشد. - وجود یک نقص اساسی و یا کلیدی در نام/ نام خانوادگی نام/ نام خانوادگی با صفحه عنوان پایان‌نامه مطابقت ندارد.	- یک حرف اشتباه تایپ شده است (پ به جای ژ یا ...) - نام/ نام خانوادگی بصورت لاتین نوشته شده است - بخشی از یک نام/ نام خانوادگی چندبخشی نوشته نشده باشد.

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان داخل کشور

عدم انطباق‌های بالقوه		وضعیت مطلوب		نام فیلد
نوع B ^۲	نوع A ^۱			
- یک حرف اشتباها تایپ شده است. - نام/ نام خانوادگی بصورت فارسی نوشته شده است - بخشی از یک نام/ نام خانوادگی چندبخشی نوشته نشده باشد.	- وجود یک نقص اساسی و یا کلیدی در نام/ نام خانوادگی لاتین - نام/ نام خانوادگی با صفحه عنوان لاتین پایان‌نامه مطابقت ندارد.	نباید کمتر از ۲ و بیشتر از ۲۵ حرف باشد	نام (انگلیسی)	
		مطابقت با متن پایان‌نامه/ رساله	نام خانوادگی (انگلیسی)	
- در حال حاضر کنترل دستی بر روی این فیلدها صورت نمی‌گیرد. - کنترل فرمت بصورت اتومات توسط سامانه انجام می‌گردد و اجازه تکمیل این فیلد برای دانشجو در صورت عدول از این فرمت داده نمی‌شود.		نباید کمتر از ۲ و بیشتر از ۵۰ حرف باشد	نام خانوادگی (انگلیسی)	
		مطابقت با متن پایان‌نامه/ رساله	رایانامه	
		حتما باید فرمت زیر را رعایت کند: xxx@xxx.com	شماره تلفن همراه	
		باید به صورت عدد و با فرمت صحیح باشد، یعنی ۰۹ شروع شود.	شماره ملی	
		مطابقت با الگوریتم شماره		

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان داخل کشور

نام فیلد		وضعیت مطلوب		عدم انطباق‌های بالقوه	
				نوع A ^۱	نوع B ^۲
		ملی			
چکیده(فارسی) ^۱		مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۳۰۰۰ حرف باشد		- چکیده وارد نشده است. - متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص و تفاوت اساسی دارد. - حروف کلمات مختلف به هم چسبیده است - عنوان در چکیده وارد شده است - در چکیده متن عربی یا انگلیسی وارد شده است. - ایمیل شخص در چکیده ذکر شده است.	- قرار ندادن فرمول‌های شیمیایی و ریاضی در فیلد چکیده - متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص و تفاوت جزئی دارد.
چکیده(انگلیسی)		مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۳۰۰۰ حرف باشد		- چکیده وارد نشده است. - متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص و تفاوت اساسی دارد. - حروف کلمات مختلف به هم چسبیده است - ایمیل شخص در چکیده ذکر شده است.	- قرار ندادن فرمول‌های شیمیایی و ریاضی در فیلد چکیده - متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص و تفاوت جزئی دارد.
نوع پایان نامه/ رساله		انتخاب فقط یکی از گزینه‌های بنیادی، کاربردی و توسعه‌ای		در حال حاضر کنترل نمی‌شود.	
فهرست مطالب		مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰ حرف		- عدم همخوانی شماره صفحات با فهرست مطالب - به هم‌ریختگی ساختار و ظاهر فهرست مطالب	

^۱ حتی در پایان نامه به زبان خارجی چکیده و صفحه عنوان فارسی الزامی است.

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان داخل کشور

عدم انطباق‌های بالقوه		وضعیت مطلوب	نام فیلد
نوع B ^۲	نوع A ^۱		
• چسبیدگی کلمات در فهرست مطالب		باشد	
در حال حاضر باتوجه به اینکه در متن پایان‌نامه/ رساله این موضوع چک می‌گردد لذا در سطح فراداده این موضوع کنترل نمی‌شود.		مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد	فهرست منابع فارسی
		مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد	فهرست منابع غیر فارسی
• در یک فیلد بیش از یک کلید واژه وارد شده باشد.	- با کلیدواژه موجود در متن مطابقت داشته باشد - یک کلید واژه در چند فیلد تکرار شده باشد. - مطابقت نداشتن کلید واژه فارسی و لاتین - جابجا وارد شدن کلید واژه (انگلیسی به جای فارسی یا بالعکس) - فینگلیش نوشتن کلید واژه انگلیسی بر مبنای کلیدواژه فارسی	کامل بودن فیلدها هر کلید واژه در یک فیلد باشد (به ازای هر کلید واژه یک فیلد)	کلید واژه‌ها
-	آپلود فایلی با فرمتی غیر از pdf توضیحات: موارد و آیتم‌های قابل کنترل در فایل pdf در جدول دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل PDF) ذکر شده است.	انطباق کامل فایل پایان‌نامه با قواعد نگارشی و محتوایی مرکز آموزشی آپلود الزامی فایل با فرمت pdf	فایل pdf پایان نامه/ رساله
-	آپلود فایلی با فرمتی غیر از doc	انطباق کامل فایل پایان‌نامه با قواعد نگارشی و محتوایی مرکز آموزشی	فایل word پایان نامه/ رساله

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان داخل کشور

نام فیلد	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A ^۱	نوع B ^۲
	آپلود الزامی فایل با فرمت doc	توضیحات: موارد و آیتم‌های قابل کنترل در فایل ورد در جدول دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل ورد) ذکر شده است.	
دیگر فایل‌های پایان‌نامه / رساله مانند Latex و ...	انطباق کامل فایل‌های ضمیمه (غیر از فایل اصلی پایان‌نامه) با فایل‌های الزام شده در مرکز آموزشی (بعنوان مثال فایل لاتکس در رشته ریاضی و ...)	• آپلود فایلی با فرمتی غیر از txt	-

جدول ۵-۴: طرح نهایی کیفیت فراداده‌ها در سامانه ثبت (فایل Pdf)

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه سامانه ثبت دانشجویان داخل کشور (فایل PDF) ^۱			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام محتوا
نوع B	نوع A		
	- کامل نبودن آیتم‌های صفحه عنوان - کامل نبودن نام دانشگاه (در دانشگاه‌های پیام نور/ آزاد نام استان، شهر و ... کامل ذکر نشده است) - ناقص بودن نام خانوادگی در متن فایل در مقایسه با اطلاعات وارد شده در سامانه ثبت^۲	- کامل بودن آیتم‌های زیر در صفحه عنوان - لوگوی دانشگاه - نام دانشگاه (نام واحد) - دانشکده - مقطع (ارشد/ دکتری) - رشته و گرایش - عنوان پایان‌نامه - نام و نام خانوادگی نویسنده - نام استاد راهنما - نام استاد مشاور - تاریخ دفاع	صفحه عنوان فارسی
چینش نامناسب در صفحه	- قرار نگرفتن چکیده در پایان‌نامه - ایمیل یا شماره تلفن شخص در چکیده ذکر شده است. - مناسب نبودن فرمت در چکیده (مواردی مانند به هم چسبیدن کلمات، خوانا نبودن حروف و ...)	- قرار گرفتن یک چکیده بصورت کامل در یک صفحه - چینش مناسب چکیده در وسط صفحه - مناسب بودن فرمت چکیده	چکیده فارسی

^۱ کلیه موارد و آیتمهای کنترلی در فایل ورد هم مشابه فایل پی دی اف میباشد

^۲ از آنجاییکه در جستجوهای آتی توسط کاربران، اطلاعات ثبت شده در فیلدهای سامانه مهم است این عدم انطباق میتواند جزئی تلقی گردد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه سامانه ثبت دانشجویان داخل کشور (فایل PDF)^۱

عدم انطباق‌های بالقوه		وضعیت مطلوب	نام محتوا
نوع B	نوع A		
	- نبود شماره صفحه - عدم وجود فصولی مانند نتیجه گیری - مطابق نبودن شماره صفحه فصل پایانی فایل با تعداد صفحات فایل - به هم چسبیدن کلمات (نامناسب بودن فرمت نوشتاری فهرست مطالب)^۱	- وجود شماره صفحه - مطابق بودن تعداد فصلها با متن - وجود ترتیب منطقی میان فصلهای پایان‌نامه	فهرست مطالب
	- عدم تطابق میان شماره صفحه ذکر شده در فهرست جداول و شماره صفحه قرار گرفته شده در متن - قرار ندادن جدول در متن در حالیکه در فهرست ذکر شده باشد. - در فهرست جداول فقط شماره آمده است (عنوان ذکر نشده است)	- کامل بودن فهرست جداول - مطابقت میان شماره صفحه جداول ذکر شده در فهرست با محل قرارگیری آنها در متن	فهرست جداول
	- عدم تطابق میان شماره صفحه ذکر شده در فهرست اشکال و شماره صفحه قرار گرفته شده در متن - قرار ندادن شکل در متن در حالیکه در فهرست ذکر شده باشد. - در فهرست اشکال فقط شماره آمده است (عنوان ذکر نشده است)	- کامل بودن فهرست اشکال - مطابقت میان شماره صفحه اشکال ذکر شده در فهرست با محل قرارگیری آنها در متن	فهرست اشکال
	عدم تطابق میان شماره صفحه ذکر شده در فهرست نمودار و	کامل بودن فهرست نمودارها	فهرست نمودار

^۱ اگر فایل ورد ارسالی توسط نویسنده مناسب باشد، از آن برای اصلاح فایل PDF استفاده خواهد شد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه سامانه ثبت دانشجویان داخل کشور (فایل PDF)^۱

نام محتوا	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
	• مطابقت میان شماره صفحه نمودارها ذکر شده در فهرست با محل قرارگیری آنها در متن	• شماره صفحه قرار گرفته شده در متن • قرار ندادن نمودار در متن در حالیکه در فهرست ذکر شده باشد. • در فهرست نمودار فقط شماره آمده است (عنوان ذکر نشده است)	
متن پایان‌نامه	• کامل بودن متن پایان‌نامه • فایل پایان‌نامه بدون پس‌نقش (واترمارک) باشد. • فایل پایان‌نامه بدون آدرس و مشخصات شخصی دانشجو و استاد راهنما باشد.	• ناقص بودن پایان‌نامه (فصلها کم است) • قرار ندادن پیوست در پایان‌نامه در حالیکه در فهرست وجود دارد • نداشتن شماره صفحه • به هم ریختگی فونت پایان‌نامه • قفل شدن فایل پی‌دی‌اف توسط نویسنده^۱ • فایل ارسالی دارای واترمارک باشد.	• وجود صفحات سفید^۲ (خصوصاً در سی صفحه اول)
جداول	جداول مندرج در متن دارای عنوان و شماره باشند.	• نبود عنوان یا شماره برای جداول	
اشکال	اشکال مندرج در متن دارای عنوان و شماره باشند	• نبود عنوان و شماره برای اشکال	
نمودار	• نمودارهای مندرج در متن دارای عنوان و شماره باشند	• نبود عنوان و شماره برای نمودارها	

^۱ در صورت کامل بودن فایل ورد از آن پی‌دی‌اف تهیه شده و آپلود می‌گردد

^۲ در صورتی که صفحه سفید شماره صفحه داشته باشد، حذف نمی‌گردد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه سامانه ثبت دانشجویان داخل کشور (فایل PDF)^۱

نام محتوا	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
منابع	وجود بخش منابع در متن	نبود بخش منابع	وجود به هم‌ریختگی در متن منابع
پیوست	درج شدن پیوست در پایان متن در صورتیکه در متن یا فهرست به آن اشاره شده باشد.	مشاهده نشدن بخشی یا کل پیوست‌ها در پایان متن در صورتیکه در متن یا فهرست به آن اشاره شده باشد.	
صفحه عنوان انگلیسی	کامل بودن آیتم‌های زیر در صفحه عنوان - لوگوی دانشگاه - نام دانشگاه (نام واحد) - دانشکده - مقطع (ارشد/ دکتری) - رشته و گرایش - عنوان پایان‌نامه - نام و نام خانوادگی نویسنده - نام استاد راهنما - نام استاد مشاور - تاریخ دفاع	- کامل نبودن آیتم‌های صفحه عنوان - کامل نبودن نام دانشگاه (در دانشگاه‌های پیام نور/ آزاد نام استان، شهر و ... کامل ذکر نشده است) - ناقص بودن نام خانوادگی در متن فایل در مقایسه با اطلاعات وارد شده در سامانه ثبت^۱	

^۱ از آنجاییکه در جستجوهای آتی توسط کاربران، اطلاعات ثبت شده در فیلدهای سامانه مهم است این عدم انطباق میتواند جزئی تلقی گردد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه سامانه ثبت دانشجویان داخل کشور (فایل PDF)^۱

نام محتوا	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
چکیده انگلیسی	- قرار گرفتن یک چکیده بصورت کامل در یک صفحه - چینش مناسب چکیده در وسط صفحه - مناسب بودن فرمت چکیده	- قرار نگرفتن چکیده در پایان‌نامه - ایمیل یا شماره تلفن شخص در چکیده ذکر شده است. - مناسب نبودن فرمت در چکیده (مواردی مانند به هم چسبیدن کلمات، خوانا نبودن حروف و ...)	چینش نامناسب در صفحه

جدول ۵-۵: طرح نهایی کیفیت فراداده‌ها در سامانه ثبت (فایل Word)

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه سامانه ثبت دانشجویان داخل کشور (فایل WORD)^۱

نام محتوا	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
صفحه عنوان فارسی	- کامل بودن آیتم‌های زیر در صفحه عنوان - لوگوی دانشگاه - نام دانشگاه (نام واحد) - دانشکده - مقطع (ارشد/ دکتری) - رشته و گرایش	- کامل نبودن آیتم‌های صفحه عنوان - کامل نبودن نام دانشگاه (در دانشگاه‌های پیام نور/ آزاد نام استان، شهر و ... کامل ذکر نشده است) - ناقص بودن نام خانوادگی در متن فایل در مقایسه با اطلاعات وارد شده در سامانه ثبت^۲	

^۱ اغلب موارد و آیتمهای کنترلی در فایل ورد هم مشابه فایل پی دی اف میباشد.

همچنین دو متاتگست (فرامتن) شامل فهرست مطالب و فهرست منابع ضروری بوده و بایستی با دقت بیشتری کنترل گردد.

^۲ از آنجاییکه در جستجوهای آتی توسط کاربران، اطلاعات ثبت شده در فیلدهای سامانه مهم است این عدم انطباق میتواند جزئی تلقی گردد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه سامانه ثبت دانشجویان داخل کشور (فایل WORD^۱)

عدم انطباق‌های بالقوه		وضعیت مطلوب	نام محتوا
نوع B	نوع A		
		- عنوان پایان‌نامه - نام و نام خانوادگی نویسنده - نام استادراهنما - نام استاد مشاور - تاریخ دفاع	
چینش نامناسب چکیده در صفحه	- قرار نگرفتن چکیده در پایان‌نامه - ایمیل یا شماره تلفن شخص در چکیده ذکر شده است. - نامناسب نبودن فرمت در چکیده (مواردی مانند به هم چسبیدن کلمات، خوانا نبودن حروف و ...)	- قرار گرفتن یک چکیده بصورت کامل در یک صفحه - چینش مناسب چکیده در وسط صفحه - نامناسب بودن فرمت چکیده	چکیده فارسی
	- نبود شماره صفحه - عدم وجود فصولی مانند نتیجه گیری - مطابق نبودن شماره صفحه فصل پایانی فایل با تعداد صفحات فایل - به هم چسبیدن کلمات (نامناسب بودن فرمت نوشتاری فهرست مطالب)^۱	- وجود شماره صفحه - مطابق بودن تعداد فصلها با متن - وجود ترتیب منطقی میان فصلهای پایان‌نامه	فهرست مطالب
	- عدم تطابق میان شماره صفحه ذکر شده در فهرست جداول و شماره صفحه قرار گرفته شده در متن - قرار ندادن جدول در متن در حالیکه در فهرست ذکر شده باشد. - در فهرست جداول فقط شماره آمده است (عنوان ذکر نشده است)	- کامل بودن فهرست جداول - مطابقت میان شماره صفحه جداول ذکر شده در فهرست با محل قرارگیری آنها در متن	فهرست جداول

^۱ اگر فایل ورد ارسالی توسط نویسنده مناسب باشد، از آن برای اصلاح فایل PDF استفاده خواهد شد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه سامانه ثبت دانشجویان داخل کشور (فایل WORD^۱)

عدم انطباق‌های بالقوه		وضعیت مطلوب	نام محتوا
نوع B	نوع A		
	- عدم تطابق میان شماره صفحه ذکر شده در فهرست اشکال و شماره صفحه قرار گرفته شده در متن - قرار ندادن شکل در متن در حالیکه در فهرست ذکر شده باشد. - در فهرست اشکال فقط شماره آمده است (عنوان ذکر نشده است)	- کامل بودن فهرست اشکال - مطابقت میان شماره صفحه اشکال ذکر شده در فهرست با محل قرارگیری آنها در متن	فهرست اشکال
	- عدم تطابق میان شماره صفحه ذکر شده در فهرست نمودار و شماره صفحه قرار گرفته شده در متن - قرار ندادن نمودار در متن در حالیکه در فهرست ذکر شده باشد. - در فهرست نمودار فقط شماره آمده است (عنوان ذکر نشده است)	- کامل بودن فهرست نمودارها - مطابقت میان شماره صفحه نمودارها ذکر شده در فهرست با محل قرارگیری آنها در متن	فهرست نمودار
وجود صفحات سفید^۲ (خصوصاً در سی صفحه اول)	- ناقص بودن پایان‌نامه (فصلها کم است) - قرار ندادن پیوست در پایان‌نامه در حالیکه در فهرست وجود دارد - نداشتن شماره صفحه - به هم ریختگی فونت پایان‌نامه - قفل شدن فایل پی‌دی اف توسط نویسنده^۱ - فایل ارسالی دارای واترمارک باشد.	- کامل بودن متن پایان‌نامه - فایل پایان نامه بدون پس‌نقش (واترمارک) باشد. - فایل پایان نامه بدون آدرس و مشخصات شخصی دانشجو و استاد راهنما باشد.	متن پایان‌نامه
	نبود عنوان یا شماره برای جداول	جداول مندرج در متن دارای عنوان و شماره باشند.	جداول
	نبود عنوان و شماره برای اشکال	اشکال مندرج در متن دارای عنوان و	اشکال

^۱ در صورت کامل بودن فایل ورد از آن پی دی اف تهیه شده و آپلود میگردد

^۲ در صورتی که صفحه سفید شماره صفحه داشته باشد، حذف نمی‌گردد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه سامانه ثبت دانشجویان داخل کشور (فایل WORD^۱)

نام محتوا	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A	نوع B
	شماره باشند		
نمودار	نمودارهای مندرج در متن دارای عنوان و شماره باشند	نبود عنوان و شماره برای نمودارها	
منابع	وجود بخش منابع در متن	نبود بخش منابع	وجود به هم‌ریختگی در متن منابع
پیوست	درج شدن پیوست در پایان متن در صورتیکه در متن یا فهرست به آن اشاره شده باشد.	مشاهده نشدن بخشی یا کل پیوست‌ها در پایان متن در صورتیکه در متن یا فهرست به آن اشاره شده باشد.	
صفحه عنوان انگلیسی	کامل بودن آیتم‌های زیر در صفحه عنوان - لوگوی دانشگاه - نام دانشگاه (نام واحد) - دانشکده - مقطع (ارشد/ دکتری) - رشته و گرایش - عنوان پایان‌نامه - نام و نام خانوادگی نویسنده - نام استادراهنما - نام استاد مشاور	- کامل نبودن آیتم‌های صفحه عنوان - کامل نبودن نام دانشگاه (در دانشگاه‌های پیام نور/ آزاد نام استان، شهر و ... کامل ذکر نشده است) - ناقص بودن نام خانوادگی در متن فایل در مقایسه با اطلاعات وارد شده در سامانه ثبت^۱	

^۱ از آنجاییکه در جستجوهای آتی توسط کاربران، اطلاعات ثبت شده در فیلدهای سامانه مهم است این عدم انطباق میتواند جزئی تلقی گردد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه سامانه ثبت دانشجویان داخل کشور (فایل WORD^۱)

عدم انطباق‌های بالقوه		وضعیت مطلوب	نام محتوا
نوع B	نوع A		
		تاریخ دفاع	
چینش نامناسب در صفحه	- قرار نگرفتن چکیده در پایان‌نامه - ایمیل یا شماره تلفن شخص در چکیده ذکر شده است. - مناسب نبودن فرمت در چکیده (مواردی مانند به هم چسبیدن کلمات، خوانا نبودن حروف و ...)	- قرار گرفتن یک چکیده بصورت کامل در یک صفحه - چینش مناسب چکیده در وسط صفحه - مناسب بودن فرمت چکیده	چکیده انگلیسی

۵-۲، طرح نهایی کیفیت فراداده در سامانه ثبت پایان نامه/ رساله‌های دانش‌آموختگان خارج از کشور
سامانه ثبت پایان نامه/ رساله‌های دانش‌آموختگان خارج از کشور بمنظور ثبت و اشاعه پایان نامه/ رساله‌هایی
طراحی شده است که دانشجویان ایرانی در دانشگاه‌های خارج از کشور تدوین نموده و براساس آن فارغ
التحصیل شده‌اند. این سامانه دارای فیلدهایی است که تاحد بسیار زیادی مشابه سامانه ثبت پایان نامه/
رساله‌های دانش‌آموختگان داخل کشور است. کنترل‌های صورت گرفته در این سامانه به جز موارد زیر عینا
مانند ثبت پایان نامه داخل کشور است.

- به جای فیلد وابستگی سازمانی، فیلد کشور محل تحصیل وجود دارد که فهرست ثابتی است که دانشجویان از آن نام کشور محل تحصیل را وارد می‌کند؛
- فیلد دانشگاه هم مانند وضعیت فعلی است که به فیلد کشور متصل است و فهرست دانشگاه‌های مورد تأیید وزارت علوم در آن درج شده و دانشجو باید از این فهرست انتخاب کند؛
- فیلد دانشکده هم به صورت دستی از سوی دانشجو ورود اطلاعات می‌شود.

در این فرایند ایرانداک نقشی بر عهده نداشته و تأیید کننده فرایند هم اداره کل دانش‌آموختگان وزارت عتف بوده است.

در جداول ۵-۶ تا ۵-۷ بر مبنای نظر خبرگان مشابه آنچه در بخش قبلی مورد اشاره قرار گرفت طرح کیفیت پایان نامه/ رساله‌های دانش‌آموختگان خارج از کشور قابل مشاهده است.

جدول ۵-۶: طرح نهایی کیفیت فراداده‌ها در سامانه ثبت دانشجویان خارجی

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان خارج کشور			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام فیلد
نوع B ^۲	نوع A ^۱		
توضیحات: - از آنجاکه امکان اصلاح نام و نام خانوادگی در سامانه وجود ندارد در صورت مشاهده هرگونه عدم انطباقی مدرک به دانشجو عودت داده خواهد شد و لذا عدم انطباق نوع B در این موارد وجود ندارد. - مبنای اصلاحات در بخش نام و نام خانوادگی متن پایان‌نامه خواهد بود.	- ناقص بودن نام در سامانه ثبت در موارد زیر - مواردی مانند ال، حروف اضافه و ... به درستی نوشته نشده است. - یک حرف اشتباه تایپ شده است (پ به جای ژ یا ...) - مواردی مانند سید، سیده، میر در ابتدای اسم آورده نشده است. - مواردی مانند مهندس، دکتر، حجه الاسلام در ابتدای اسم آورده شده است. - نام بصورت لاتین نوشته شده است. - بطور کلی نام با صفحه عنوان پایان‌نامه مطابقت ندارد. - بخشی از یک نام چندبخشی نوشته نشده باشد.	مطابقت با متن پایان‌نامه / رساله	نام
	ناقص بودن نام خانوادگی در سامانه ثبت	مطابقت با متن پایان‌نامه / رساله	نام خانوادگی

^۱ **عدم انطباق نوع A:** این دسته از عدم انطباق‌ها شامل مواردی است که بحرانی بوده و شخص کنترل‌کننده بایستی در صورت مشاهده یکی از این موارد مدرک را به دانشجو عودت دهد.

^۲ **عدم انطباق نوع B:** این دسته از عدم انطباق‌ها شامل مواردی بوده که غیر بحرانی بوده و در صورت مشاهده این موارد کنترل‌کننده می‌تواند اصلاحات را در سامانه انجام دهد. وجود چندین عدم انطباق نوع B می‌تواند منجر به عدم انطباق نوع A در مدرک شده و عودت داده شود.

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان خارج کشور			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام فیلد
نوع B ^۲	نوع A ^۱		
	• مواردی مانند ال، حروف اضافه و ... به درستی نوشته نشده است. • یک حرف اشتباه تایپ شده است (پ به جای ژ یا ...) • نام خانوادگی بصورت لاتین نوشته شده است • بطور کلی نام خانوادگی با صفحه عنوان پایان‌نامه مطابقت ندارد. • بخشی از یک نام خانوادگی چندبخشی نوشته نشده باشد. • در صورتیکه پسوند اگر معنادار نباشد (مانند پور، زاده و ...) • نیم‌فاصله و اگر معنادار باشد فاصله داشته باشد.		
	• رایانامه از آنجاکه که مبنای مکاتبات و تعامل با دانشجویان است لذا بصورت غیر مستقیم و در عمل مورد صحه‌گذاری قرار می‌گیرد و کنترلی روی آن انجام نخواهد شد.	مطابقت با فرمت زیر: xxx@xxx.com	رایانامه
	• در حال حاضر با توجه به عدم امکان استفاده از سرویس ثبت احوال امکان کنترل شماره ملی وجود ندارد و کنترلی روی آن انجام نخواهد شد. • الگوریتم و رابطه میان اعداد کدملی در حال حاضر در این بخش چک خواهد شد.	مطابقت با الگوریتم شماره ملی مطابقت با وب سرویس تطبیق کد ملی و نام و نام خانوادگی در سامانه ثبت احوال	شماره ملی
• وجود هرگونه مغایرت میان صفحه عنوان و فیلد تکمیل شده در سامانه ثبت در صورت اطمینان از مغایرت مدرک با اطلاعات وارد شده در سیستم (در این حالت کاربر سیستم در صورت امکان با دانشجوی مربوطه تماس می‌گیرد و اصلاح مربوطه را انجام می‌دهد)	• برگرداندن مدرک در صورت عدم امکان تماس تلفنی با کاربر و عدم اطمینان از مقطع تحصیلی درج شده در سیستم	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله	مقطع تحصیلی

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان خارج کشور			
نام فیلد	وضعیت مطلوب	عدم انطباق‌های بالقوه	
		نوع A ^۱	نوع B ^۲
کشور محل تحصیل	انتخاب از گزینه‌های مندرج در فایل	عدم تطابق با صفحه عنوان پایان‌نامه/ رساله	
موسسه آموزشی/پژوهشی	انتخاب از گزینه‌های مندرج در لیست کرکره‌ای	عدم تطابق با صفحه عنوان پایان‌نامه/ رساله	
دانشکده/پژوهشکده	معنادار و صحیح بودن متن	خطای نوشتاری در عنوان دانشکده/ پژوهشکده	مغایرت با نام دانشکده/پژوهشکده مندرج در صفحه عنوان پایان‌نامه/ رساله
زبان اصلی پایان نامه	مطابقت با متن پایان‌نامه/ رساله (بایستی یکی از زبانهای فارسی، انگلیسی، فرانسوی، آلمانی یا عربی باشد).	هرگونه مغایرت میان زبان انتخاب شده و متن نوشته شده در پایان‌نامه/ رساله	
عنوان فارسی	مطابقت با عنوان انگلیسی	وجود مغایرت معنایی اساسی (چند کلمه یا بیشتر) میان فیلد عنوان و متن پایان‌نامه	- وجود کاراکترهای اضافی مانند space - وجود غلط‌های املایی در متن عنوان
عنوان انگلیسی	مطابقت با عنوان مندرج در صفحه عنوان انگلیسی پایان‌نامه/ رساله	- عدم انطباق کامل با متن صفحه عنوان انگلیسی پایان‌نامه - مطابق نبودن با عنوان فارسی	- وجود کاراکترهای اضافی مانند space - وجود غلط‌های املایی در متن عنوان
رشته	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله		مغایرت با نام رشته مندرج در صفحه عنوان پایان‌نامه/ رساله
گرایش	مطابقت با مشخصات مندرج در صفحه اول پایان‌نامه/ رساله		مغایرت با نام گرایش مندرج در صفحه عنوان پایان‌نامه/ رساله
تاریخ دفاع	مطابقت با متن پایان‌نامه/ رساله	غیر واقعی بودن (تاریخ دفاع دانشجو بعد از تاریخ ثبت در سامانه ثبت باشد)	مطابقت نداشتن با اطلاعات صفحه عنوان پایان‌نامه/ رساله

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان خارج کشور				
نام فیلد	وضعیت مطلوب		عدم انطباق‌های بالقوه	
			نوع A ^۱	نوع B ^۲
استاد راهنما/استاد مشاور	نام	نباید کمتر از ۲ و بیشتر از ۲۵ حرف باشد مطابقت با متن پایان‌نامه/رساله	• ناقص بودن نام در سامانه ثبت در موارد زیر • مواردی مانند ال، حروف اضافه و ... به درستی نوشته نشده است. • یک حرف اشتباه تایپ شده است (پ به جای ژ یا ...) • مواردی مانند سید، سیده، میر در ابتدای اسم آورده نشده است.	
	نام خانوادگی	نباید کمتر از ۲ و بیشتر از ۵۰ حرف باشد مطابقت با متن پایان‌نامه/رساله	• مواردی مانند مهندس، دکتر، حجه الاسلام در ابتدای اسم آورده شده است. • نام بصورت لاتین نوشته شده است. • بطور کلی نام با صفحه عنوان پایان‌نامه مطابقت ندارد. • بخشی از یک نام چندبخشی نوشته نشده باشد.	
	نام (انگلیسی)	نباید کمتر از ۲ و بیشتر از ۲۵ حرف باشد مطابقت با متن پایان‌نامه/رساله	• وجود یک نقص اساسی و یا کلیدی در نام/نام خانوادگی لاتین • نام/نام خانوادگی با صفحه عنوان لاتین پایان‌نامه مطابقت ندارد.	• یک حرف اشتباه تایپ شده است. • نام/نام خانوادگی بصورت فارسی نوشته شده است • بخشی از یک نام/نام خانوادگی چندبخشی نوشته نشده باشد.
	نام خانوادگی (انگلیسی)	نباید کمتر از ۲ و بیشتر از ۵۰ حرف باشد مطابقت با متن		

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان خارج کشور				
نام فیلد		وضعیت مطلوب		عدم انطباق‌های بالقوه
				نوع A ^۱
				نوع B ^۲
		پایان‌نامه / رساله		
		رایانامه	حتما باید فرمت زیر را رعایت کند: xxx@xxx.com	- در حال حاضر کنترل دستی بر روی این فیلدها صورت نمی‌گیرد. - کنترل فرمت بصورت اتومات توسط سامانه انجام می‌گردد و اجازه تکمیل این فیلد برای دانشجو در صورت عدول از این فرمت داده نمی‌شود.
چکیده(فارسی) ^۱	مطابقت معنایی با متن پایان‌نامه / رساله نباید کمتر از ۲۰ و بیشتر از ۳۰۰۰ حرف	- چکیده وارد نشده است. - متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص	قرار ندادن فرمول‌های شیمیایی و ریاضی در فیلد چکیده	

^۱ حتی در پایان نامه به زبان خارجی چکیده و صفحه عنوان فارسی الزامی است.

دسته‌بندی عدم انطباق‌ها در اقسام اطلاعاتی سامانه ثبت دانشجویان خارج کشور			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام فیلد
نوع B ^۲	نوع A ^۱		
متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص و تفاوت جزئی دارد.	- و تفاوت اساسی دارد. - حروف کلمات مختلف به هم چسبیده است - عنوان در چکیده وارد شده است - در چکیده متن عربی یا انگلیسی وارد شده است. - ایمیل شخص در چکیده ذکر شده است.	باشد	
- قرار ندادن فرمول‌های شیمیایی و ریاضی در فیلد چکیده - متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص و تفاوت جزئی دارد.	- چکیده وارد نشده است. - متن چکیده وارد شده در سامانه نسبت به متن پایان‌نامه نقص و تفاوت اساسی دارد. - حروف کلمات مختلف به هم چسبیده است - ایمیل شخص در چکیده ذکر شده است.	مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۳۰۰۰ حرف باشد	چکیده(انگلیسی)
در حال حاضر کنترل نمی‌شود.		انتخاب فقط یکی از گزینه‌های بنیادی، کاربردی و توسعه‌ای	نوع پایان نامه/ رساله
- عدم همخوانی شماره صفحات با فهرست مطالب - به هم‌ریختگی ساختار و ظاهر فهرست مطالب - چسبیدگی کلمات در فهرست مطالب		مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد	فهرست مطالب
در حال حاضر باتوجه به اینکه در متن پایان‌نامه/ رساله این موضوع چک می‌گردد لذا در سطح فراداده این موضوع کنترل نمی‌شود.		مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف باشد	فهرست منابع فارسی
		مطابقت با متن پایان‌نامه/ رساله نباید کمتر از ۲۰ و بیشتر از ۲۰۰۰۰ حرف	فهرست منابع غیر فارسی

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان خارج کشور			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام فیلد
نوع B ^۲	نوع A ^۱		
		باشد	
در یک فیلد بیش از یک کلید واژه وارد شده باشد.	- با کلیدواژه موجود در متن مطابقت داشته باشد - یک کلید واژه در چند فیلد تکرار شده باشد. - مطابقت نداشتن کلید واژه فارسی و لاتین - جابجا وارد شدن کلید واژه (انگلیسی به جای فارسی یا بالعکس) - فینگیلیش نوشتن کلید واژه انگلیسی بر مبنای کلیدواژه فارسی	کامل بودن فیلدها هر کلید واژه در یک فیلد باشد (به ازای هر کلید واژه یک فیلد)	کلید واژه ها
-	آپلود فایلی با فرمتی غیر از pdf توضیحات: موارد و آیتم‌های قابل کنترل در فایل pdf در جدول دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل PDF) ذکر شده است.	انطباق کامل فایل پایان‌نامه با قواعد نگارشی و محتوایی مرکز آموزشی آپلود الزامی فایل با فرمت pdf	فایل pdf پایان نامه/ رساله
-	آپلود فایلی با فرمتی غیر از doc توضیحات: موارد و آیتم‌های قابل کنترل در فایل ورد در جدول دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل ورد) ذکر شده است.	انطباق کامل فایل پایان‌نامه با قواعد نگارشی و محتوایی مرکز آموزشی آپلود الزامی فایل با فرمت doc	فایل word پایان نامه/ رساله
-	آپلود فایلی با فرمتی غیر از txt	انطباق کامل فایل‌های ضمیمه (غیر از فایل اصلی پایان‌نامه) با فایل‌های الزام شده در مرکز آموزشی (بعنوان مثال فایل لاتکس در رشته ریاضی و ...)	دیگر فایل‌های پایان نامه/ رساله مانند Latex و ...

دسته‌بندی عدم انطباق‌ها در اقلام اطلاعاتی سامانه ثبت دانشجویان خارج کشور			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام فیلد
نوع B ^۲	نوع A ^۱		

جدول ۵-۷: طرح نهایی کیفیت فراداده‌ها در سامانه ثبت دانشجویان خارجی (فایل Pdf)

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل PDF) ^۱			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام محتوا
نوع B	نوع A		
	- کامل نبودن آیتم‌های صفحه عنوان - کامل نبودن نام دانشگاه - ناقص بودن نام خانوادگی در متن فایل در مقایسه با اطلاعات وارد شده در سامانه ثبت^۲	- کامل بودن آیتم‌های زیر در صفحه عنوان • لوگوی دانشگاه • نام دانشگاه (نام واحد) • دانشکده • مقطع (ارشد/ دکتری) • رشته و گرایش • عنوان پایان‌نامه • نام و نام خانوادگی نویسنده • نام استادراهنما • نام استاد مشاور • تاریخ دفاع	صفحه عنوان انگلیسی
چینش نامناسب در صفحه	- قرار نگرفتن چکیده در پایان‌نامه - ایمیل یا شماره تلفن شخص در چکیده ذکر شده است. - مناسب نبودن فرمت در چکیده (مواردی مانند به هم چسبیدن کلمات، خوانا نبودن حروف و ...)	- قرار گرفتن یک چکیده بصورت کامل در یک صفحه - چینش مناسب چکیده در وسط صفحه - مناسب بودن فرمت چکیده	چکیده فارسی

^۱ کلیه موارد و آیتمهای کنترلی در فایل ورد هم مشابه فایل پی دی اف میباشد

^۲ از آنجاییکه در جستجوهای آتی توسط کاربران، اطلاعات ثبت شده در فیلدهای سامانه مهم است این عدم انطباق میتواند جزئی تلقی گردد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل PDF) ^۱			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام محتوا
نوع B	نوع A		
	- نبود شماره صفحه - عدم وجود فصولی مانند نتیجه‌گیری - مطابق نبودن شماره صفحه فصل پایانی فایل با تعداد صفحات فایل - به هم چسبیدن کلمات (نامناسب بودن فرمت نوشتاری فهرست مطالب)^۱	- وجود شماره صفحه - مطابق بودن تعداد فصلها با متن - وجود ترتیب منطقی میان فصلهای پایان‌نامه	فهرست مطالب
	- عدم تطابق میان شماره صفحه ذکر شده در فهرست جداول و شماره صفحه قرار گرفته شده در متن - قرار ندادن جدول در متن در حالیکه در فهرست ذکر شده باشد. - در فهرست جداول فقط شماره آمده است (عنوان ذکر نشده است)	- کامل بودن فهرست جداول - مطابقت میان شماره صفحه جداول ذکر شده در فهرست با محل قرارگیری آنها در متن	فهرست جداول
	- عدم تطابق میان شماره صفحه ذکر شده در فهرست اشکال و شماره صفحه قرار گرفته شده در متن - قرار ندادن شکل در متن در حالیکه در فهرست ذکر شده باشد. - در فهرست اشکال فقط شماره آمده است (عنوان ذکر نشده است)	- کامل بودن فهرست اشکال - مطابقت میان شماره صفحه اشکال ذکر شده در فهرست با محل قرارگیری آنها در متن	فهرست اشکال
	عدم تطابق میان شماره صفحه ذکر شده در فهرست نمودار و	کامل بودن فهرست نمودارها	فهرست نمودار

^۱ اگر فایل ورد ارسالی توسط نویسنده مناسب باشد، از آن برای اصلاح فایل PDF استفاده خواهد شد.

دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل PDF) ^۱			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام محتوا
نوع B	نوع A		
	- شماره صفحه قرار گرفته شده در متن - قرار ندادن نمودار در متن در حالیکه در فهرست ذکر شده باشد. - در فهرست نمودار فقط شماره آمده است (عنوان ذکر نشده است)	مطابقت میان شماره صفحه نمودارها ذکر شده در فهرست با محل قرارگیری آنها در متن	
وجود صفحات سفید^۲ (خصوصاً در سی صفحه اول)	- ناقص بودن پایان‌نامه (فصلها کم است) - قرار ندادن پیوست در پایان‌نامه در حالیکه در فهرست وجود دارد - نداشتن شماره صفحه - به هم ریختگی فونت پایان‌نامه - قفل شدن فایل پی‌دی‌اف توسط نویسنده^۱ - فایل ارسالی دارای واترمارک باشد.	- کامل بودن متن پایان‌نامه - فایل پایان‌نامه بدون پس‌نقش(واترمارک) باشد. - فایل پایان‌نامه بدون آدرس و مشخصات شخصی دانشجو و استاد راهنما باشد.	متن پایان‌نامه
	نبود عنوان یا شماره برای جداول	جداول مندرج در متن دارای عنوان و شماره باشند.	جداول
	نبود عنوان و شماره برای اشکال	اشکال مندرج در متن دارای عنوان و شماره باشند	اشکال
	نبود عنوان و شماره برای نمودارها	نمودارهای مندرج در متن دارای عنوان و شماره باشند	نمودار

^۱ در صورت کامل بودن فایل ورد از آن پی‌دی‌اف تهیه شده و آپلود میگردد

^۲ در صورتی که صفحه سفید شماره صفحه داشته باشد، حذف نمی‌گردد.

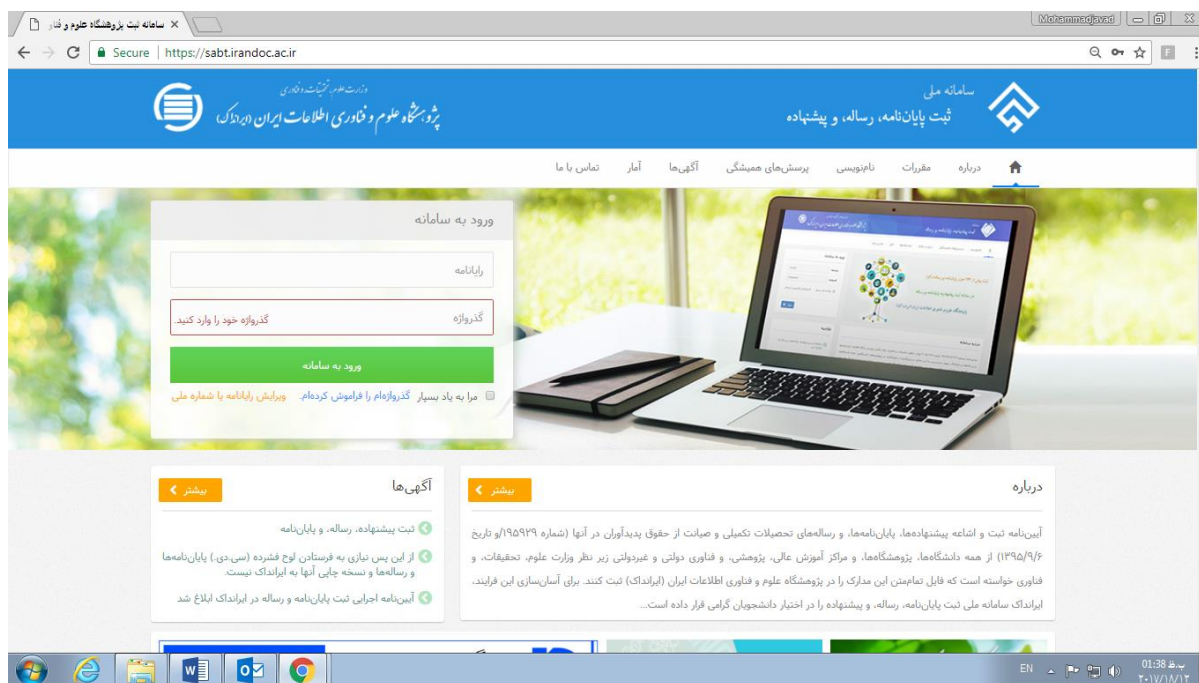
دسته‌بندی عدم انطباق‌ها در محتوای پایان‌نامه (فایل PDF) ^۱			
عدم انطباق‌های بالقوه		وضعیت مطلوب	نام محتوا
نوع B	نوع A		
• وجود به هم‌ریختگی در متن منابع	• نبود بخش منابع	• وجود بخش منابع در متن	منابع
	• مشاهده نشدن بخشی یا کل پیوست‌ها در پایان متن در صورتیکه در متن یا فهرست به آن اشاره شده باشد.	• درج شدن پیوست در پایان متن در صورتیکه در متن یا فهرست به آن اشاره شده باشد.	پیوست
• چینش نامناسب در صفحه	• قرار نگرفتن چکیده در پایان‌نامه • ایمیل یا شماره تلفن شخص در چکیده ذکر شده است. • مناسب نبودن فرمت در چکیده (مواردی مانند به هم چسبیدن کلمات، خوانا نبودن حروف و ...)	• قرار گرفتن یک چکیده بصورت کامل در یک صفحه • چینش مناسب چکیده در وسط صفحه • مناسب بودن فرمت چکیده	چکیده انگلیسی

۳-۵ فرایند کنترل کیفیت در سامانه ثبت

همانگونه که در بخشهای پیشین مطرح گردید کارشناسان واحد ثبت و فراهم آوری اطلاعات، مراحل کنترل کیفیت فراداده در سامانه ثبت را بصورت مرحله به مرحله صورت خواهند داد. مراحل مذکور به شرح زیر می باشد.

الف) ورود به سامانه ثبت با نام کاربری و رمز عبور

همانگونه که در شکل ۴-۵ مشاهده می گردد. کارشناس اداره ثبت و فراهم آوری اطلاعات پس از ورود به سامانه ثبت با درج عنوان sabt.irandoc.ac.ir مراحل کنترل پایان نامه / رساله را آغاز می نماید.

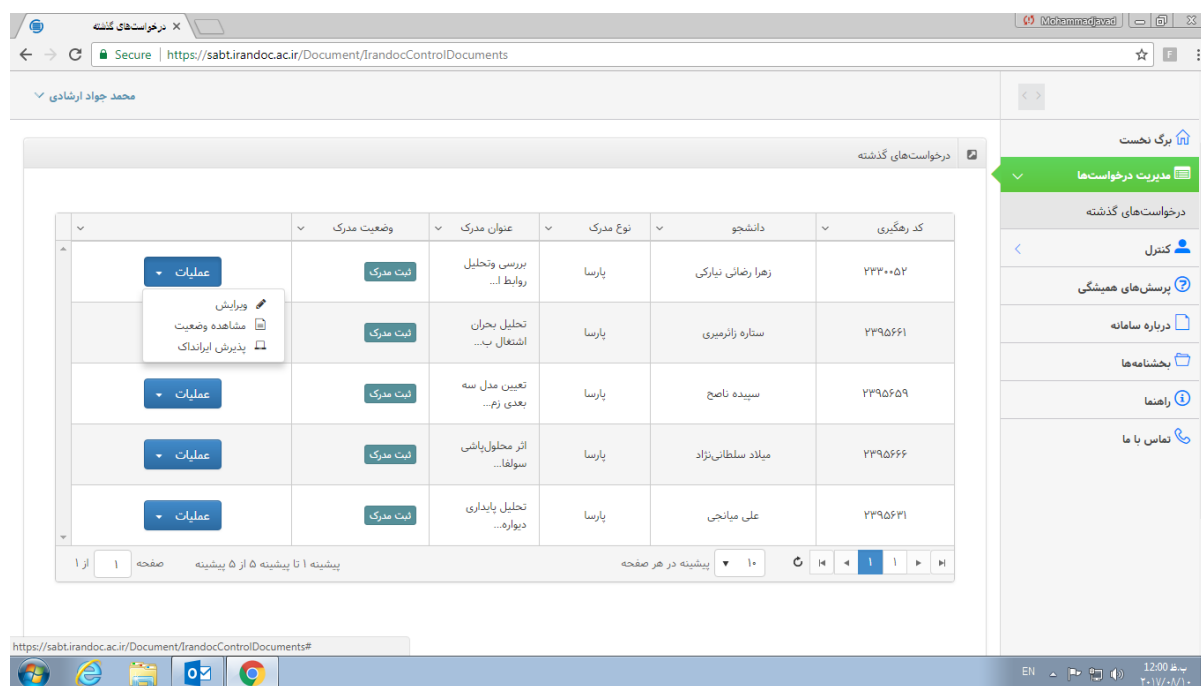


شکل ۵- ۴. صفحه ورود به سامانه ثبت به منظور کنترل مدرک ارسال شده توسط دانشجویان

ب) انتخاب پایان نامه / رساله بمنظور کنترل کیفیت فراداده

طراحی مدل کیفیت فراداده در سامانه ثبت ... □ ۲۳۸

پس از ورود کارشناس به سامانه ثبت، در تب مدیریت درخواست‌ها کلیه پایان‌نامه / رساله‌های موجود در کارتابل را مشاهده می‌نماید. به ترتیب از بالا زودترین مدرک انتخاب شده و پس از کلیک روی تب عملیات و سپس انتخاب گزینه "پذیرش ایرانداک" مراحل کنترل کیفیت آغاز می‌گردد. (شکل ۵-۵)



شکل ۵-۵. مشاهده فهرست درخواست‌های در حال انتظار در سامانه ثبت

ج) مشاهده و دانلود فایل‌های Pdf و Word

در این مرحله کارشناس در آغاز مرحله کنترل فایل‌های Pdf و Word را دانلود نموده و مراحل کنترل کیفیت فراداده را آغاز می‌نماید. (شکل ۵-۶)

شکل ۵-۶. آغاز فرایند بررسی مدرک ارسال شده توسط دانشجو

(د) کنترل داده‌های شخصی و تحصیلی

در این مرحله مطابق طرح کیفیت فراداده و چک‌لیست‌های کنترل کیفیت داده‌های شخصی و تحصیلی دانشجو مورد بررسی و کنترل قرار می‌گیرد و در صورت مشاهده مغایرت مطابق طرح کیفیت فراداده عمل خواهد شد. (شکل ۵-۷)

فایل Pdf پایان نامه / رساله	فایل Word پایان نامه / رساله	دیگر فایل های پایان نامه / رساله
مشاهده	مشاهده	-

اطلاعات شخصی و تحصیلی			
نام و نام خانوادگی	زهرا رضائی نیارکی	رایانامه	niyaraki94@gmail.com
شماره ملی	۰۰۵۴۱۴۱۶۴۸	شماره دانشجویی	۹۴۱۹۰۵۲
مقطع تحصیلی	کارشناسی ارشد	وابستگی سازمانی	مؤسسه های آموزش عالی غیردولتی-غیرانتفاعی
مؤسسه آموزشی / پژوهشی	مؤسسه آموزش عالی دانش البرز	دانشکده / پژوهشکده	مؤسسه آموزش عالی البرز

اطلاعات پایان نامه / رساله	
عنوان (فارسی)	بررسی و تحلیل، روابط ایران، معماری، دوره سقراط، دوم (با تأکید بر جنبش، تفکر)

چاپ
بارگذاری سریع
ویرایش
پذیرش ایرادها
رد
بازگشت

2330052_e1a710....docx 6.0/11.0 MB, 8 sec
2330052_9a31cda....pdf
Show all

EN
12:03
۱۳۹۷/۰۷/۱۰

شکل ۵- ۷. صفحه اطلاعات شخصی و تحصیلی دانشجو

ه) کنترل اطلاعات پایان نامه / رساله

در این مرحله اطلاعات پایان نامه / رساله شامل عنوان فارسی و انگلیسی، گروه تحصیلی، رشته و گرایش، زبان اصلی، تاریخ دفاع، استاد راهنمای نخست و دوم و نیز استاد مشاور نخست و دوم مورد بررسی قرار می گیرد.

پذیرش ایراندک

مؤسسه آموزشی / پژوهشی	مؤسسه آموزش عالی دانش البرز	دانشکده / پژوهشکده	مؤسسه آموزش عالی البرز
-----------------------	-----------------------------	--------------------	------------------------

اطلاعات پایان نامه / رساله

عنوان (فارسی)	بررسی و تحلیل روابط ایران و عمان در دوره پهلوی دوم (با تأکید بر جنبش نظام)		
عنوان (انگلیسی)	Analysis of relations between Iran and Oman in the second Pahlavi era (with an emphasis on movement Dhofer)		
گروه تحصیلی	علوم انسانی	رشته و گرایش تحصیلی	تاریخ - تاریخ انقلاب اسلامی
زبان اصلی پایان نامه / رساله	فارسی	تاریخ دفاع	۱۳۹۶/۶/۳۱
استاد راهنمای نخست	امیرحسین جانی	استاد راهنمای دوم	
استاد مشاور نخست		استاد مشاور دوم	

محتوای پایان نامه / رساله

2330052_e1a710....docx 2330052_9a31cda....pdf Show all

12:05 ۲۰۱۷/۰۸/۱۰

شکل ۵-۸. صفحه اطلاعات پایان نامه / رساله

و) کنترل محتوای پایان نامه / رساله

در این بخش محتوای پایان نامه / رساله و میزان انطباق آن با فایل های Pdf و Word بر مبنای طرح کیفیت فراداده طراحی شده مورد ارزیابی و کنترل قرار می گیرد.

طراحی مدل کیفیت فراداده در سامانه ثبت ... □ ۲۴۲

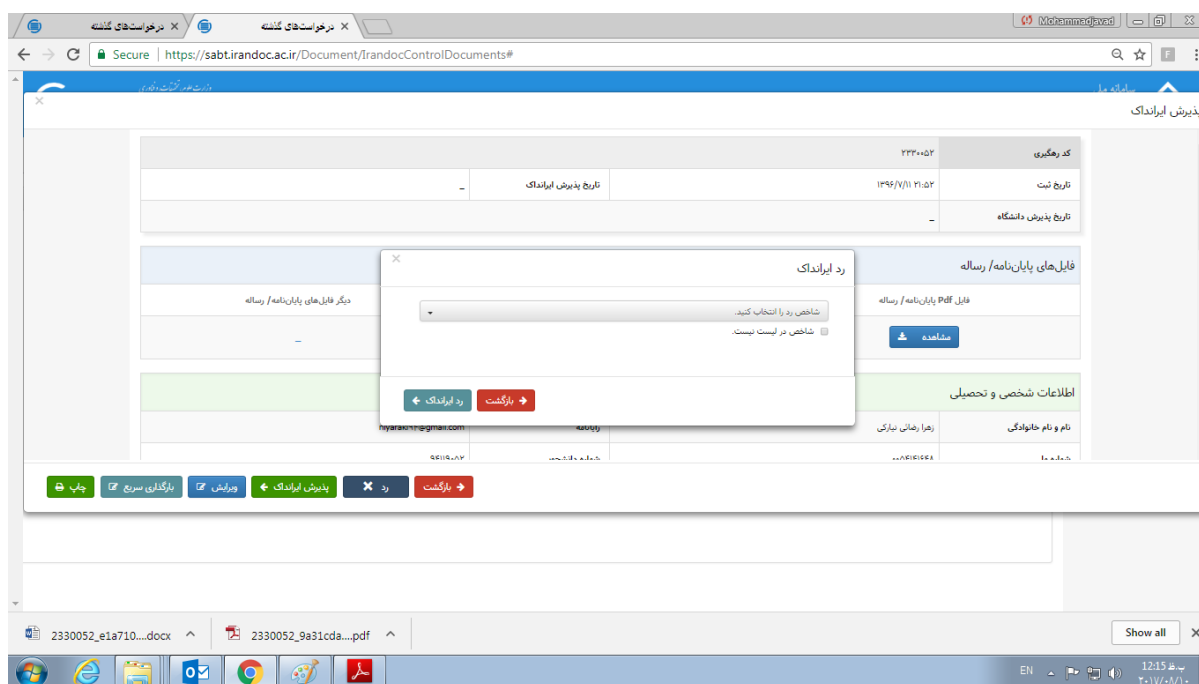
شکل ۵-۹. صفحه اطلاعات چکیده‌های پایان‌نامه/رساله

شکل ۵-۱۰. صفحه اطلاعات فهرست و کلیدواژه‌ها پایان‌نامه/رساله

اشکال ۵-۹ و ۵-۱۰ مواردی که در بخش محتوای پایان‌نامه/ رساله مورد کنترل و ارزیابی قرار می‌گیرند را نمایش می‌دهد. مواردی مانند فهرست مطالب، چکیده فارسی و انگلیسی، فهرست منابع و نیز کلیدواژه‌های فارسی و انگلیسی در این بخش کنترل خواهند شد.

ز) اعمال نظر نهایی در خصوص پایان‌نامه/ رساله

در این بخش پس از بررسی اطلاعات مندرج در فایل‌های Pdf و Word و نیز اطلاعات ثبت شده در سیستم و بر مبنای طرح کیفی فراداده‌ها در سامانه ثبت کارشناس مربوطه تصمیم به رد یا پذیرش مدرک خواهد گرفت.



شکل ۵- ۱۱. صفحه اعلام نظر نهایی در خصوص پایان نامه/ رساله

در صورت رد ایراداک همانطور که در شکل بالا مشاهده می گردد شاخص رد ایراداک شامل دلایل آن نیز می گردد انتخاب خواهد شد.

چک لیست های کنترل کیفیت فراداده ها در سامانه ثبت به شرح زیر بر مبنای طرح کیفیت فراداده ها قابل ارائه خواهد بود.

چک لیست کنترل فراداده ها در سامانه ثبت پایان نامه/ رساله ها					
نام کارشناس کنترل کننده:			پایان نامه □	رساله □	
نوع پایان نامه/ رساله: دانشجوین داخل □		انشجویان خارج از کشور □		تاریخ کنترل:	
ردیف	نوع داده/ فراداده	نام قلم اطلاعاتی	نام مشخصه کنترلی	وضعیت انطباق ^۱	مغایرت مشاهده شده
۱	فراداده	نام	بررسی مواردی مانند ال، حروف اضافه در نام نویسنده		
۲			آیا حروف نام نویسنده به درستی درج شده است؟		
۳			تطابق مواردی مانند سید، سیده، میر در ابتدای نام نویسنده با صفحه عنوان		
۴			عدم درج مواردی مانند مهندس، دکتر، حجه الاسلام		
۵			عدم درج نام به صورت لاتین		
۶			تطابق کامل نام نویسنده با صفحه عنوان در صورت چندبخشی بودن		
۷		نام خانوادگی	بررسی مواردی مانند ال، حروف اضافه در نام خانوادگی نویسنده		

^۱ در صورت منطبق بودن آیتم مربوطه از علامت ✓ و در صورت نامنطبق بودن از علامت × استفاده گردد.

چک لیست کنترل فراداده‌ها در سامانه ثبت پایان‌نامه / رساله‌ها					
نام کارشناس کنترل کننده:			پایان نامه □ رساله □		
نوع پایان‌نامه / رساله: دانشجویان داخل □ دانشجویان خارج از کشور □		تاریخ کنترل:			
				آیا حروف نام خانوادگی نویسنده به درستی درج شده است؟	۸
				عدم درج نام خانوادگی به صورت لاتین	۹
				تطابق کامل نام خانوادگی نویسنده با صفحه عنوان در صورت چندبخشی بودن	۱۰
				نیم‌فاصله بودن در صورتیکه پسوند معنادار نباشد (مانند پور، زاده و ...) و فاصله داشتن اگر معنادار باشد.	۱۱
				آیا مقطع تحصیلی ثبت شده در سامانه ثبت با صفحه عنوان مغایرت دارد؟	۱۲
				آیا مقطع وابستگی سازمانی ثبت شده در سامانه ثبت با صفحه عنوان مغایرت دارد؟	۱۳
	موسسه آموزشی/پژوهشی			آیا در صورتیکه "دانشگاههای وابسته به حوزه علمیه" انتخاب شده است، پایان نامه مربوط به حوزه علمیه است؟	۱۴
				بررسی تطابق کامل میان نام موسسه انتخاب شده با صفحه عنوان	۱۵
				در دانشگاههای پیام نور/ آزاد آیا واحد مشخص شده است؟	۱۶
	دانشکده/پژوهشکده			بررسی اینکه در صورتیکه نام دانشکده/ پژوهشکده بصورت دستی توسط دانشجو در سامانه ثبت وارد شده است، در لیست کرکره‌ای موجود نباشد.	۱۷

چک‌لیست کنترل فراداده‌ها در سامانه ثبت پایان‌نامه/ رساله‌ها					
نام کارشناس کنترل کننده:			پایان نامه □ رساله □		
نوع پایان‌نامه/ رساله: دانشجویان داخل □ دانشجویان خارج از کشور □		تاریخ کنترل:			
۱۸		بررسی تطابق میان نام دانشکده/ پژوهشکده و صفحه عنوان پایان‌نامه/ رساله			
۱۹	زبان اصلی پایان نامه	بررسی هرگونه مغایرت میان زبان انتخاب شده و متن نوشته شده در پایان‌نامه/ رساله			
۲۰	عنوان فارسی	بررسی هرگونه مغایرت (کلمات عنوان) میان فیلد عنوان و صفحه عنوان فارسی پایان‌نامه			
۲۱		وجود کاراکترهای اضافی مانند space			
۲۲		وجود غلط‌های املایی در متن عنوان			
۲۳	عنوان انگلیسی	بررسی هرگونه مغایرت (کلمات عنوان) میان فیلد عنوان و صفحه عنوان انگلیسی پایان‌نامه			
۲۴		وجود کاراکترهای اضافی مانند space			
۲۵		وجود غلط‌های املایی در متن عنوان			
۲۶	رشته	انطباق با نام رشته مندرج در صفحه عنوان پایان‌نامه/ رساله			
۲۷	گرایش	انطباق با نام گرایش مندرج در صفحه عنوان پایان‌نامه/ رساله			
۲۸	تاریخ دفاع	انطباق با اطلاعات صفحه عنوان پایان‌نامه/ رساله			
۲۹		واقعی بودن (تاریخ دفاع دانشجو بعد از تاریخ ثبت در سامانه ثبت باشد)			
۳۰	نام استادراهنما/	بررسی مواردی مانند ال، حروف اضافه			

چک‌لیست کنترل فراداده‌ها در سامانه ثبت پایان‌نامه / رساله‌ها				
نام کارشناس کنترل کننده:		پایان نامه □ رساله □		
نوع پایان‌نامه / رساله: دانشجوین داخل □ انشجویان خارج از کشور □		تاریخ کنترل:		
استاد مشاور		در نام نویسنده		
		آیا حروف نام نویسنده به درستی درج شده است؟		۳۱
		تطابق مواردی مانند سید، سیده، میر در ابتدای نام نویسنده با صفحه عنوان		۳۲
		عدم درج مواردی مانند مهندس، دکتر، حجه الاسلام		۳۳
		عدم درج نام به صورت لاتین		۳۴
		تطابق کامل نام نویسنده با صفحه عنوان در صورت چندبخشی بودن		۳۵
		بررسی مواردی مانند ال، حروف اضافه در نام خانوادگی نویسنده		۳۶
		آیا حروف نام خانوادگی نویسنده به درستی درج شده است؟		۳۷
		عدم درج نام خانوادگی به صورت لاتین		۳۸
		تطابق کامل نام خانوادگی نویسنده با صفحه عنوان در صورت چندبخشی بودن		۳۹
نام خانوادگی استاد مشاور		نیم‌فاصله بودن در صورتیکه پسوند معنادار نباشد (مانند پور، زاده و ...) و فاصله داشتن اگر معنادار باشد.		۴۰
		بررسی هرگونه نقص اساسی و یا کلیدی در نام/ نام خانوادگی لاتین		۴۱
		بررسی مطابقت نام/ نام خانوادگی با صفحه عنوان لاتین پایان‌نامه		۴۲
		تایپ اشتباه حروف		۴۳
		فارسی نوشته نشدن نام انگلیسی		۴۴
نام/ نام خانوادگی انگلیسی		در صورت چندبخشی بودن بررسی		۴۵

چک‌لیست کنترل فراداده‌ها در سامانه ثبت پایان‌نامه / رساله‌ها					
نام کارشناس کنترل کننده:			پایان‌نامه □ رساله □		
نوع پایان‌نامه / رساله: دانشجوین داخل □ انشجویان خارج از کشور □		تاریخ کنترل:			
چکیده فارسی					درج درست کلیه بخشها
					وارد شدن صحیح چکیده
					عدم تفاوت میان این فیلد و چکیده درج شده در متن
					به هم چسبیده نبودن کلمات
					نبود متن فارسی یا عربی در چکیده
					عدم درج ایمیل یا شماره تلفن در چکیده
					اطمینان از وارد نشدن فرمول یا رابطه ریاضی یا شیمیایی
					وارد شدن صحیح چکیده
					عدم تفاوت میان این فیلد و چکیده درج شده در متن
					به هم چسبیده نبودن کلمات
چکیده انگلیسی					عدم درج ایمیل یا شماره تلفن در چکیده
					اطمینان از وارد نشدن فرمول یا رابطه ریاضی یا شیمیایی
					همخوانی شماره صفحات با فهرست مطالب
					عدم هم‌ریختگی ساختار و ظاهر فهرست مطالب
					چسبیدگی کلمات در فهرست مطالب
کلید واژه ها					تطابق با کلیدواژه موجود در متن
					تکرار نشدن یک کلید واژه در چند فیلد
					مطابقت داشتن کلید واژه فارسی و لاتین
					جابجا وارد نشدن کلید واژه (انگلیسی

چک‌لیست کنترل فراداده‌ها در سامانه ثبت پایان‌نامه / رساله‌ها					
نام کارشناس کنترل کننده:			پایان نامه ☐ رساله ☐		
نوع پایان‌نامه / رساله: دانشجویان داخل ☐ دانشجویان خارج از کشور ☐		تاریخ کنترل:			
		به جای فارسی یا بالعکس)			
۶۴		فینگلیش نوشته نشدن کلید واژه انگلیسی بر مبنای کلیدواژه فارسی			
۶۵		در یک فیلد بیش از یک کلید واژه وارد نشده باشد.			
۶۶	فایل Pdf	کامل بودن آیتم‌های زیر در صفحه عنوان • لوگوی دانشگاه • نام دانشگاه (نام واحد) • دانشکده • مقطع (ارشد/ دکتری) • رشته و گرایش • عنوان پایان‌نامه • نام و نام خانوادگی نویسنده • نام استادراهنما • نام استاد مشاور • تاریخ دفاع	صفحه عنوان		
۶۷		در دانشگاه‌های پیام نور/ آزاد آیا واحد مشخص شده است؟			
۶۸		بررسی ناقص نبودن نام خانوادگی در متن فایل در مقایسه با اطلاعات وارد شده در سامانه			
۶۹	چکیده فارسی	چینش مناسب در صفحه			
۷۰		به هم چسبیده نبودن کلمات			
۷۱		نبود متن فارسی یا عربی در چکیده			
۷۲		عدم درج ایمیل یا شماره تلفن در چکیده			
۷۳		اطمینان از وارد نشدن فرمول یا رابطه			

چک لیست کنترل فراداده‌ها در سامانه ثبت پایان نامه / رساله‌ها					
نام کارشناس کنترل کننده:			پایان نامه ☐ رساله ☐		
نوع پایان نامه / رساله: دانشجوین داخل ☐ انشجویان خارج از کشور ☐		تاریخ کنترل:			
	فهرست مطالب	ریاضی یا شیمیایی			
		به هم چسبیده نبودن کلمات			۷۴
		درج صحیح شماره صفحه			۷۵
		اطمینان از هم چسبیده نبودن کلمات (نامناسب بودن فرمت نوشتاری فهرست مطالب			۷۶
		مطابقت شماره صفحه فصل پایانی فایل با تعداد صفحات فایل			۷۷
		بررسی وجود فصولی مانند نتیجه گیری			۷۸
	فهرست جداول	مطابقت میان شماره صفحه ذکر شده در فهرست جداول و شماره صفحه قرار گرفته شده در متن			۷۹
		قرار دادن جدول در متن در حالیکه در فهرست ذکر شده باشد			۸۰
		کامل بودن فهرست جداول (عنوان و شماره صفحه)			۸۱
		مطابقت میان شماره صفحه جداول ذکر شده در فهرست با محل قرارگیری آنها در متن			۸۲
	فهرست اشکال	مطابقت میان شماره صفحه ذکر شده در فهرست اشکال و شماره صفحه قرار گرفته شده در متن			۸۳
		قرار دادن شکل در متن در حالیکه در فهرست ذکر شده باشد			۸۴
		کامل بودن فهرست اشکال (عنوان و شماره صفحه)			۸۵
		مطابقت میان شماره صفحه اشکال ذکر شده در فهرست با محل قرارگیری آنها در متن			۸۶

چک‌لیست کنترل فراداده‌ها در سامانه ثبت پایان‌نامه / رساله‌ها					
نام کارشناس کنترل کننده:			پایان نامه □	رساله □	
نوع پایان‌نامه / رساله: دانشجوین داخل □		انشجویان خارج از کشور □	تاریخ کنترل:		
۸۷	فهرست نمودار				مطابقت میان شماره صفحه ذکر شده در فهرست نمودار و شماره صفحه قرار گرفته شده در متن
					قرار دادن نمودار در متن در حالیکه در فهرست ذکر شده باشد
					کامل بودن فهرست نمودار (عنوان و شماره صفحه)
					مطابقت میان شماره صفحه نمودار ذکر شده در فهرست با محل قرارگیری آنها در متن
۸۸	متن پایان‌نامه				نداشتن پس‌نقش (واترمارک)
۸۹					ناقص نبودن پایان‌نامه
۹۰					اطمینان از قرار ندادن پیوست در پایان‌نامه در حالیکه در فهرست وجود دارد.
۹۱					اطمینان از داشتن شماره صفحه
۹۲					اطمینان از عدم به هم ریختگی فونت پایان‌نامه
۹۳					اطمینان از نبود صفحات سفید
۹۴					اطمینان از اینکه پایان‌نامه بدون آدرس و مشخصات شخصی دانشجو و استاد راهنما باشد
۹۵					قفل نشدن فایل پی‌دی‌اف توسط نویسنده
۹۶					جداول مندرج در متن دارای عنوان و شماره باشند.
۹۷					اشکال مندرج در متن دارای عنوان و شماره باشند.
۹۸	جداول				نمودارهای مندرج در متن دارای عنوان و شماره باشند.
۹۹	اشکال				
۱۰۰	نمودار				
۱۰۱					

چک‌لیست کنترل فراداده‌ها در سامانه ثبت پایان‌نامه / رساله‌ها					
نام کارشناس کنترل کننده:			پایان نامه ☐ رساله ☐		
نوع پایان‌نامه / رساله: دانشجویان داخل ☐ دانشجویان خارج از کشور ☐		تاریخ کنترل:			
۱۰۲	منابع	وجود بخش منابع در متن			
۱۰۳		اطمینان از به هم‌ریخته نبودن متن منابع			
۱۰۴	پیوست	درج شدن پیوست در پایان متن در صورتیکه در متن یا فهرست به آن اشاره شده باشد.			
۱۰۵	صفحه عنوان انگلیسی	کامل بودن آیتم‌های زیر در صفحه عنوان • لوگوی دانشگاه • نام دانشگاه (نام واحد) • دانشکده • مقطع (ارشد/ دکتری) • رشته و گرایش • عنوان پایان‌نامه • نام و نام خانوادگی نویسنده • نام استادراهنما • نام استاد مشاور • تاریخ دفاع			
۱۰۶		در دانشگاه‌های پیام نور/ آزاد آیا واحد مشخص شده است؟			
۱۰۷		بررسی ناقص نبودن نام خانوادگی در متن فایل در مقایسه با اطلاعات وارد شده در سامانه			
۱۰۸	چکیده انگلیسی	چینش مناسب در صفحه			
۱۰۹		به هم چسبیده نبودن کلمات			
۱۱۰		نبود متن فارسی یا عربی در چکیده			
۱۱۱		عدم درج ایمیل یا شماره تلفن در چکیده			
۱۱۲		اطمینان از وارد نشدن فرمول یا رابطه ریاضی یا شیمیایی			

چک لیست کنترل فراداده‌ها در سامانه ثبت پایان نامه / رساله‌ها					
نام کارشناس کنترل کننده:			پایان نامه ☐ رساله ☐		
نوع پایان نامه / رساله: دانشجویان داخل ☐ دانشجویان خارج از کشور ☐		تاریخ کنترل:			
۱۱۳			به هم چسبیده نبودن کلمات		
۱۱۴	فایل ورد	صفحه عنوان	کامل بودن آیتم‌های زیر در صفحه عنوان		
			• لوگوی دانشگاه • نام دانشگاه (نام واحد) • دانشکده • مقطع (ارشد/ دکتری) • رشته و گرایش • عنوان پایان نامه • نام و نام خانوادگی نویسنده • نام استادراهنما • نام استاد مشاور • تاریخ دفاع		
۱۱۵	فایل ورد		در دانشگاه‌های پیام نور/ آزاد آیا واحد مشخص شده است؟		
۱۱۶			بررسی ناقص نبودن نام خانوادگی در متن فایل در مقایسه با اطلاعات وارد شده در سامانه		
۱۱۷			چینش مناسب در صفحه		
۱۱۸			به هم چسبیده نبودن کلمات		
۱۱۹			نبود متن فارسی یا عربی در چکیده		
۱۲۰			عدم درج ایمیل یا شماره تلفن در چکیده		
۱۲۱			اطمینان از وارد نشدن فرمول یا رابطه ریاضی یا شیمیایی		
۱۲۲			به هم چسبیده نبودن کلمات		
۱۲۳			درج صحیح شماره صفحه		
۱۲۴				فهرست مطالب	اطمینان از هم چسبیده نبودن کلمات (نامناسب بودن فرمت نوشتاری فهرست مطالب)

چک‌لیست کنترل فراداده‌ها در سامانه ثبت پایان‌نامه / رساله‌ها					
نام کارشناس کنترل کننده:			پایان نامه □ رساله □		
نوع پایان‌نامه / رساله: دانشجویان داخل □ دانشجویان خارج از کشور □		تاریخ کنترل:			
۱۲۵				مطابقت شماره صفحه فصل پایانی فایل با تعداد صفحات فایل	
				بررسی وجود فصولی مانند نتیجه گیری	
				مطابقت میان شماره صفحه ذکر شده در فهرست جداول و شماره صفحه قرار گرفته شده در متن	
				قرار دادن جدول در متن در حالیکه در فهرست ذکر شده باشد	
				کامل بودن فهرست جداول (عنوان و شماره صفحه)	
				مطابقت میان شماره صفحه جداول ذکر شده در فهرست با محل قرارگیری آنها در متن	
۱۲۶				مطابقت میان شماره صفحه ذکر شده در فهرست اشکال و شماره صفحه قرار گرفته شده در متن	
				قرار دادن شکل در متن در حالیکه در فهرست ذکر شده باشد	
				کامل بودن فهرست اشکال (عنوان و شماره صفحه)	
				مطابقت میان شماره صفحه اشکال ذکر شده در فهرست با محل قرارگیری آنها در متن	
۱۲۷				مطابقت میان شماره صفحه ذکر شده در فهرست نمودار و شماره صفحه قرار گرفته شده در متن	
				قرار دادن نمودار در متن در حالیکه در فهرست ذکر شده باشد	
				کامل بودن فهرست نمودار (عنوان و شماره صفحه)	
۱۲۸					
۱۲۹					
۱۳۰					
۱۳۱					
۱۳۲					
۱۳۳					
۱۳۴					
۱۳۵					
۱۳۶					
۱۳۷					

چک‌لیست کنترل فراداده‌ها در سامانه ثبت پایان‌نامه / رساله‌ها					
نام کارشناس کنترل کننده:			پایان نامه □ رساله □		
نوع پایان‌نامه / رساله: دانشجویان داخل □ دانشجویان خارج از کشور □		تاریخ کنترل:			
۱۳۸	متن پایان‌نامه	شماره صفحه)			
		مطابقت میان شماره صفحه نمودار ذکر شده در فهرست با محل قرارگیری آنها در متن			
		نداشتن پس‌نقش (واترمارک)			۱۳۹
		ناقص نبودن پایان‌نامه			۱۴۰
		اطمینان از قرار ندادن پیوست در پایان‌نامه در حالیکه در فهرست وجود دارد.			۱۴۱
		اطمینان از داشتن شماره صفحه			۱۴۲
		اطمینان از عدم به هم ریختگی فونت پایان‌نامه			۱۴۳
		اطمینان از نبود صفحات سفید			۱۴۴
		اطمینان از اینکه پایان‌نامه بدون آدرس و مشخصات شخصی دانشجو و استاد راهنما باشد			۱۴۵
		قفل نشدن فایل پی‌دی اف توسط نویسنده			۱۴۶
۱۴۷	جداول	جداول مندرج در متن دارای عنوان و شماره باشند.			
۱۴۸	اشکال	اشکال مندرج در متن دارای عنوان و شماره باشند.			
۱۴۹	نمودار	نمودارهای مندرج در متن دارای عنوان و شماره باشند.			
۱۵۰	منابع	وجود بخش منابع در متن			
۱۵۱		اطمینان از به هم‌ریخته نبودن متن منابع			
۱۵۲	پیوست	درج شدن پیوست در پایان‌نامه در صورتیکه در متن یا فهرست به آن اشاره شده باشد.			

چک‌لیست کنترل فراداده‌ها در سامانه ثبت پایان‌نامه / رساله‌ها					
نام کارشناس کنترل کننده:			پایان نامه ☐ رساله ☐		
نوع پایان‌نامه / رساله: دانشجویان داخل ☐ دانشجویان خارج از کشور ☐		تاریخ کنترل:			
۱۵۳	صفحه عنوان انگلیسی				کامل بودن آیتم‌های زیر در صفحه عنوان
					• لوگوی دانشگاه
					• نام دانشگاه (نام واحد)
۱۵۴					• دانشکده
					• مقطع (ارشد/ دکتری)
					• رشته و گرایش
۱۵۵					• عنوان پایان‌نامه
					• نام و نام خانوادگی نویسنده
					• نام استاد راهنما
۱۵۶					• نام استاد مشاور
					• تاریخ دفاع
					در دانشگاه‌های پیام نور/ آزاد آیا واحد مشخص شده است؟
۱۵۷					بررسی ناقص نبودن نام خانوادگی در متن فایل در مقایسه با اطلاعات وارد شده در سامانه
					چینش مناسب در صفحه
					به هم چسبیده نبودن کلمات
۱۵۸					نبود متن فارسی یا عربی در چکیده
					عدم درج ایمیل یا شماره تلفن در چکیده
					اطمینان از وارد نشدن فرمول یا رابطه ریاضی یا شیمیایی
۱۵۹					به هم چسبیده نبودن کلمات

شکل ۵- ۱۲. چک‌لیست نهایی کنترل فراداده‌ها در سامانه ثبت

۵-۴ ارائه پیشنهادات در حوزه طراحی ویرایش‌های آتی سامانه ثبت

پیشنهادات سازنده مطرح شده ذیل در کارگروه بررسی روایی طرح کیفی داده‌ها در سامانه ثبت بمنظور ایجاد بهبود در توسعه‌های آتی سامانه ثبت طرح و بحث گردیده‌اند.

۱. در حال حاضر از آنجاکه برخی کنترلرها در دو واحد اداره ثبت و فراهم‌آوری اطلاعات و سازماندهی و تحلیل اطلاعات مشترک هستند لذا امکانسنجی یکپارچه نمودن کنترلرها در واحد فراهم‌آوری و ثبت و نیز در سامانه ویرایش بایستی در آینده نزدیک بررسی گردد.
۲. امکان اتصال به سامانه ثبت احوال فراهم گردد تا پس از وارد شدن کد ملی توسط متقاضی بلافاصله نام و نام خانوادگی فرد بصورت خودکار در سیستم درج گردد.
۳. در سامانه گنج‌ترتیبی داده خواهد شد که در زمان جستجوی عبارات یا اسامی هر دو حالت با فاصله یا بدون فاصله را در پاسخ به جستجو ارائه نماید. پسوند اگر معنادار نباشد (مانند پور، زاده و ...) نیم‌فاصله و اگر معنادار باشد فاصله داشته باشد. میتوان پسوندهای بی معنا را از پایگاه داده اداره ثبت احوال دریافت نمود.
۴. در سامانه ترتیبی ایجاد شده است که برخی پیشنندهای اضافه مانند مهندس، دکتر، حجه الاسلام پیش از ارسال به سامانه ویرایش بصورت خودکار حذف گردد. این مکانیزم میتواند به مرور با اضافه شدن کلمات مختلف تکمیل گردد.
۵. آموزشهایی در قالب TOOL TIP و یا دستورالعمل آنلاین برای آموزش کاربران در زمان تکمیل فیلدهای مختلف سامانه ثبت ایجاد شود.
۶. باتوجه به اینکه در وزارت بهداشت یک پایان‌نامه ممکن است توسط دو دانشجو نوشته شود و مذاکره در خصوص پایان‌نامه‌های این وزارتخانه در حال پیگیری است در سامانه برای این موضوع بایستی در آینده ترتیبی دیده شود.
۷. در صورت امکان اسکن کد ملی در سیستم از کاربر اخذ شود. این اسکن بصورت آنلاین توسط کاربر سامانه ثبت دیده شود و با فیلد مربوطه کنترل گردد.
۸. در سامانه ثبت این ویژگی دیده شود که حتی در صورت عدم انتخاب وابستگی سازمانی بصورت اتومات پس از انتخاب دانشگاه، وابستگی سازمانی انتخاب شود.

۹. در سامانه ثبت برای دانشگاههای پیام نور با توجه به حجم زیاد تعداد دانشگاهها
۱۰. بصورت TOOL TIP به کاربر گفته شود که اگر فیلدی را اشتباه وارد کرد پایان نامه عودت داده می شود.
۱۱. ساز و کاری دیده شود که یک نفر به عنوان رابط با مراجع ذیربط در ارتباط باشد و در صورت اضافه شدن یک واحد به دانشگاهها عنوان آن دانشگاه در سیستم اضافه شود.
۱۲. در صورت امکان، با موسسات آموزش عالی مختلف مکاتبه شود به محض اضافه شدن یک رشته گرایش در یک دانشگاه به ایرانداک اطلاع داده شود.
۱۳. در صورت امکان استفاده از پایگاههای داده سازمان سنجش که در آن اطلاعات هر دانشجو، شامل نام دانشگاه، نام دانشکده، رشته و گزارش کاملاً مشخص شده است در بهبودهای آتی سامانه ثبت دیده شود. بدین معنا که بصورت Auto Fill پس از وارد کردن کد ملی دانشجو این اطلاعات تکمیل گردد.
۱۴. ساختار دانشکدههای هر دانشگاه بصورت خوداظهاری از دانشگاهها اخذ و پس از اخذ ساختار در سیستم به روز گردد. پس از نهایی شدن فهرست دانشکدههای هر دانشگاه، امکان ثبت دستی از کاربر در فیلد دانشکده برداشته شود.
۱۵. بطور کلی اگر کاربر بصورت دستی فیلدی مانند دانشکده را تکمیل نمود، بایستی ترتیبی دیده شود که در سامانه ثبت به کاربر کنترل آلامر داده شود تا فیلد دستی تکمیل شده را کنترل نماید.
۱۶. تا حد امکان، بایستی تکمیل دستی فیلدها در کمترین حد ممکن در سامانه ثبت توسط دانشجوی تحصیلات تکمیلی صورت گیرد و بایستی در بلند مدت صفر شود.
۱۷. در فیلد "زبان اصلی پایان نامه" لازم است کل زبانهای موجود پشت سیستم قرار داده شود.
۱۸. باتوجه به اینکه رشته/ گرایشهای دانشگاه آزاد و وزارت علوم متفاوت هستند در آینده برای انتخاب فیلد رشته/ گرایش قبل از انتخاب این امکان فراهم گردد که کاربر وابستگی سازمانی را انتخاب نماید و بر اساس این وابستگی رشته/ گرایش را انتخاب نماید.

۱۹. بررسی و رفع مشکلات فعلی رشته/ گرایش‌ها در آینده در برنامه‌های مرکز قرار گیرد. مواردی مانند خط تیره، فواصل و نیم‌فاصله، غلط‌های املائی و ... در عنوان رشته/ گرایش‌ها نیز اصلاح گردد. همچنین رشته/ گرایش‌های مشابه و نیز یکپارچگی آنها می‌تواند در آینده مورد بررسی قرار گیرد.
۲۰. در حال حاضر در سامانه ثبت مدرکی دال بر دفاع دانشجوی تحصیلات تکمیلی که در سامانه ثبت پایان‌نامه خود را آپلود می‌کند از دانشجو دریافت نمی‌شود.
۲۱. تغییر فرایند موجود به گونه‌ای که تایید ایرنداگ، پس از تایید دانشگاه باشد می‌تواند اطمینان از نهایی شدن پایان‌نامه را در زمانیکه پایان‌نامه در سامانه ثبت آپلود می‌گردد، فراهم نماید. این کار تا حد زیادی می‌تواند به بهبود کیفیت پایان‌نامه‌های ثبت شده در سامانه ثبت منجر شود.
۲۲. بمنظور فراهم آوردن شرایطی که بتوان اطمینان ایجاد کرد که دانشجو در سامانه ثبت ایمیل استاد راهنما/ مشاور را صحیح تکمیل می‌کند بصورت TOOL TIP به دانشجو در زمان تکمیل مراحل در سامانه ثبت تذکر داده شود که ایمیلی بمنظور تایید برای استاد راهنما/ مشاور ارسال می‌گردد.
۲۳. بمنظور فراهم آوردن شرایطی که بتوان اطمینان ایجاد کرد که دانشجو در سامانه ثبت شماره تلفن استاد راهنما/ مشاور را صحیح تکمیل می‌کند بصورت TOOL TIP به دانشجو در زمان تکمیل مراحل در سامانه ثبت تذکر داده شود که پیامکی بمنظور تایید برای استاد راهنما/ مشاور ارسال می‌گردد.
۲۴. در سامانه ثبت در فیلد چکیده امکان وارد کردن تصویر فراهم گردد. این کار به ذخیره‌سازی فرمول‌های شیمیایی و ریاضی در سامانه ثبت کمک شایانی می‌کند. به عبارتی دانشجو در زمان وارد کردن فرمول ریاضی در فیلد چکیده در سامانه ثبت از نرم افزار (Editor) مربوط به اینکار استفاده میکند و پس از درج فرمول آن عبارت بصورت تصویر در چکیده ذخیره میگردد.
۲۵. بمنظور بهبود کیفیت چکیده در مواردی که دانشجو بایستی فرمول‌ها و روابط ریاضی را در چکیده وارد نماید می‌توان از یک دستورالعمل بعنوان راهنما استفاده نمود. به این معنا که فرمول‌ها و روابط را بصورت شماره‌گذاری شده (مثلاً فرمول شماره یک، فرمول شماره دو و ...) در چکیده وارد کند و در متن طبق همین شماره به آنها اشاره کند. این راهنما می‌تواند به محض تشخیص رشته دانشجو و اطمینان از اینکه وی نیاز به ثبت فرمول ریاضی در چکیده دارد برای کاربر نمایان خواهد شد. بعلاوه

امکان این نیز وجود دارد که این راهنما در اختیار نمایه‌سازان قرار داده شود و مراحل نمایه‌سازی نیز مطابق این رویه صورت گیرد.

۲۶. از آنجا که فهرست مطالب در سامانه گنج (نسخه بتا) نمایش داده می‌شود در دریافت فهرست مطالب از دانشجو در سامانه ثبت ملاحظات لازم فراهم گردد.

۲۷. باتوجه به عدم امکان کنترل فهرست مطالب در سامانه گنج و جلوگیری از نمایش داده بی کیفیت، فهرست مطالب از بخشهای نمایش داده شده در سامانه گنج بتا حذف شود.

۲۸. ایجاد یک وحدت رویه در نوشتن فهرست مطالب می‌تواند به بهبود کیفیت این فیلد در داده‌های دریافت شده در سامانه ثبت کمک ویژه‌ای نماید.

۲۹. در بهبود سامانه ثبت در خصوص ارتقاء کیفیت داده‌های پایان‌نامه/ رساله‌ها اقدامات زیر صورت گیرد.

- a. استفاده از چک‌باکس برای دانشجو در حالتیکه فهرست منابع فارسی یا لاتین ندارد.
- b. به شکل همزمان با بند قبلی یک TOOL TIP برای کاربر ایجاد شود که در زمان زدن تیک چک باکس بند قبلی به دانشجو هشدار داده شود در صورتیکه چک‌باکس فعال شود و داشتن فهرست مربوطه محرز شود پایان‌نامه/ رساله عودت داده می‌شود. و این موضوع در توسط همکاران کنترل چک شود.

۳۰. بطور کلی در ابتدای ثبت پایان‌نامه/ رساله توسط دانشجو، یک صفحه باز شود و به دانشجو اعلام گردد که محتوای ارسالی به سامانه توسط کارشناسان کنترل می‌گردد و در صورت مشاهده هرگونه مغایرت میان آنچه ثبت شده و آنچه آپلود شده مشاهده شود مدرک عودت داده می‌شود. در این خصوص استفاده از گرافهایی مانند نمودارهای ویزیو و ... می‌تواند مفید واقع شود.

۳۱. استفاده از اصطلاحنامه‌های موجود ایرانداک بمنظور تهیه معادل انگلیسی از کلیدواژه فارسی و بالعکس بصورت اتومات از میان کلیدواژه‌های کنترل شده

۳۲. اضافه نمودن حداقل تعداد کلیدواژه‌ها (انگلیسی یا فارسی) از ۲ به ۳ در طراحی آتی سامانه ثبت

۳۳. اضافه نمودن حداکثر تعداد کلیدواژه‌ها (انگلیسی یا فارسی) به ۸ در طراحی آتی سامانه ثبت

۳۴. استفاده از تکنیک‌هایی مانند Text Mining و پیشنهاد کلید واژه به دانشجو پس از آپلود چکیده و انتخاب توسط دانشجو می‌تواند در توسعه‌های آتی سامانه ثبت مورد توجه قرار گیرد.

۳۵. در آینده می‌توان در توسعه سامانه ثبت امکان ارسال هدفمند و تخصیص هوشمند پایان‌نامه/رساله را فراهم نمود.

۳۶. در سامانه ثبت در زمان ثبت پایان‌نامه/رساله و در زمان آپلود سایر فایلها از دانشجو خواسته شود نوع فایل یا نرم‌افزاری که برای بازکردن آن لازم است را از یک فهرست انتخاب نماید. این موضوع در ویرایش‌های آتی سامانه ثبت فراهم گردد.

۳۷. در کنترل‌های سامانه ثبت توسط کارشناسان کنترل‌کننده، یک قانون کلی این است که محتوای فایل پیوست شده تغییر نیابد و تغییرات صورت گرفته در حد حذف مواردی مانند اینترها و فواصل اضافه است.

۳۸. در خصوص فایل‌های پیوست شده در سامانه ثبت ایجاد یک راهنما یا دستورالعمل در زمان آپلود فایل بمنظور کاهش حجم یا دسته‌بندی فایل‌های پایان‌نامه به کاهش مشکلات آتی کمک نماید. بعنوان مثال می‌توان فایل‌های عکس حجیم را در بخش فایل‌های سایر (etc) قرار داد. این موضوع برای برخی رشته‌ها مانند معماری که عکس‌های با حجم زیاد در پایان‌نامه وجود دارد به دانشجو در زمان آپلود فایل‌ها کمک می‌کند.

۳۹. تهیه چکلیست از طرح کیفیت فراداده و ارائه آن به دانشجو در زمان تکمیل فیلدهای سامانه ثبت کمک می‌کند که در زمان تکمیل سامانه ثبت کیفیت آیتم‌های مختلف بهبود یابد و دانشجو از مواردی که در ایرانداک کنترل می‌گردد اطلاع یابد.

۴۰. ارائه راهنمایی‌هایی مانند یک فایل آموزشی تصویری به کاربر در زمان تکمیل فیلدهای مختلف سامانه ثبت کمک شایانی خواهد نمود.

۴۱. بمنظور کاهش مشکلات مختلف فایل پی‌دی‌اف ارسالی توسط دانشجویان راهکارهای مختلف زیر را می‌توان انجام داد.

a. راهنمایی کردن کاربر به سویی که با برخی نرم‌افزارهای خاص پی‌دی‌اف کننده این کار را انجام دهند.

b. این موضوع توسط کارشناسان کنترل گردد. بعنوان مثال با پی‌دی‌اف‌ساز ورد این کار انجام گردد.

c. ارائه نرم‌افزار یکسان پی‌دی‌اف‌ساز به دانشگاه و اجبار دانشجویان جهت استفاده از این نرم‌افزار. بدین منظور لازم است مطالعه اولیه‌ای در خصوص کیفیت خروجی، نقاط ضعف و قوت همه پی‌دی‌اف‌سازها صورت گیرد.

d. دریافت فایل ورد با کیفیت از دانشجو و تهیه فایل پی‌دی‌اف در ایرانداک با کیفیت مناسب

۴۲. در برخی از موارد که فراداده پایان‌نامه / رساله‌های قبلی ناقص است به کمک OCR نمودن صفحه اول فایل آن مدرک می‌توان تا دقت بالایی نواقص را برطرف نمود.

۴۳. ارائه ویرایش بعدی طرح کیفی می‌تواند موقوف به اندازه‌گیری کیفیت فراداده‌های موجود در سامانه ثبت و مشاهده مشاهده شده در این فرایند گردد.

۴۴. مشاهده و تجزیه تحلیل مشکلاتی که در سامانه پیشینه، تالار جستجوی اشاعه، کاربران گنج و ... دیده شده‌اند به بهبود کیفیت در سامانه ثبت منجر خواهد شد.

۴۵. توسعه فرایندهای موجود در سامانه ثبت بگونه‌ای که کنترل‌های صورت گرفته در این سامانه در یک ساز و کار مشخصی مجدداً بصورت تصادفی کنترل گردد به بهبود خروجی پایان‌نامه / رساله‌های اشاعه داده شده منجر خواهد شد.

۴۶. تدوین شاخص‌های کیفی بمنظور مقایسه وضعیت کیفیت داده‌های پایان‌نامه / رساله پیش از بهبود کیفیت سامانه (تدوین و پیاده‌سازی طرح کیفی) و پس از آن به بهبود مستمر سامانه در آینده کمک شایانی خواهد کرد. (بعنوان مثال میزان عدم انطباق‌های نوع A پیش و پس از استقرار سیستم کنترل کیفیت داده یکی از این شاخصهاست)

۴۷. در آینده در صورت امکان نصب نرم‌افزار مناسب برای مشاهده Latex برای همکاران کنترل کننده سامانه ثبت میتوان موارد کنترلی فایل ورد را عیناً در فایل Latex کنترل نمود.

۴۸. ارائه دسته‌بندی برای سایر فایل‌های ضمیمه شده توسط دانشجویان (مانند کدها، عکسها، نقشه‌ها، فیلمها و ...) به مدیریت این دسته از فایل‌ها در آینده کمک ویژه‌ای خواهد نمود. در زمان آپلود

فایلها نوع فایل را میتوان از دانشجو پرسش نموده و دسته‌بندی کرد. مکانیزم آرشیو این دسته از فایل‌ها بایستی طراحی شود.

۴۹. ایجاد ساز و کاری که طی آن از بصورت منظم مجموعه فایل‌های پایان‌نامه/رساله بک‌آپ گرفته شود. همچنین مطابق نظم و قواعد مشخصی دسته‌بندی گردند.

۵۰. بررسی فرمت ایمیل‌های کاربران مانند قالب "IR." و همچنین فرمت شماره ملی کاربران بک‌مک الگوریتم سازمان ثبت احوال و نیز الگوی شماره تلفن موبایل بصورت خودکار در سامانه ثبت در بهبودهای آتی این سامانه پیشنهاد می‌گردد.

همچنین موارد زیر توسط کارشناسان اداره ثبت و فراهم‌آوری اطلاعات بمنظور بهبود کیفیت سامانه مطرح شده است.

- برخی از اقدامات خطا ناپذیرساز لازم است بصورت فرایندی صورت گیرد. مثلاً بایستی شماره ملی که در اطلاعات اولیه از فرد گرفته میشود در ادامه همان ثبت گردد.
- چند اسم‌ه بودن نویسنده مثلاً نویسندگان عربی سه اسم‌ه هستند
- عربی بودن متن پایان‌نامه در دانشگاه جامعه المصطفی
- در سامان ثبت، در حالت رد ایرانداک، بایستی در منوی کرکره‌ای علت‌های رد ایرانداک ذکر شود و بتوان تحلیل‌های مناسبی انجام داد.
- عبارت "پذیرش ایرانداک" هم در ابتدای بررسی مدرک و هم در انتهای بررسی مدرک در سامان مورد استفاده قرار گرفته است. شاید در ابتدا عبارت "بررسی مدارک" مناسبتر باشد.
- در سامانه ترتیبی داده شود وقتی کاربر به هر دلیل "نام پژوهشکده/ دانشکده" را خودش اضافه می‌نماید در ایرانداک کارشناس مربوط به پایگاه اسامی پژوهشکده/ دانشکده از آن اطلاع یافته و پس از بررسی در سامانه اضافه نماید.
- مورد قبل در خصوص اضافه شدن نام رشته نیز می‌تواند لحاظ شود.

- قید شدن کلیدواژه فارسی در بخش کلید واژه فارسی و کلید واژه انگلیسی در بخش کلید واژه انگلیسی
- مشخص شدن عنوان کلیدواژه فارسی و کلیدواژه انگلیسی در بالای هر کدام
- در سامانه ترتیبی داده شود که مشکلات پایان نامه به تدریج و مرحله و مرحله در یک پنجره ثبت شود. در حال حاضر اگر در پنجره رد یک نکته اضافه شود و از یک تب به تب دیگر روند مطلب نوشته شده حذف می‌گردد.
- اگر دانشجو، دانشگاه را وارد کرد اتومات وابستگی سازمانی را اصلاح و درج نماید.
- ایجاد امکان اصلاح بعد از چند دقیقه از اظهارنظر در مورد سامانه ثبت (در وضعیت فعلی بلافاصله خارج شده و به مرحله بعد می‌رود)
- ترتیب آیتمهایی که چک میشود با ترتیب قرارگیری در پایان نامه یکی نیست. مثلاً الان اول اسم است بعد دانشگاه در حالیکه در سامانه ثبت برعکس است.
- در بخش نام و نام خانوادگی در سامانه امکان وارد شدن اسم لاتین را ندهد.
- ایجاد 'FAQ برای برخی موارد شامل
 - در جامعه المصطفی بایستی سایر دستگاههای اجرایی انتخاب شود در حالیکه حوزه های علمیه نوشته میشود.
 - دانشگاه علامه و پژوهشگاه علامه جابجا میشوند.
- ممکن است کارشناس در سیستم رد یا پذیرش مدرک را بزند ولی از کار تابلش خارج نشود
- مشخص نیست برخی متن ها که بایستی تیک زده شوند چه استفاده ای دارند مثلاً فرمول ریاضی در عنوان یا تقاضا محور بودن پایان نامه
- در فهرست زبانهای پایان نامه زبان اردو نیست در حالیکه در جامعه المصطفی این زبان هست.
- شماره ملی اساتید راهنما و مشاور و رایانامه آنها اجباری نیست
- در چکیده اسپیس به عنوان کاراکتر قبول میکند و لذا امکان چکیده خالی باشد
- در سیستم به جای چکیده فارسی عربی وارد شود سیستم قبول میکند

- ارتباط میان کلید واژه فارسی و معادل انگلیسی آن در سیستم دیده شود
- در صورتیکه در سیستم اسم دانشگاه نبود امکان پیشنهاد دادن به سیستم توسط دانشجو باشد
- در صورت امکان فهرست‌های فایل پی دی اف بصورت لینک باشد
- نام و نام خانوادگی و کدملی قابل اصلاح شدن نیست در سیستم
- در صورتیکه نماینده دانشگاه وجود نداشته باشد در سیستم میماند
- نام، نام خانوادگی، کدملی و ایمیل مواردی هستند که امکان ویرایش برای آنها وجود ندارد.
- نام دانشگاه استاد راهنما هم در سامانه اخذ شود ممکن است دانشگاه استاد راهنما و دانشگاه دانشجو متفاوت باشد

فهرست منابع

منابع فارسی

- ارشادی، محمدجواد؛ احترامی، تینا. (۱۳۹۵)، طرحریزی و استقرار نظام تضمین کیفیت در سامانه‌های گردآوری و ثبت، سازمان‌دهی، و اشاعه اطلاعات پایان‌نامه‌ها/ رساله‌های دانش‌آموختگان داخل کشور. تهران، پژوهشگاه علوم و فناوری اطلاعات ایران.
- ارشادی، محمدجواد؛ رجبی، تقی؛ شیرانی، فرهاد و رضایی، نساء. (۱۳۹۴)، کاربرد تکنیک تحلیل ریشه در حل مشکلات کیفی سامانه‌های اطلاعاتی تحقیقاتی: مطالعه موردی سامانه اشاعه اطلاعات پایان‌نامه‌ها/ رساله‌های دانش‌آموختگان داخل کشور (گنج). نشریه مدیریت اطلاعات، ۱.
- علیدوستی، سیروس. و عصاره، حسن. ۱۳۸۴. طراحی فرایند کنترل کیفیت نمایه سازی، گزارش طرح پژوهشی، تهران، پژوهشگاه اطلاعات و مدارک علمی ایران.

منابع لاتین

- Abiteboul, S., Buneman, P., and Suciu, D. 2000. Data on the Web: From Relations to Semistructured Data and XML. Morgan Kaufmann Publishers.
- Aiken, P. 1996. Data Reverse Engineering. McGraw Hill.
- Atzeni, P., & De Antonellis, V. (1993). Relational database theory. Benjamin-Cummings Publishing Co., Inc.

Arenas, M., Bertossi, L., and Chomicki, J. 1999. Consistent query answers in inconsistent databases. In Proceedings of the 18th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS). ACM, New York, 68–79.

Atzeni, P., Merialdo, P., and Sindoni, G. 2001. Web site evaluation: Methodology and case study. In Proceedings of International Workshop on data Semantics in Web Information Systems (DASWIS).

Ballou, D. and Pazer, H. 1985. Modeling data and process quality in multi-input, multi-output information systems. *Manag. Sci.* 31, 2.

Ballou, D., Wang, R., Pazer, H., & Tayi, G. K. (1998). Modeling information manufacturing systems to determine information product quality. *Management Science*, 44(4), 462-484.

Basile, A., Batini, C., Grega, S., Mastrella, M., and Maurino, A. 2007. Orme: A new methodology for information quality and basel II operational risk. In Proceedings of the 12th International Conference of Information Quality, Industrial Track.

Basili, V., Caldiera, C., Rombach, H. 1994. Goal question metric paradigm.

Baskarada, S., Koronios, A., and Gao, J. 2006. Towards a capability maturity model for information quality management: a tqdm approach. In Proceedings of the 11th International Conference on Information Quality.

Batini C., Cappiello C., Francalanci C. and Maurino A. 2009. Methodologies for data quality and assessment and Improvement. *ACM computing surveys*, 41(3).

Batini, C. and Scannapieco, M. 2006. *Data Quality: Concepts, Methodologies and Techniques*. Springer Verlag.

Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. *ACM computing surveys (CSUR)*, 41(3), 16.

Bertolazzi, P., Santis, L. D., and Scannapieco, M. 2003. Automatic record matching in cooperative information systems. In Proceedings of the ICDT International Workshop on Data Quality in Cooperative Information Systems (DQCIS).

Bettschen, P. 2005. Master data management (MDM) enables IQ at Tetra Pak. In Proceedings of the 10th International Conference on Information Quality.

Bilke, A., Bleiholder, J., Bohm, C., Draba, K., Naumann, F., and Weis, M. September 2005. Automatic data fusion with HumMer. In Proceedings of the VLDB Demonstration Program.

Bovee, M., Srivastave, R., and Mak, B. 2001. A conceptual framework and belief-function approach to assessing overall information quality. In Proceedings of the 6th International Conference on Information Quality.

Buneman, P. 1997. Semi-structured data. In Proceedings of the 16th ACM Symposium on Principles of Database Systems (PODS).

Buneman, P., Khanna, S., and Tan, W. C. 2001. Why and where: A characterization of data provenance. International Conference on Database Theory.

CAL`1, A., Calvanese, D., De Giacomo, G., and Lenzerini, M. 2004. Data integration under integrity constraints. Inform. Syst. 29, 2, 147–163.

Calvanese, D., De Giacomo, D., and Lenzerini, M. 1999. Modeling and querying semi-structured data. Network. Inform. Syst. J. 2, 2, 253–273.

Cappiello, C., Francalanci, C., and Pernici, B. 2003. Preserving Web sites: A data quality approach. In Proceedings of the 7th International Conference on Information Quality (ICIQ).

Cappiello, C., Francalanci, C., Pernici, B., Plebani, P., AND SCANNAPIECO, M. 2003b. Data quality assurance in cooperative information systems: a multi-dimension certificate. In Proceedings of the ICDT International Workshop on Data Quality in Cooperative Information Systems (DQCIS).

Catarci, T., and Scannapieco, M. 2002. Data quality under the computer science perspective. Archivi Computer 2.

Chapman, A., Richards, H., and Hawken, S. 2006. Data and information quality at the Canadian institute for health information. In Proceedings of the 11th International Conference on Information Quality.

Chengalur-Smith, I. N., Ballou, D. P., & Pazer, H. L. (1999). The impact of data quality information on decision making: an exploratory analysis. IEEE Transactions on Knowledge and Data Engineering, 11(6), 853-864.

Corey, D., Cobler, L., Haynes, K., and Walker, R. 1996. Data quality assurance activities in the military health services system. In Proceedings of the 1st International Conference on Information Quality. 127–153.

Dasu, T., & Johnson, T. (2003). Exploratory data mining and data cleaning (Vol. 479). John Wiley & Sons.

Data Warehousing Institute. 2006. Data quality and the bottom line: Achieving business success through a commitment to high quality data. <http://www.dw-institute.com/>

De Amicis, F., & Batini, C. (2004). A methodology for data quality assessment on financial data. *Studies in Communication Sciences*, 4(2), 115-137.

De Amicis, F., Barone, D., and Batini, C. 2006. An analytical framework to analyze dependencies among data quality dimensions. In *Proceedings of the 11th International Conference on Information Quality (ICIQ)*. 369–383.

De Michelis, G., Dubois, E., Jarke, M., Mattes, F., Mylopoulos, J., Papazoglou, M., Pohl, K., Schmidt, J., Woo, C., and Yu, E. 1997. *Cooperative Information Systems: A Manifesto*. In *Cooperative Information Systems: Trends & Directions*, M. Papazoglou and G. Schlageter, Eds. Academic-Press.

De Santis, L., Scannapieco, M., and Catarci, T. 2003. Trusting data quality in cooperative information systems. In *Proceedings of the 11th International Conference on Cooperative Information Systems (CoopIS)*. Catania, Italy.

Dedeke, A. 2005. Building quality into the information supply chain. *Advances in Management Information Systems-Information Quality Monograph (AMIS-IQ) Monograph*. R. Wang, E. Pierce, S. Madnick, and Fisher C.W., Eds.

DQI. 2004. Data quality initiative framework. Project report. [www.wales.nhs.uk/sites/documents/319/DQI Framwork Update Letter 160604.pdf](http://www.wales.nhs.uk/sites/documents/319/DQI_Framwork_Update_Letter_160604.pdf)

English, L. (2009). *Information quality applied: best practices for improving business information, processes, and systems*. Indianapolis: Wiley.

English, L. 1999. *Improving Data Warehouse and Business Information Quality*. Wiley & Sons.

English, L. 2002. Process management and information quality: how improving information production processes improved information (product) quality. In *Proceedings of the 7th International Conference on Information Quality (ICIQ)*. 206–211.

Eppler, M. and Helfert, M. 2004. A classification and analysis of data quality costs. In *Proceedings of the 9th International Conference on Information Systems (ICIQ)*.

Eppler, M. and Munzenmaier, P. 2002. Measuring information quality in the Web context: A survey of state-of-the-art instruments and an application methodology. In *Proceedings of the 7th International Conference on Information Systems (ICIQ)*.

Falorsi, P. D., & Righi, P. (2008). A balanced sampling approach for multi-way stratification designs for small area estimation. *Survey Methodology*, 34(2), 223-234.

Fellegi, I. P., and Holt, D. 1976. A systematic approach to automatic edit and imputation. *J. Amer. Stat. Assoc.* 71, 353, 17–35.

Fisher, C. W., & Kingma, B. R. (2001). Criticality of data quality as exemplified in two disasters. *Information & Management*, 39(2), 109-116.

FRATERNALI, P., LANZI, P., MATERA, M., AND MAURINO, A. 2004. Model-driven Web usage analysis for the evaluation of Web application quality. *J. Web Eng.* 3, 2, 124–152.

Gackowski, Z. 2006. Redefining information quality: the operations management approach. In *Proceedings of the 11th International Conference on Information Quality (ICIQ)*. 399–419.

Hammer, M. 1990. Reengineering work: Don't automate, obliterate. *Harvard Bus. Rev.* 104–112.

Hammer, M. and Champy, J. 2001. *Reengineering the Corporation: A Manifesto for Business Revolution*, Harper Collins.

Hernandez, M. and Stolfo, S. 1998. Real-world data is dirty: Data cleansing and the merge/purge problem. *J. Data Min. Knowl. Dis.* 1, 2.

Inmon, W. H. (1992), *Building the Data Warehouse*, New York, NY: Wiley.

Isakowitz, T., Stohr, E. A., & Balasubramanian, P. (1995). RMM: a methodology for structured hypermedia design. *Communications of the ACM*, 38(8), 34-44.

Isakowitz, T., Stohr, E., and Balasubramanian, P. 1995. RMM: A methodology for structured hypermedia design. *Comm. ACM* 58, 8.

ISTAT. 2004. Guidelines for the data quality improvement of localization data in public administration (in Italian). www.istat.it

Jarke, M., Gellersdörfer, R., Jeusfeld, M. A., Staudt, M., & Eherer, S. (1995). ConceptBase—a deductive object base for meta data management. *Journal of Intelligent Information Systems*, 4(2), 167-192.

Jarke, M., Lenzerini, M., Vassiliou, Y., and Vassilioudis, P., . 1995. *Fundamentals of Data Warehouses*. Springer Verlag.

Jeusfeld, M., Quix, C., and Jarke, M. 1998. Design and analysis of quality information for data warehouses. In *Proceedings of the 17th International Conference on Conceptual Modeling*.

Kerr, K. and Norris, T. 2004. The development of a healthcare data quality framework and strategy. In *Proceedings of the 9th International Conference on Information Quality*.

Kettinger, W. and Grover, V. 1995. Special section: Toward a theory of business process change management. *J. Manag. Inform. Syst.* 12, 1, 9–30.

Kovac, R. and Weickert, C. 2002. Starting with quality: Using TDQM in a start-up organization. In *Proceedings of the 7th International Conference on Information Quality (ICIQ)*. Boston, 69–78.

Lee, Y.W., Strong, D. M., Kahn, B. K., and Wang, R. Y. 2002. AIMQ: A methodology for information quality assessment. *Information Management*. 40, 2, 133–460.

Lenzerini, M. 2002. Data integration: A theoretical perspective. In *Proceedings of the 21st ACM Symposium on Principles of Database Systems (PODS)*.

Liu, L. and Chi, L. 2002. Evolutionary data quality. In *Proceedings of the 7th International Conference on Information Quality*.

Long, J. A., & Seko, C. E. (2005). A cyclic-hierarchical method for database data-quality evaluation and improvement. *Information quality*, 1, 52.

Long, J. and Seko, C. April 2005. A cyclic-hierarchical method for database data-quality evaluation and improvement. In *Advances in Management Information Systems-Information Quality Monograph (AMISIQ)*

Loshin, D. 2004. *Enterprise Knowledge Management - The Data Quality Approach*. Series in Data Management Systems, Morgan Kaufmann, chapter 4.

Luisa De Biagi, Roberto Puccinelli, Massimiliano Saccone, Luciana Trufelli. 2014. Research product repositories: strategies for data and metadata quality control. National Research Council of Italy.

Lyman, P. and Varian, H. R. 2003. How much information. <http://www.sims.berkeley.edu/how-much-info-2003>.

Madnick, S. and Wang, R. Y. 1992. Introduction to total data quality management (TDQM) research program. TDQM-92-01, Total Data Quality Management Program, MIT Sloan School of Management.

Mecca, G., Atzeni, P., Masci, M., Merialdo, P., and Sindoni, G. 1998. The AraneusWeb-based management system. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, L. M. Haas and A. Tiwary, Eds. ACM Press, 544–546.

MECCA, G., Merialdo, P., Atzeni, P., and Crescenzi, V. 1999. The (short) araeneus guide toWeb site development. In *Proceedings of the 2nd International Workshop on theWeb and Databases (WebDB) Conjunction with Sigmod*.

Motro, A. and Anokhin, P. 2005. Fusionplex: Resolution of data inconsistencies in the data integration of heterogeneous information sources. *Inform. Fusion*, 7, 2, 176–196.

Muthu, S., Withman, L., and Cheraghi, S.H. 1999. Business process re-engineering : a consolidated methodology. In *Proceedings of the 4th annual International Conference on Industrial Engineering Theory, Applications and Practice*.

Nadkarni, P. 2006. Delivering data on time: The assurant health case. In *Proceedings of the 11th International Conference on Information Quality*.

Naumann, F. 2002. Quality-driven query answering for integrated information systems. *Lecture Notes in Computer Science*, vol. 2261.

Nelson, R. Ryan, Peter A. Todd, and Barbara H. Wixom. 2005. Antecedents of information and system quality: an empirical examination within the context of data warehousing. *Journal of management information systems*. 21(4), 199-235.

Oakland, J. 1989. *Total Quality Management*. Springer.

Office of Management and Budget. 2006. Information quality guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated by agencies. <http://www.whitehouse.gov/omb/fedreg/reproducible.html>.

Pernici, B. and Scannapieco, M. 2003. Data quality in Web information systems. *J. Data Semant.* 1, 48–68.

Pipino, L., Lee, Y., and Wang, R. 2002. Data quality assessment. *Commun. ACM* 45, 4.

Raghunathan, S. 1999. Impact of information quality and decision-maker quality on decision quality: a theoretical model and simulation analysis. *Decis. Supp. Syst.* 26, 275–286.

Rahm, E., Thor, A., AUMULLER, D., Hong-Hai, D., Golovin, N., and Kirsten, T. June 2005. iFuice information fusion utilizing instance correspondences and peer mappings. In *Proceedings of the 8th International Workshop on the Web and Databases (WebDB)*. located with SIGMOD.

Rao, R. 2003. From unstructured data to actionable intelligence. *IT Professional* 535, 6, 29–35.

Redman, T. 1996. *Data Quality for the Information Age*. Artech House.

Redman, T. 1998. The impact of poor data quality on the typical enterprise. *Comm. ACM* 41, 2, 79–82.

Richard W., Stuart E. Madnick. 1990. A Polygen Model for Heterogeneous Database Systems: The Source Tagging Perspective, Proceedings of the 16th International Conference on Very Large Data Bases, 519-538.

Richard Y. Wang, Diane M. Strong. 1996. Beyond accuracy: what data quality means to data consumers, Journal of Management Information Systems, 12 (4), 5-33.

Richard Y. Wang, Henry B. Kon, Stuart E. Madnick. 1993. Data Quality Requirements Analysis and Modeling, Proceedings of the Ninth International Conference on Data Engineering, 670-677.

Scannapieco, M., Virgillito A., Marchetti, M., Mecella, M., and Baldoni, R. 2004. The DaQuinCIS architecture: a platform for exchanging and improving data quality in Cooperative Information Systems. Inform. Syst. 29, 7, 551–582.

Scannapieco, M., Perinci, B., and Pierce, E. 2002. IP-UML: Towards a Methodology for Quality Improvement based on the IP-MAP Framework. In Proceedings of the 7th International Conference on Information Quality (ICIQ). Boston.

Scannapieco, M., Pernici, B., and Pierce, E. 2005. IP-UML: A methodology for quality improvement-based on IP-MAP and UML. In Information Quality, Advances in Management Information Systems, Information Quality Monograph (AMIS-IQ), R. Wang, E. Pierce, S. Madnik, and C. Fisher, Eds.

Sessions, V. 2007. Employing the TDQM methodology: An assessment of the SC SOR. In Proceedings of the 12th International Conference on Information Quality. 519–537.

Shankaranarayanan, G., Wang, R. Y., and Ziad, M. 2000. Modeling the manufacture of an information product with IP-MAP. In Proceedings of the 6th International Conference on Information Quality (ICIQ 2000). Boston.

Shankaranarayanan, G. and Wang, R. 2007. IPMAP: Current state and perspectives. In Proceedings of the 12th International Conference on Information Quality.

Sheng, Y. 2003. Exploring the mediating and moderating effects of information quality on firm's endeavor on information systems. In Proceedings of the 8th International Conference on Information Quality 2003 (ICIQ). 344–352.

Sheng, Y. and Mykytyn, P. 2002. Information technology investment and firm performance: A perspective of data quality. In Proceedings of the 7th International Conference on Information Quality (ICIQ). DC, 132–141.

Stoica, I., Morris, R., Liben-Nowell, D., Karger, D. R., Kaashoek, M. F., Dabek, F., & Balakrishnan, H. (2003). Chord: a scalable peer-to-peer lookup protocol for internet applications. *IEEE/ACM Transactions on Networking (TON)*, 11(1), 17-32.

Storey, V. and Wang, R. Y. 1998. Modeling quality requirements in conceptual database design. *Proceedings of the International Conference on Information Quality*, 64—87.

Su, Y. and Jin, Z. 2004. A methodology for information quality assessment in the designing and manufacturing processes of mechanical products. In *Proceedings of the 9th International Conference on Information Quality (ICIQ)*. 447—465.

Vassiliadis, P., Vagena, Z., Skiadopoulos, S., Karayannidis, N., and Sellis, T. 2001. ARTKOS: toward the modeling, design, control and execution of ETL processes. *Inform. Syst.* 26, 537–561.

Vermeer, B. 2000. How important is data quality for evaluating the impact of edi on global supply chains. In *Proceedings of the 33rd Hawaii Conference on Systems Sciences*.

Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*, 39(11), 86-95.

WANG, R. 1998. A product perspective on total data quality management. *Comm. ACM* 41, 2.

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.

Zachman, J. 2006. Zachman institute for framework advancement (ZIFA). www.zifa.com.

Abstract

Registration system of theses and dissertations of graduated students plays as a critical system for dissemination of scientific records and therefore an inevitable role for development of researches. Therefore suitable quality level of this system has main effect on augmenting the quality of researches. Some main fields of this system are such as Name, National code, Grade, University branch, Supervisor, Proposal characteristics and Full text characteristics. For improving the quality level of this system determining of appropriate level of these mentioned fields is essential. For this reason different data quality models are extracted from literature and the good condition of all of fields are mapped. Current quality level of all fields, policies and strategies on data quality are completely studied in this project. The roles of each stakeholder in the registrations system such as student and university representative are determined through this project. A framework for grading the nonconformities and also corrective actions are illustrated. These actions are such as corrections of metadata, removing, accepting and finally sending to indexing section. Produced documents of this project is called quality plan.

Keywords: Metadata; Quality plan; Data quality, Research information system



Iranian Research Institute for Information Science and Technology

Research Project

Designing A Metadata Quality Model on Registration System of Theses and Dissertations

By:

Mohammad Javad Ershadi

Co-Author:

Dr. Jalal A. Nasiri

Dr. Farhad Shirani

Winter 2018