



وزارت علوم، تحقیقات و فناوری
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)

طرح پژوهشی

**ارائه طرح نمونه برداری پذیرش در فرایند ویراستاری و نمایه سازی
پایان نامه ها و رساله ها
(گزارش نهایی)**

مجری

محمد جواد ارشادی

همکار

زهرا دهسرایبی

آذر ۱۳۹۹

حقوق: پژوهشگاه علوم و فناوری اطلاعات ایران

طرح پژوهشی "ارائه طرح نمونه‌برداری پذیرش در فرایند ویراستاری و نمایه‌سازی پایان‌نامه‌ها و رساله‌ها" پیرو قرارداد شماره ۳۳/۶۶۱۸ مورخ ۱۳۹۸/۰۷/۰۲ میان پژوهشگاه علوم و فناوری اطلاعات ایران (کارفرما) و آقای محمدجواد ارشادی اجرا شده است. گزارش حاضر یک جلد از مستندات این طرح است.

این گزارش و تمامی حقوق مادی آن بر اساس قانون حمایت حقوق مولفان و مصنفان و هنرمندان، مصوب ۱۳۴۸ و اصلاحیه‌های بعدی آن و همچنین آیین‌نامه‌های اجرایی این قانون متعلق به پژوهشگاه علوم و فناوری اطلاعات ایران و هرگونه 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 بر عهده مجری طرح پژوهشی است.

۳۳/۸۰۷۷

۱۳۹۹/۰۹/۱۹

ندارد

بنام خدا

ایرانداک در نیم قرن دوم کار خود همچنان جوان و کوشا، همراه هر پژوهش، پژوهشگر، و سیاست‌گذار

۱۳۴۲-۱۳۹۹

گواهی

انجام طرح پژوهشی:

◇ عنوان: ارائه طرح نمونه‌برداری پذیرش در فرایند ویراستاری و نمایه‌سازی پایان‌نامه‌ها و رساله‌ها

◇ مجری: دکتر محمدجواد ارشادی

◇ همکار: زهرا دهنسرای

◇ شماره قرارداد: ۳۳/۷۰۰۸

◇ تاریخ قرارداد: ۱۳۹۸/۷/۲

◇ تاریخ شروع: ۱۳۹۸/۷/۲

◇ تاریخ پایان: ۱۳۹۹/۹/۸

◇ ناظر: دکتر آرمان ساجدی‌نژاد

در پژوهشگاه علوم و فناوری اطلاعات ایران گواهی می‌شود. گزارش پایانی این طرح، بر پایه داوری در نشست ۳۹۴ (تاریخ ۱۳۹۹/۹/۱۸) شورای پژوهشی پژوهشگاه به تصویب رسیده است. لازم می‌دانم مراتب قدردانی خود را به مجری، همکار، و ناظر این طرح تقدیم دارم.

با آرزوی پیروزی
پرویز شهریاری
معاون پژوهش و آموزش

ارائه طرح نمونه‌برداری پذیرش در فرایند ویراستاری و نمایه‌سازی پایان‌نامه‌ها و رساله‌ها

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

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

صندوق پستی: ۱۳۱۵۷۷۳۳۱۴ تلفن: ۰۲۱-۶۶۴۹۴۹۸۰ (داخلی ۴۷۸)

وبگاه: <http://irandoc.ac.ir/u/mj-ershadi>

رایانامه: Ershadi@irandoc.ac.ir

همکار(ان) اصلی: زهرا دهنسرای

چکیده

در مدارک پایان‌نامه/رساله (پارسا) دانش‌آموختگان کل کشور نمایه‌سازی یک فرایند کلیدی است که در صورت انجام درست، امکان بازیابی مناسب را در سامانه گنج فراهم می‌سازد. وضعیت موجود ثبت، نمایه‌سازی، و ویراستاری مدارک پارسا بدین گونه است که پس از ثبت مدارک توسط دانشجویان تحصیلات تکمیلی در سامانه ثبت و کنترل نهایی توسط همکاران واحد فراهم‌آوری، مدارک تایید شده برای انجام نمایه‌سازی به صورت خودکار وارد سامانه‌ای با نام «ویرایش و سازماندهی اطلاعات علمی» می‌شوند. در این سامانه، افزون بر کنترل دوباره همه اقلام اطلاعاتی پایان‌نامه/رساله‌ها، نمایه‌سازی انجام شده و سپس به همکاران بخش ویراستاری ارسال می‌شود. در این بخش، بر روی نمایه‌سازی‌های صورت گرفته کنترل نهایی انجام شده و در صورت وجود مشکل بازگردانده شده و در غیر این صورت برای اشاعه به سامانه گنج ارسال می‌شود. در وضعیت موجود علاوه بر کنترل‌های قبلی، همه مدارک دوباره توسط مسئول واحد بمنظور ارزیابی کیفیت کار انجام شده مورد بررسی قرار می‌گیرد. از این رو همانگونه که مشاهده می‌شود بر روی برخی اقلام ممکن است چندین کنترل صورت پذیرد. اهمیت این پژوهش زمانی بیشتر خواهد شد که در سامانه ثبت جدید از آنجا که برخی اقلام اطلاعاتی (مانند نام دانشگاه، رشته و گرایش و...) به صورت فهرست از پیش تعیین شده توسط کاربر انتخاب می‌شود، امکان بروز خطا وجود ندارد. از این رو نیازی به کنترل این اقلام در سامانه ویرایش و سازماندهی اطلاعات وجود ندارد. در این پژوهش چارچوبی برای نمونه‌برداری در فرایندهای نمایه‌سازی و ویراستاری فراهم شده است. در اجرای این پژوهش، در مرحله ارائه روش نمونه‌برداری افزون بر استفاده از روش گروه تمرکزی، به منظور طراحی آماری از یک مدل بهینه‌سازی برای کمینه‌سازی خطاهای نوع I و II استفاده شده است. نتایج این پژوهش نشان داد، با توجه به کیفیت قابل قبول بسیاری از اقلام اطلاعاتی پارسا پس از نمایه‌سازی، نمونه‌برداری کاری حیاتی در ارتقاء کارایی واحد سازماندهی و تحلیل اطلاعات است. دسته‌بندی اقلام

اطلاعاتی به سه دسته حیاتی، اصلی و جزئی و تعیین تعداد و چگونگی نمونه برداری برای هر دسته از دیگر نتایج این پژوهش است.

کلیدواژه‌ها: نمایه سازی؛ نمونه برداری؛ کنترل کیفیت؛ پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)؛ سازماندهی و تحلیل اطلاعات

فهرست مطالب

عنوان	شماره صفحه
چکیده.....	أ.....
قدردانی.....	ذ.....
فصل اول- کلیات پژوهش.....	۱.....
۱-۱- مقدمه.....	۲.....
۲-۱- کیفیت اطلاعات.....	۵.....
۳-۱- فراداده.....	۸.....
۴-۱- اهداف طرح.....	۹.....
۵-۱- روش پژوهش.....	۹.....
۶-۱- بیان مساله/ موضوع پژوهش.....	۱۱.....
فصل دوم- پیشینه پژوهش.....	۱۵.....
۱-۲- اهمیت کیفیت فراداده.....	۱۶.....
۲-۲- اهمیت کنترل کیفیت آماری در فرایند نمایه سازی و ویرایش.....	۱۹.....
فصل سوم- ارزیابی داده/فراداده های ثبت شده در سامانه ثبت جدید.....	۲۴.....

۲۵	۱-۳- مقدمه
۲۶	۲-۳- معرفی اقلام اطلاعاتی در فرایند نمایه‌سازی و ویراستاری
۲۶	۱-۲-۳- اقلام اطلاعاتی فرایند نمایه‌سازی
۲۸	۱-۲-۳- اقلام اطلاعاتی فرایند ویراستاری
۲۸	۳-۳- چارچوب ارزیابی کیفیت مدارک ثبت شده در فرایند نمایه‌سازی
۳۰	۴-۳- چارچوب ارزیابی کیفیت مدارک ثبت شده در فرایند ویراستاری
۳۱	۵-۳- نتایج ارزیابی کیفیت مدارک ثبت شده در فرایند نمایه‌سازی
۳۱	۱-۵-۳- ارزیابی کیفیت اقلام صفحه عنوان
۳۳	۲-۵-۳- ارزیابی کیفیت چکیده‌های فارسی و لاتین
۳۵	۳-۵-۳- ارزیابی کیفیت فهرست‌ها
۳۶	۶-۳- نتایج ارزیابی کیفیت مدارک ثبت شده در فرایند ویراستاری
۳۷	۱-۶-۳- ارزیابی کیفیت اقلام صفحه عنوان
۳۹	۲-۶-۳- ارزیابی کیفیت چکیده‌های فارسی و لاتین
۴۰	۳-۶-۳- ارزیابی کیفیت فهرست‌ها
۴۲	۴-۶-۳- ارزیابی کیفیت اقلام اضافه شده در نمایه‌سازی
۴۴	۷-۳- جمع‌بندی و نتیجه‌گیری
۴۸	فصل چهارم- معرفی روش‌های نمونه‌برداری
۴۹	۱-۴- مقدمه
۵۱	۱-۱-۴- معرفی انواع روش‌های نمونه‌گیری برای پذیرش
۵۴	۲-۱-۴- مقایسه میان طرح‌های یک، دو و چندبار نمونه‌برداری
۵۵	۲-۴- بازرسی اصلاحی
۵۷	۳-۴- روش نمونه‌برداری بچ‌به‌بچ از دیدگاه مبانی احتمال
۵۹	۱-۳-۴- منحنی OC
۶۴	۴-۴- روش نمونه‌برداری پیوسته (CSP)
۶۷	۵-۴- بررسی ویژگی‌ها و تمایزهای روش‌های نمونه‌برداری
۶۹	فصل پنجم- طراحی روش‌های نمونه‌برداری برای فرایندهای نمایه‌سازی و ویراستاری
۷۰	۱-۵- مقدمه

۷۱	۵-۲- چارچوب کلان مراحل یافتن طرح نمونه‌برداری مناسب.....
۷۳	۵-۳- طراحی الگوریتم دستیابی به پارامترهای بهینه طرح نمونه‌برداری
۷۵	۵-۴- ارائه طرح نمونه‌برداری برای فرایندهای نمایه‌سازی و ویراستاری.....
۷۹	۵-۵- تعیین مسئول اصلی در فرایند کنترل اقلام پارسا
۸۳	۵-۶- منحنی OC برای طرح‌های نمونه‌برداری
۶۹	فصل ششم- نتیجه‌گیری و ارائه رهنمودهای کاربردی
۸۷	۶-۱- نتایج کلیدی پژوهش
۸۸	۶-۱- ارائه رهنمودهای کاربردی
۹۰	مراجع
۹۱	فهرست منابع فارسی
۹۱	فهرست منابع لاتین

فهرست جداول

عنوان	شماره صفحه
جدول ۱-۳ ناهمخوانی‌های بالقوه در فرایند نمایه‌سازی پارسا	۲۸
جدول ۲-۳ نتایج ارزیابی کیفیت صفحه عنوان پارساها در نمایه‌سازی	۳۲
جدول ۳-۳ نتایج ارزیابی کیفیت چکیده‌های فارسی و لاتین پارساها در نمایه‌سازی	۳۴
جدول ۴-۳ نتایج ارزیابی کیفیت فهرست‌های پارساها در نمایه‌سازی	۳۵
جدول ۵-۳ نتایج ارزیابی کیفیت صفحه عنوان پارساها در نمایه‌سازی	۳۷
جدول ۶-۳ نتایج ارزیابی اقلام چکیده‌های فارسی و لاتین پارساها در ویراستاری	۳۹
جدول ۷-۳ نتایج ارزیابی کیفیت فهرست‌های پارساها در ویراستاری	۴۱
جدول ۸-۳ نتایج ارزیابی کیفیت اقلام اضافه شده توسط نمایه‌ساز در ویراستاری	۴۳
جدول ۹-۳ بررسی میزان بهبود در اقلام اطلاعاتی مدارک پارسا	۴۵
جدول ۱-۵ مقیاس‌دهی به پارامترهای AQL و LTPD برپایه نوع مشخصه کیفی	۷۶
جدول ۲-۵ تعیین حدود قابل قبول برای تعداد نمونه	۷۷
جدول ۳-۵ طرح نمونه‌برداری بهینه برای مراحل گوناگون ثبت و سازماندهی اطلاعات	۷۸
جدول ۴-۵ شناسایی مسئول اصلی اقلام اطلاعاتی سامانه ثبت	۷۹

فهرست اشکال و نمودارها

عنوان	شماره صفحه
شکل ۱-۱ مراحل کنترل کیفیت در واحدهای سازمانی ایرانداک	۱۲
شکل ۱-۲ فرایند نمایه‌سازی و ویرایش پارسا	۲۲
شکل ۱-۳ نمودار مقایسه کیفیت اقلام صفحه عنوان پارساها در نمایه‌سازی	۳۳
شکل ۲-۳ نمودار مقایسه کیفیت اقلام چکیده‌های فارسی و لاتین پارساها در نمایه‌سازی	۳۴
شکل ۳-۳ نمودار مقایسه کیفیت اقلام فهرست‌های پارساها در نمایه‌سازی	۳۶
شکل ۴-۳ نمودار مقایسه کیفیت اقلام صفحه عنوان پارساها در ویراستاری	۳۸
شکل ۵-۳ نمودار مقایسه کیفیت اقلام اطلاعات چکیده پارساها در ویراستاری	۴۰
شکل ۶-۳ نمودار مقایسه کیفیت اقلام فهرست‌های پارساها در ویراستاری	۴۲
شکل ۷-۳ نمودار مقایسه کیفیت اقلام اضافه شده توسط نمایه‌ساز در ویراستاری	۴۳
شکل ۱-۴ نمای شماتیک از روش نمونه‌برداری یک مرحله‌ای	۵۹
شکل ۲-۴ نمای شماتیک از منحنی OC	۶۰
شکل ۳-۴ منحنی OC در حالت ایده‌آل	۶۱
شکل ۴-۴ جایگاه ریسکهای مصرف کننده و تولید کننده در منحنی OC	۶۳
شکل ۵-۴ مقایسه دو حالت ایده‌آل و واقعی در منحنی OC	۶۴
شکل ۱-۵ گامهای اصلی یافتن طرح نمونه‌برداری در فرایندهای نمایه‌سازی و ویراستاری	۷۲
شکل ۲-۵ الگوریتم تعیین پارامترهای بهینه طرح نمونه‌برداری	۷۴
شکل ۳-۵ منحنی OC برای طرح نمونه‌برداری مشخصه نوع A	۸۴
شکل ۴-۵ منحنی OC برای طرح نمونه‌برداری مشخصه نوع B	۸۴

شکل ۵-۵ منحنی OC برای طرح نمونه برداری مشخصه نوع C ۸۵

قدردانی

از ریاست گرامی پژوهشگاه که همواره پشتیبان اصلی پژوهش‌های کاربردی بوده‌اند و در این پژوهش نیز از آغاز بنده را حمایت کردند سپاسگزاری می‌کنم. همچنین، از همکار سخت‌کوش، دانا و بادرایت این پروژه، سرکار خانم دهسرای، که این پژوهش بدون کمک ایشان به سرانجام نمی‌رسید، کمال تشکر را دارم. از جناب آقای محمودیان که دیدگاه‌های اجرایی ایشان جنبه‌های کاربردی این پژوهش را پررنگ‌تر کرد نیز بسیار قدردانی می‌کنم. از همکاران گرامی مدیریت سازماندهی و تحلیل اطلاعات، به‌ویژه خانم‌ها صادقیان، گنجیان، موسیوند و محمودیان که در بخش مطالعه میدانی در سامانه ویرایش بنده را یاری رساندند، کمال تشکر را دارم. در نهایت، از اعضای گرامی شورای پژوهش ایرانداک برای ارائه دیدگاه‌های ارزشمندشان سپاسگزارم.

فصل اول

کلیات پژوهش

۱-۱- مقدمه

امروزه سازمان‌های گوناگون با رشد فزاینده پایگاه‌های داده و اطلاعات مواجه هستند. مدارک و مستندات به‌عنوان بخشی از این پایگاه‌ها با توجه به نقش ویژه‌ای که در بقا و توسعه دانش سازمانی دارند به‌عنوان یک پایگاه داده بیشتر مورد توجه هستند. سوابق و مستندات بی‌کیفیت، منسوخ و حتی کم‌اهمیت در صورتی که در قالب یک فرایند کنترل کیفیت غربال‌گری نشوند و سطح کیفیتشان ارتقاء پیدا نکند در آینده سازمان‌ها را با حجم فزاینده داده بی‌کیفیت مواجه خواهند ساخت که به شدت چابکی و پویایی آنها را با تهدید مواجه خواهند ساخت. پایگاه‌های اطلاعاتی که در دسترس عموم قرار می‌گیرند مانند گنجینه‌های پایان‌نامه/ رساله (پارسا)ها برپایه همین اصل و قاعده باید از جنبه‌های کیفیت، بهره‌وری و تعالی مورد توجه قرار گیرند تا جمعیت زیاد کاربران آنها دچار سردرگمی و گمراهی نشوند و در یک نگرش کلان، جامعه به رشد و تعالی برنامه‌ریزی شده دست یابد.

از سوی دیگر، در بهبود کیفیت مدارک و مستندات، اصل ۲۰-۸۰ که برای اولین بار جوران آن را توسعه داد بسیار تاثیرگذار است (Wilkinson, 2006). برپایه این اصل در یک جامعه هشتاد درصد از معلولها فقط از بیست درصد علتها ناشی می‌شوند. آمارشناسان از اصطلاح "اثر تمرکز"^۱ برای توصیف پدیده‌ای استفاده می‌کنند که در آن جمعیتی از اعضاء برای یک اثر مشترک مشارکت داشته باشند و تعداد کمی از مشارکت‌کنندگان بخش عمده اثر را تشکیل می‌دهند. در بازیابی مدارک و مستندات تعداد واژه‌های زیاد کتابشناختی بازیابی شده، کیفیت بازیابی را کاهش می‌دهد. به بیان دیگر،

¹ Concentration Effect

در مواردی که نتیجه یک جستجو توسط کاربر تعداد زیادی مدرک باشد نارضایتی را در پی خواهد داشت. اگر کیفیت مدارک بازیابی شده نیز نامناسب باشد سطح نارضایتی بالاتر خواهد بود. امروزه ابزارهایی که در فیلتر کردن مدارک بازیابی شده استفاده می شود تا حدی به افزایش کیفیت بازیابی مدارک جستجو شده توسط کاربر کمک می کند و نقش اثر تمرکز را نمایان تر می سازد. سامانه های پیشنهاددهنده نیز ابزارهای مفیدی در ارتقاء سطح کیفیت بازیابی اطلاعات هستند.

همچنین کیفیت محتوای مدارک و مستندات بعنوان یک بعد حیاتی در سطح کیفیت یک پایگاه داده از مدارک و مستندات شناخته شده و اهمیت بسزایی در رضایت کاربران پایگاه دارد. این مقوله طیف گسترده ای را از ویژگی های ساختاری یک مدرک علمی مانند جدول، نمودار، فهرست گرفته تا جنبه های درونی آن مانند اصالت، سطح پابندی به اصول اخلاقی، منسوخ نبودن مطالب علمی استفاده شده پوشش می دهد. از سوی دیگر، رشد کمی داده ها، اهمیت بسزای بازیابی مدارک و همچنین کیفیت محتوای یک مدرک علمی، سازمان ها را به سمت توجه به کنترل کیفیت مستندات سوق می دهند. همانطور که کایزر (۱۹۹۰) می گوید "مدیر خوب بودن به معنای پذیرفتن مسئولیت کارهایی است که اداره شما تولید می کند" (در موضوع مورد اشاره ما این اطلاعات است)، مدیران پایگاه های اطلاعاتی و دانشی نیز در این زمینه مسئولیت بسیار مهمی برعهده دارند. همچنین، مشتریان و کاربران نهایی جستجوی بانک های اطلاعاتی از موضوعات مربوط به کیفیت آگاهی پیدا کرده و خواستار اطمینان، دقت و به طور کلی کیفیت بالاتری متناسب با هزینه و زمانی که در استفاده از آنها ایجاد شده است هستند. همراه با افزایش تقاضا برای کیفیت بالاتر، بهم پیوستگی و اعتماد بیشتر بین مشتریان و تأمین کنندگان محتوای پایگاه های اطلاعاتی موضوعی اجتناب ناپذیر است.

تنوپیر^۲ (۱۹۹۲) در مقاله خود با نام "یک روز در زندگی تولیدکننده بانک اطلاعاتی" تجربه اش را در بازدید و مشاهده همه مراحل ایجاد و اشاعه در سه شرکت بزرگ تولید بانک اطلاعاتی ارائه می کند. وی این مراحل را که از تأمین کننده مدرک آغاز شده و به مشتری (کاربر) مدرک ختم می شود در سه رابطه خلاصه می کند. اولین رابطه میان بین کسانی است که تعیین می کنند چه عناوین (قلم های اطلاعاتی) در بانک اطلاعاتی گنجانده شود و افرادی که باید آنها را بدست آورند. رابطه بعدی بین فراهم آوران و ناشران وجود دارد و به دنبال آن یک رابطه بین ناشران و بخش دریافت کننده برقرار

² Tenopir

می شود. رابطه سوم (نهایی) میان مشتریان بخش دریافت کننده (نمایه سازها) و دریافت کنندگان است. نمایه سازها، کاربران، ویراستاران و فهرست نویسانی هستند که فهرستی از عناوین موضوع، کلیدواژگان و نمایه ها را تهیه می کنند. همانطور که تنویر روند را شرح می دهد، بسیاری از بررسی ها و تأییدها ادامه می یابد. تأییدکننده ها، چه انسان یا ماشین، دریافت کنندگان خروجی نمایه سازی هستند. تنویر گروهی از متخصصان کنترل کیفیت را که نظارت بر کیفیت کلی مراحل مختلف دارند و خطاها را تصحیح می کنند، شناسایی می کند و نقش آنها را در کیفیت نهایی بسیار تاثیرگذار می داند. از این رو، به نظر می رسد روند کنترل کیفیت برای تولیدکنندگان پایگاه داده به بازرسی بسیار وابسته است. برخی بازرسی ها برپایه ماشینی سازی یا کنترل خودکار به کمک نرم افزار است مانند چک کردن کامل بودن ارقام داده، خطاهای تایپی و فرمت داده ها و همچنین دوبارگی ها. برخی دیگر به شکل کامل به صورت چشمی و برپایه کنترل انسانی صورت می گیرد مانند کنترلی که کارشناسان نمایه ساز بر روی نمایه ها و همچنین اطلاعات کتابشناختی صورت می دهند.

از دیگر سو، متخصصان کنترل کیفیت در نهایی کردن مدرک پیش از اشاعه و همچنین تصحیح خطاهای معرفی شده توسط سایر کاربران و کارشناسان نقش ویژه ای برعهده دارند. بنابراین، تعیین و طراحی فرایندهایی برای نظارت بر کیفیت ورودی داده ها و درست کردن پیوسته خطاها کاری ضروری است، حتی اگر این فرایندها همیشه به بهبود کیفیت نیانجامند. رویکردهای تضمین کیفیت نیز می توانند همزمان با طراحی مناسب فرایندهای کنترل کیفیت مستندات، استفاده از نرم افزارها و الگوریتم های از پیش طراحی شده خطایاب و نیز ممیزی های درست، یک رویکرد بهبود مستمر را در سازمان نهایی کنند. شاخص هایی مانند کامل بودن، دقت، صحت و سازگاری می توانند در تضمین کیفیت مستندات مورد توجه قرار گیرند که در یک فرایند ارزیابی عملکرد پایگاه داده به شکل مستمر اندازه گیری خواهند شد.

در پیاده سازی تضمین کیفیت پایگاه های داده استفاده از رویکردهای پیشگیری از بروز خطا در همه فرایندها باید بطور ویژه ای مورد توجه قرار گیرد. در طراحی سامانه ها و نرم افزارها نیز رویکردهای مشتری مدارانه و توجه به خواسته های مشتری به عنوان یک امر کلیدی در انتخاب ارقام اطلاعاتی و کارکردهای درست سامانه یا نرم افزار به حساب می آید. این موضوع از آن جهت اهمیت دارد که

مشتریان و کاربران نهایی استفاده‌کننده از پایگاه داده به‌عنوان ذی‌نفع اصلی استفاده‌کننده از پایگاه داده، شناخته شده و خطاهای پایگاه داده بیشترین تاثیر را در رضایت آنها خواهد داشت.

برای شناسایی و برطرف کردن خطاها همراستا با پیاده‌سازی مفاهیم تضمین کیفیت در فرایند کنترل کیفیت داده موضوع نمونه‌برداری مورد توجه است تا در کنار مفاهیم تضمین کیفیت اطمینان کامل‌تری از کیفیت خروجی فرایند فراهم شود. براین اساس در پژوهش جاری به چگونگی انجام این کار در فرایند نمایه‌سازی ویراستاری مدارک پارسا خواهیم پرداخت. از آنجاکه مفهوم نمونه‌برداری با مفهوم کیفیت آمیخته است و زمینه پژوهش در حوزه داده و اطلاعات است در ادامه به مرور مقدماتی در مورد کیفیت اطلاعات خواهیم پرداخت.

۲-۱- کیفیت اطلاعات

اطلاعات در کوتاه‌ترین تعریف، «داده‌های پردازش شده» است. داده‌های خام بالقوه محتوای معنی‌داری هستند که ما آنها را در راستای شناخت و فهم هر مفهوم مادی یا غیرمادی، به کمک روش‌های پژوهشی، و با استفاده از ابزارهای شناختی به دست می‌آوریم. داده یک شرح مقدماتی از یک پدیده، اتفاق، فعالیت یا تعاملات است که ثبت، دسته‌بندی شده و ذخیره شده‌است؛ اما سازماندهی نشده و برای یک منظور مشخص آماده نشده است. داده‌ها عناصر اصلی اطلاعات هستند. داده‌ها در صورتی به اطلاعات تبدیل می‌شوند که افراد بخواهند برای درک بیشتر از آنها استفاده کنند (پرایس و شانکس^۳، ۲۰۱۶). اطلاعات، داده‌های خلاصه‌ای هستند که گروه‌بندی، ذخیره، پالایش و سازماندهی شده‌اند تا بتوانند معنی‌دار شوند. اطلاعات زمانی ارزش پیدا می‌کنند که برای یک بُعد خاص، یک فرد خاص، یک هدف خاص و در زمان خاص گردآوری و آماده شوند، لذا اطلاعاتی که برای یک مدیر، جنبه اطلاعاتی دارد، برای مدیر دیگر ممکن است اصلاً ارزشی نداشته باشد. بدین ترتیب، اطلاعات آگاهی‌های به دست آمده از عناصرها و رویدادهای جهان هستی است. به زبان محدود تکنیکی، مجموعه‌ای از نمادهای زبانی معنی‌دار و پیوسته درباره موجودات است. در زبان انگلیسی، اطلاعات از نظم ساختاری و ذاتی خبر می‌دهد (باتینی و اسکناپیکو^۴، ۲۰۱۶).

³ Price & Shanks

⁴ Batini & Scannapieco

اطلاعات را از دو دیدگاه ساختاری و معنایی تقسیم‌بندی می‌شوند به‌صورتی که هر کدام از آنها خود گروه‌های دیگری را دربر می‌گیرد (اثنی عشری و اسدی، ۱۳۹۴).

الف. جنبه ساختاری: ساختار اطلاعات، مستقل از جنبه‌های معنایی آن است و ممکن است ایستا یا پویا ارزیابی شود. زمانی که اطلاعات تغییرناپذیر باشد از نوع ایستاست؛ مانند اطلاعات کتاب، تصویر، صفحه فشرده. هرگاه اطلاعات در جریان فرایند، دستخوش هرگونه تغییر یا پردازش شود از نوع پویاست. نوع اول را فرآورده و نوع دوم را فرایند می‌خوانند. (اثنی عشری و اسدی، ۱۳۹۴)

الف-۱ اطلاعات به‌صورت فرآورده: زمانی که اطلاعات را کالا تلقی کنیم، مانند هر فرآورده‌ای قابلیت تولید، ذخیره‌سازی، و انتقال دارد و همچنین می‌تواند آفریده شود، رشد یابد، و نیست شود. این نوع اطلاعات مانند کتاب، تصویر، صوت، قابلیت نگهداری دارد. اطلاعات از این دیدگاه، گرچه در دسته "فرآورده" قرار می‌گیرد، می‌تواند برپایه هدف تحلیل‌گر در هر "فرایندی" شکل یابد. (کاردناس-گارسیا و همکاران^۵، ۲۰۱۹)

الف-۲ اطلاعات در قالب یک فرایند: این جنبه از اطلاعات دارای ویژگی پیچیده و متفاوت از جنبه پیشین است، فرایند محور بودن اطلاعات را از دو دیدگاه زیر می‌توان مورد توجه قرار داد.

۱- اطلاعات برپایه زمان و مکان در حال تغییر است؛ پدیده‌ای ممکن است در یک زمان یا مکان اطلاعات باشد و همان پدیده در زمان یا مکانی دیگر ارزش اطلاعاتی چندانی نداشته باشد. این نوع نگرش جنبه معناشناختی اطلاعات است.

۲- اطلاعات در ذات خود جنبه فرایندی را داراست (این دیدگاه با رویکرد قبلی متفاوت است و تاحدی در مقابل آن است). در این دیدگاه اطلاعات چیزی جز فرایند نیست؛ و بنابراین، ماهیت پویا دارد. آنچه از جنبه ساختاری ماهیت ایستا دارد در دو سوی این فرایند قرار می‌گیرد. از یک سو داده (سوی آغازین) و سوی دیگر آن دانش است. اطلاعات برخلاف داده یا دانش تغییرپذیر نیست، بلکه داده‌ها که همان مواد خام اطلاعات هستند، و دانش به‌عنوان صورت سازماندار و هدفمند قابلیت تغییر دارند (اثنی عشری و اسدی، ۱۳۹۴).

⁵ Cárdenas-García et al.

مارچاند^۶ (۱۹۹۰) هشت بعد کیفیت اطلاعات را به عنوان چارچوبی برای تجزیه و تحلیل به صورت زیر مشخص می کند.

۱. ارزش واقعی یک محصول یا خدمات اطلاعاتی که ممکن است برای یک کاربر اطلاعات داشته باشد.

۲. ویژگی های اطلاعات: ابعادی مانند صحت یا کامل بودن در این بخش گنجانده می شوند.

۳. قابلیت اطمینان محصول یا خدمات اطلاعات

۴. زمان مند بودن اطلاعات: معنی اطلاعات در چرخه حیات خود باید مورد تفسیر قرار گیرد.

۵. ارتباط (مرتبط بودن) اطلاعات: درجه و میزان همخوانی اطلاعات با استانداردها یا معیارهای کاربر. برای طراح سیستم های اطلاعاتی، ارتباط ممکن است این همخوانی هماهنگی اطلاعات با ویژگی های سامانه باشد در حالی که برای کاربر این ممکن است میزان فایده و استفاده در زمان مناسب باشد.

۶. اعتبار اطلاعات: روایی چگونگی جمع آوری یا تجزیه و تحلیل اطلاعات، شهرت کسی که آن را جمع آوری می کند، و یا چگونگی ارائه نتایج میزان اعتبار اطلاعات را مشخص می کند.

۷. زیبایی شناسی اطلاعات. ویژگی های ذهنی کاربران اطلاعات است که به چگونگی ارائه و اشاعه اطلاعات وابسته است (Lau et al., 2007).

۸. ارزش درک شده از اطلاعات.

برپایه این ۸ بعد کیفیت اطلاعات، امروزه نگرانی از کیفیت و قابلیت اعتماد اطلاعات، نه تنها در بین جستجوگران آنلاین و متخصصان اطلاعات، بلکه در بین عموم مردم رو به افزایش است (راسل راز و همکاران^۷، ۲۰۱۸). متخصصان اطلاعات برپایه این اصل به مشتریان در فهم محدودیت های منابع داده کمک می کنند و به پرورش تفکر نقد به داده ها می پردازند، و همچنین دقت و اعتبار اطلاعات جمع آوری شده را ارزیابی می کنند. از این رو، امروزه بهبود کیفیت به روشنی موضوعی مهم و حیاتی بوده و لازم است به عنوان یک راهبردی کلیدی در سازمان ها مورد توجه قرار گیرد. در تمام زنجیره

^۶ Marchand

^۷ Russell-Rose

تولید داده از تولیدکننده پایگاه داده گرفته تا مصرف‌کننده نهایی آن ایجاد همکاری و یک تعهد همه‌جانبه در سیستم برای بهبود پیوسته کیفیت ضروری و اجتناب‌ناپذیر است. از میان مراحل مختلف تولید، کنترل و اشاعه اطلاعات در پایگاه‌های اطلاعاتی علمی، نمایه‌سازی و کنترل (ویراستاری) اطلاعات نقش ویژه‌ای در پردازش اطلاعات برعهده دارند که در این پژوهش به این بخش خواهیم پرداخت. همچنین باتوجه به این که در فرایند نمایه‌سازی و ویراستاری فراداده به عنوان مهم‌ترین بخش مورد توجه است، در بخش بعدی به تعریف عمیق‌تری از آن خواهیم پرداخت.

۳-۱- فراداده

فراداده داده‌نما، متادیتا، ابرداده - شریان حیاتی هر سامانه کتابخانه‌ای است و در دنیایی که فقط در انگلستان سالانه ۴۵۰۰۰ کتاب جدید تولید می‌شود تنها ابزاری است که جستجوی دانشگاهی را ممکن می‌سازد (راویت و دانا^۸، ۲۰۱۵). استاندارد مارک^۹ باوجود محدودیت‌هایی که به‌ویژه در حوزه کیفیت فراداده دارد امروزه به‌عنوان یک استاندارد پایه برای فراداده‌های کتابخانه‌های دانشگاهی مورد استفاده قرار می‌گیرد (راویت و دانا، ۲۰۱۵). در این موقعیت برای دستیابی به یک کیفیت کارا در فراداده‌ها چاره‌ای جز پایبندی به این استاندارد و به‌صورت همزمان، یافتن راهکارهای فراتر رفتن از آن وجود ندارد. از سوی دیگر، کیفیت فراداده در همه داده‌های علمی دیجیتالی مانند کتاب‌های الکترونیکی، پایان‌نامه/رساله‌ها یا سایر مدارکی که شکل فیزیکی/قفسه‌ای ندارند بسیار اهمیت دارد. گرچه خدمات جستجو برپایه وب (اینترنت) امروزه به‌راحتی دردسترس است اما تا زمانی که فراداده کیفیت مطلوب را نداشته باشد یافتن مدرک مورد نظر کاری بسیار دشوار است. برخلاف گذشته که مسئولیت معرفی یک مدرک علمی با کتابخانه یا واحد نشر آن مدرک بود امروزه اطلاعات کتابشناختی یک مدرک که در فراداده آن نهفته است این کار را برعهده دارد از این‌رو جستجوگر مدرک با مشاهده آن به مناسب‌بودن، سازگاری با هدف جستجو و حتی اعتبار آن مدرک پی می‌برد. باتوجه به حجم زیاد داده در بانک‌های اطلاعاتی برای اطمینان از کیفیت داده‌ها نمونه‌برداری و کنترل کیفیت نمونه‌ها و همچنین قضاوت کیفیت داده برپایه کیفیت نمونه برداشته شده کاری اجتناب‌ناپذیر

^۸ Ravit & Dana

^۹ Mark

است. گرچه برخی کنترل‌های خودکار بر روی میزان کامل بودن و یا سازگاری محتوای داده/ فراداده‌های بانک‌های اطلاعاتی پژوهشی وجود دارد، دستی بودن برخی گام‌های فرایند به‌ویژه در کار نمایه‌سازی/ ویراستاری داده‌ها اهمیت کنترل کیفیت فراداده را با توجه به اثرگذاری بر کیفیت پژوهش در جامعه علمی بیشتر نمایان می‌کند.

در بخش بعدی به معرفی اهداف طرح خواهیم پرداخت.

۴-۱- اهداف طرح

اهداف پژوهش:

- ۱) شناسایی وضعیت موجود کیفیت اقلام اطلاعاتی وارد شده به سامانه ویرایش و سازماندهی اطلاعات از سامانه ثبت جدید،
- ۲) تعیین تعداد نمونه و روش تصمیم‌گیری در فرایند نمایه‌سازی و ویراستاری براساس تعداد کل مدارک اختصاص داده شده به هر کاربر،
- ۳) تعیین روش تصمیم‌گیری در صورت مشاهده داده‌های ناهمخوان در سامانه ویرایش و سازماندهی اطلاعات توسط کاربر نمایه‌ساز و ویراستار

۵-۱- روش پژوهش

این پژوهش در دو گام اصلی انجام خواهد شد:

گام اول) شناخت وضعیت موجود سطح کیفیت اقلام اطلاعاتی پس از طی مراحل ثبت (خروجی فرایند ثبت و فراهم‌آوری)

برای انجام این گام اقدامات زیر صورت خواهد پذیرفت:

- ۱-۱) نمونه‌برداری از داده/ فراداده‌های ثبت شده در سامانه ثبت جدید پس از فرایند ثبت و فراهم‌آوری،

ارائه طرح نمونه برداری پذیرش در فرایند ویراستاری □ ۱۰

- لازم به توضیح است طراحی فرمها و جداول و چارچوب نمونه برداری در این مرحله توسط مجری پروژه و اجرای نمونه برداری و تعیین مناسب/نامناسب بودن اقلام داده ها توسط همکاران معاونت اطلاعات علم و فناوری انجام خواهد شد.
- تعداد داده مناسب در این مرحله پس از تأیید معاونت گرامی اطلاعات علم و فناوری تعیین خواهد شد.

۲-۱) تحلیل نتایج نمونه برداری و تعیین درصد داده/فرا داده ناهمخوان پارساها

گام دوم) تعیین طرح نمونه برداری و مسئول اصلی کیفیت (کارشناس واحد فراهم آوری/نماینده ساز/ ویراستار) برای هریک از اقلام اطلاعاتی

در این گام دو روش اصلی نمونه برداری در کنترل کیفیت شامل نمونه برداری به صورت بچ به بچ و نمونه برداری به صورت پیوسته برپایه منابع علمی معرفی و ارزیابی خواهند شد. مزایا و معایب این دو روش مورد بررسی قرار خواهد گرفت. سپس تحلیل داده های جمع آوری شده در مرحله قبل صورت خواهد پذیرفت. همچنین، کارگروه کیفیت پارساها تشکیل خواهد شد. در این کارگروه برپایه نتایج تحلیل، تعداد نمونه (n) و عدد پذیرش (C) برای هر قلم اطلاعاتی و نیز مسئول اصلی کیفیت (کارشناس واحد فراهم آوری/نماینده ساز/ ویراستار) تعیین خواهد شد. برای تعیین عدد پذیرش (C) از شاخص میانگین در تحلیل های آماری انجام شده، استفاده خواهد شد.

لازم به توضیح است در گام دوم طرح نمونه برداری، برای نماینده ساز، ویراستار و همچنین مدیر گرامی سازماندهی و تحلیل اطلاعات جداگانه تعیین خواهد شد.

پیشنهاد ارائه شده توسط مجری پروژه در گام دوم توسط کارگروه کیفیت پارساها نهایی و پیشنهادات اصلاحی ارائه خواهد شد.

۶-۱- بیان مساله / موضوع پژوهش

واحد «سازماندهی و تحلیل اطلاعات» بعنوان یکی از مهمترین بخش های سازمانی در معاونت اطلاعات علم و فناوری به حساب می آید. ماموریت اصلی این واحد تحقیق و بررسی روش های سازماندهی و طبقه بندی مدارک و اطلاعات، سازماندهی مدارک به منظور پردازش، تحلیل و بازیابی اطلاعات و نیز انتخاب یکی از روش های نمایه سازی و کنترل واژگان با اصطلاحنامه ها است. مهمترین وظایف واحد سازماندهی و تحلیل اطلاعات بشرح زیر است.

- بررسی انواع مدارک پارسا و نمایه سازی آنها که از بخش تهیه مدارک ارسال می گردد.
- استفاده از اصطلاحنامه ها جهت انتخاب اصطلاحات مرجع مرجع و همچنین یکسان سازی و معادل یابی کلیدواژه ها به فارسی و انگلیسی.
- تلاش در جهت استانداردسازی نمایه سازی و استفاده از دستورالعمل های نمایه سازی خدمات مدیریت ورود سازماندهی و تحلیل اطلاعات:
- گردآوری، سازماندهی، ذخیره، حفظ، بازیابی، تحلیل و اشاعه مدارک و اطلاعات علمی و فنی کشور از جمله پایان نامه های کارشناسی ارشد و رساله های دکترای تخصصی دانش آموختگان داخل و دانش آموختگان ایرانی خارج از کشور، مقاله های 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، آمارنامه ها، فرهنگ نامه ها، و اصطلاح نامه ها.

همانگونه که از ماموریتها و وظایف اشاره شده مشاهده می شود و برپایه ماموریت اصلی پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک) یکی از مهمترین حوزه های کاری واحد سازماندهی و تحلیل اطلاعات، نمایه سازی و ویرایش پایان نامه / رساله (پارسا) های دانش آموختگان کل کشور است.

پارساها پس از ثبت توسط دانشجو و تأیید در واحد فراهم آوری برای نمایه سازی و ویرایش به واحد سازماندهی و تحلیل اطلاعات ارسال می شود. لازم به توضیح است که این کار به صورت سیستمی انجام شده و فعالیتهای اصلی در واحد سازماندهی و تحلیل اطلاعات در قالب سامانه ویرایش و تحلیل اطلاعات علمی صورت می پذیرد.



شکل ۱-۱ مراحل کنترل کیفیت در واحدهای سازمانی ایرانداک

بطور کلی بمنظور اطمینان از کیفیت نهایی پارسای اشاعه داده شده در سامانه گنج همانگونه که در شکل ۱-۱ نمایش داده شده است، ۴ مرحله کنترل کیفیت صورت می گیرد.

مرحله اول به واحد ثبت و فراهم آوری اطلاعات مربوط می شود که پس از ثبت مدارک توسط دانشجویان تحصیلات تکمیلی کنترل های اولیه داده (فایل پی دی اف/ ورد) و فراداده (اقلام اطلاعاتی ثبت شده در سامانه مانند نام، نام خانوادگی، رشته، گرایش و ...) پارساها کنترل می شود و در صورت تأیید نهایی این واحد و سپس نماینده دانشگاه به سامانه ویرایش و سازماندهی اطلاعات (واحد سازماندهی و تحلیل اطلاعات) وارد می شود. (مرحله دوم تا چهارم)

مرحله دوم شامل کنترل‌های بخش نمایه‌سازی است. هرچند توجه اصلی این بخش اصلاح نگارش نام‌ها، عنوان (فارسی و انگلیسی)، چکیده (فارسی و انگلیسی)، کلیدواژه «در صورت وجود» و تخصیص کلیدواژه‌های جدید» از منابع مختلف اصطلاح‌نامه‌ها، کلیدواژه‌های کمکی، و کتاب‌های منبع در اختیار است، اما در عمل همه اقلام اطلاعاتی پارسا دوباره چک شده و اصلاحات لازم صورت می‌پذیرد. به عبارتی در این مرحله بازرسی ۱۰۰ درصد بر روی تمام اقلام اطلاعاتی پارسا صورت می‌پذیرد.

در مرحله سوم کارشناسان بخش ویراستاری در واحد سازماندهی و تحلیل اطلاعات دوباره عنوان و کلیدواژه را چک کرده و در صورت بروز مشکل ممکن است آنرا اصلاح و یا به نمایه‌ساز عودت دهند. کنترل صورت گرفته در این مرحله هم ۱۰۰ درصد بوده بعبارتی تمامی مدارک نمایه‌سازی شده باید توسط کارشناسان ویراستار کنترل و اصلاح شود.

در مرحله چهارم، مدیر واحد نیز بصورت تصادفی برخی مدارک را کنترل و در صورت مشاهده مشکل آنها را اصلاح می‌کند.

لازم به یادآوری است مدارک پارسا پس از مرحله سوم اشاعه داده می‌شوند و کنترل‌های مرحله آخر پس از اشاعه انجام می‌شود.

باوجود کنترل‌های مختلفی که در ۳ مرحله گفته شده صورت گرفته است در سامانه گنج مدارک اشاعه داده شده با مشکلات کیفی مختلفی همراه است که البته بخشی از آنها را میتوان با کنترل کارا و اثربخش در ۳ مرحله گفته شده حذف نمود. مشکلاتی مانند غلط‌های املائی چکیده، عنوان و ...، نواقص اقلام اطلاعاتی و ناقص بودن نمایه‌سازی‌های صورت گرفته مواردی هستند که با بهبود فرایند کنترل کیفیت در ۳ مرحله گفته شده در شکل ۲ می‌توان تا حدود زیادی آنها را رفع کرد.

از سوی دیگر در سامانه ثبت جدید برخی از اقلام اطلاعاتی مانند رشته، گرایش، دانشگاه و ... بصورت انتخاب از لیست صورت می‌گیرد و خطاناپذیر خواهد بود لذا نیازی به کنترل در مراحل ۲ تا ۴ ندارد. لازم به یادآوری است از آنجا که سامانه ثبت جدید با هدف حل مشکلات کیفی طراحی و پیاده‌سازی شده است نیاز است رویه‌های فعلی در سامانه ویرایش برپایه اصلاحات و بهبودهای سامانه

ثبت جدید اصلاح شود و تمرکز کنترل‌های کیفیت در واحد سازماندهی و ویرایش اطلاعات بر مشکلات کیفی مدارک پارسای خروجی سامانه ثبت جدید باشد. از سوی دیگر ۲ مرحله کنترل در این واحد (که هر ۲ بازرسی ۱۰۰٪ است) کارایی لازم را ندارد و نیاز است در برخی اقلام اطلاعاتی کنترلها کاسته شده تا اثربخشی مناسب ایجاد شود. همچنین در وضعیت موجود رویه مشخصی برای ارزیابی کیفیت نمایه‌سازی و ویرایش وجود ندارد و مشخص نیست عملکرد یک کارشناس نمایه‌ساز/ ویراستار با چه کیفیتی مطلوب و در چه صورت نامطلوب است.

از این‌رو، در این پژوهش بنا داریم چارچوبی برای نمونه‌برداری در فرایندهای نمایه‌سازی و ویراستاری فراهم کنیم. به عبارتی تعیین این‌که از مجموع مدارک نهایی ارسال شده برای هر کاربر نمایه‌ساز کدام قلم اطلاعاتی نمونه‌برداری و کنترل شود و کدام قلم بدون نیاز به کنترل برای ویراستاری ارسال شود و در حین نمونه‌برداری، از حجم ورودی برای کاربر چه تعداد نمونه برداشته شده و در صورت مشاهده مشکل چگونه تصمیم‌گیری شود نیز یکی از اهداف این پژوهش است. اقدامی مشابه آنچه بیان شد برای فرایند ویراستاری نیز صورت خواهد گرفت.

فصل دوم

پیشینه پژوهش

۲-۱ اهمیت کیفیت قرارداد

امروزه همانگونه که کتابخانه‌ها در حال بررسی ارتقاء نقش خود بعنوان نگهداری اطلاعات به یک نقش کامل‌تر در حوزه انتشار داده‌های دیجیتالی هستند فهرست‌های آن‌ها و شیوه فهرست‌بندی آن‌ها نیز با تغییرات پایه‌ای مواجه شده است. نه تنها کتابخانه‌ها باید نقش سازمان دهی و ارائه دهی کتاب‌ها و نشریات منتشر شده توسط دیگران را بر عهده دارد و نقش سنتی خود را ایفا کنند، بلکه باید به طور جدی شروع به ایفای نقش ناشر و جمع کننده مطالب دیجیتالی نیز ایفا کنند. همچنین، ارزیابی قرارداد هنوز بخش روتین از کار بیشتر متخصصان قرارداد نیست چراکه متخصصان قرارداد مبتنی بر کتابخانه وظیفه خود را در وهله اول عنوان ایجاد قرارداد می‌بینند تا کنترل کیفیت آن. بروس و هیلمن^۱ (۲۰۰۴) در پژوهش خود در مورد کیفیت قرارداد، هفت معیار و چندین سطح پابندی را تعریف کردند. کامل بودن، دقت، اصالت، همخوانی با خواسته‌ها، سازگاری منطقی و انسجام، به موقع بودن و قابلیت دسترسی این هفت معیار را تشکیل می‌دهند. آن‌ها این نکته را بیان کردند که تفاوت در قرارداد ممکن است از محیطی که در آن قرارداد ساخته و مورد استفاده قرار می‌گیرد یا تفاوت‌های ذاتی موجود در آنها ناشی شده باشد. آن‌ها همچنین به این نکته اشاره می‌کنند که کیفیت را نمی‌توان در شرایط خلا مورد بحث قرار داد بلکه محدودیت‌های اقتصادی، سیاسی و فنی بخشی از هر تصمیم تاثیرگذار بر کیفیت و درک کیفیت هستند.

بارتون و همکاران (۲۰۰۳) در مورد اهمیت بافت^۲ داده در کیفیت قرارداد پژوهشی را انجام دادند. در سطح محلی، زمینه ای که قرارداد در آن ایجاد می‌شود می‌تواند بر نتیجه برنامه تضمین کیفیت تأثیر بگذارد، به ویژه که مربوط به زمینه های خاص باشد. در بعضی موارد، هرچه مجموعه داده بزرگتر باشد، احتمال بروز مشکل بیشتر می‌شود. به عنوان مثال، در جمعیت زیادی از نویسندگان، در پرونده‌های مرجع، نام ممکن است مورد نیاز باشد تا جان اسمیت ها را از دیگری جدا کند. در موارد دیگر، درجه تنوع می‌تواند مشخص کند که آیا کیفیت قرارداد به یک مساله تبدیل می‌شود یا نه. برای مثال، در آرشیوی از مقالات و گزارش‌هایی که از یک گروه تحقیق منفرد نشأت می‌گیرد، وابستگی نویسنده را می‌توان به

¹ Bruce & Hillmann

² Context

عنوان یک مقدار پیش فرض تنظیم کرد، به عنوان مثال، در جمعیت زیادی از نویسندگان، در پرونده‌های مرجع، نام ممکن است مورد نیاز باشد تا جان اسمیت ها را از دیگری جدا کند. در موارد دیگر، میزان تنوع می تواند مشخص کند که آیا کیفیت قرارداد به یک مساله تبدیل می شود یا نه. برای مثال، در آرشیوی از مقالات و گزارش هایی که از یک گروه تحقیق منحصر به فرد نشأت می گیرد، وابستگی نویسنده را می توان به عنوان یک مقدار پیش فرض تنظیم کرد، در حالی که یک آرشیو مبتنی بر موضوع ممکن است نیاز داشته باشد که حتی نام شرکت یا سازمانی که مقاله را تولید کرده است به عنوان یک قلم اطلاعاتی ثبت کرد.

تجزیه و تحلیل قرارداد باید بتواند به تعدادی پرسش مهم پاسخ دهد (تئاتر^۱، ۲۰۰۴)، به عنوان مثال:

- برای کدام قسمت های قرارداد داده وجود دارد؟
- چند درصد از تعداد کل سوابق هر فیلد دارای داده است؟
- قرارداد در زمینه های مختلف بافت خود چقدر سازگار است؟
- چه الگوهای قابل تشخیصی برای قرارداد وجود دارد؟

همه برنامه های بهبوددهنده کیفیت قرارداد بر پایه چرخه دمنگ و در چارچوب پرسش های بالا می توانند در بلند مدت به تضمین کیفیت ساختارهای توسعه دهنده قرارداد بیانجامند.

در حالی که بدیهی است که قراردادها، پایه و اساس هر خدمت برخط هستند، ایجاد و تولید سوابق قراردادها در یک سیستم به صورت کامل و سازگار کار آسانی نیست (راویت و دانا، ۲۰۱۵). طراحی فرایند کنترل کیفیت می تواند مشکلات ناهماهنگی و یکپارچگی قراردادها را در ارائه دهندگان خدمات جستجو حل کند. اهمیت این موضوع در حالی دوچندان می شود که محتوا به محض اینکه به صورت برخط قرار داده می شود در دسترس خوانندگان قرار می گیرد و نادرستی، عدم دسترسی و یا مشکلات دیگر تبعات سنگینی برای خوانندگان برجا خواهد گذاشت.

¹ Tennant

فرا داده با کیفیت شرط لازم برای رسیدن به فعالیت های کارا، قابل اعتماد و کارآمد در یک مجموعه دیجیتال است؛ همچنین، قدرت فرا داده در اجرای اثربخش نقش های اصلی اطلاعات کتابشناختی از جمله بازیابی، استفاده، انتشار و تصدیق است (پارک، ۲۰۰۹).

متأسفانه پژوهش های کمی در مورد فرا داده ها و کیفیت آنها در محیط دیجیتال صورت پذیرفته است و حتی در مواردی که نیاز به کیفیت فرا داده توسط پژوهشگر پیشنهاد شده است، راهکارهای اجرایی و خطوط راهنمای مناسبی ارائه نشده است (راویت و دانا، ۲۰۱۵). در حقیقت، یک روش استاندارد پذیرفته شده برای تعریف کیفیت فرا داده و روش هایی که بتوان آن را ارزیابی، اندازه گیری و بهبود داد، وجود ندارد (راویت و دانا، ۲۰۱۵). با این حال، چندین تعریف مفید که چارچوب عمومی را برای ارزیابی کیفیت فرا داده ارائه می دهند، وجود دارد. یکی از این تعاریف نتیجه پژوهش بروس و هیلمن (۲۰۰۴) است که ابعاد مختلف کیفیت فرا داده ها را بررسی کرده اند. آنها تعریفی از کیفیت فرا داده و ارزیابی آن به عنوان یک چارچوب متشکل از هفت بعد کامل بودن، صحت، دقت، اعتبار، سازگاری با نیازها، زمان مند بودن و دسترس پذیری ارائه کردند. استویلا و همکاران (۲۰۰۷) از یک مدل عمومی تر استفاده کرده و ارزیابی های کیفیت را در زمینه (بافت) استفاده از اطلاعات، باتکیه بر کاربردپذیری اطلاعات، ارائه می کنند. آنها در فرایند ارزیابی انواع فعالیت های کاربران استفاده کننده از اطلاعات و هنجارها و ارزش های جامعه کاربر اطلاعات مورد قضاوت قرار می دهند. استویلیا و گاسر (۲۰۰۸) ترکیبی از رویکردهای تحلیلی و تجربی ارائه می کنند که ارزش ایجاد شده از تغییر کیفیت فرا داده بعنوان پایه و خط مبنا از دیدگاه کاربران نهایی شفاف باشد.

به طور مشخص، پارک و توساکا (۲۰۱۰) دریافتند که رایج ترین معیارهای ارزیابی کیفیت فرا داده در از دیدگاه کارشناسان این زمینه، صحت و سازگاری بوده و هر دو نسبت به شاخص کامل بودن اطلاعات کتابشناختی اولویت بالاتری داشتند. پالویتسینیس و همکاران (۲۰۱۴) سازوکارهای تضمین کیفیت و تاثیر آنها در پایگاه های یادگیری دیجیتال را اندازه گیری کردند. مطالعه موردی آنها نشان داد که قرار دادن نقاط کنترلی مناسب در فرایندهای ایجاد فرا داده ها در چرخه عمر یک پایگاه داده می تواند تاحد زیادی به بهبود بیانجامد.

در دنیای نوظهور دیجیتال، که در آن کتابخانه های دیجیتال نقش کلیدی را هم در مقام انتشاردهنده و هم جمع آوری کننده بازی می کنند، نه فقط کنترل کیفیت فرا داده ها بلکه ارزیابی آنها نیز مشکل است

(راویت و دانا، ۲۰۱۵). همانگونه که هیلمن (۲۰۰۸) می نویسد "از آنجا که ساختار قدرتمند برپایه کیفیت در دنیای استاندارد مارک (شامل استاندارد بالغ، مستندسازی نهاده شده و اطلاعات کتابشناختی تسهیل شده) در دنیای فراداده شامل نقص هایی است، کاربران فراداده تصمیم به بهبود کیفیت فراداده در کتابخانه های خود گرفته اند تا بتوانند فرایندهایی را توسعه بدهند که یک همکاری مشترک سودمند را به همراه داشته باشد."

مشاهده ای آنچه در صفحه اول یک مدرک علمی ظاهر می شود این موضوع را گواهی می دهد که نگرش کاربر نهایی این اطلاعات در دهه گذشته یا قبل تر تا به امروز تغییر یافته است. همانگونه که در بیانیه فدراسیون بین المللی انجمن های کتابخانه ها و موسسات (IFLA^۱) درباره ی اصول بین المللی فهرست نویسی گفته شده است کیفیت فراداده ارتباط زیادی با آسان سازی کشف، شناسایی، انتخاب و استفاده از منابع اطلاعاتی مورد نیاز کاربران نهایی تعیین شده دارد (دی و همکاران، ۲۰۰۴). با این حال، سخنرانان در ویناری در سازمان استانداردهای ملی اطلاعات (NISO^۲) در مورد کتاب های الکترونیکی تأکید کردند که فراداده دیگر فقط ویژه استفاده ی ناشران و کتابداران نیست، بلکه این برای خوانندگان و کاربران نهایی است (دیوید و توماس، ۲۰۱۵). کیفیت خوب فراداده خرید و گردش عناوین کتاب های الکترونیکی را ترویج می کند. خوانندگان می خواهند قبل از تصمیم گیری برای خواندن آن کتاب الکترونیکی در مورد آن بدانند. با نگاه کردن به جلد، دسترسی به فهرست مطالب و کشف حقایقی در مورد آن نسخه خاص، کاربران قبل از خرید یا قرض گرفتن آن، در عمل کتاب را دنبال می کنند. شناخت این واقعیت بدان معنی است که حتی برای ارتقاء فروش عناوین کتاب های الکترونیکی نیز، فراداده باکیفیت باید همخوان با روش های مشخص ساخته و نگهداری شود (پارک و همکاران، ۲۰۱۰).

۲-۲ اهمیت کنترل کیفیت آماری در فرایند نمایه سازی و ویراستاری

کنترل کیفیت آماری^۳ (SQC) به روشهای مختلفی انجام می شود که یکی از انواع این روشها، نمونه گیری برای پذیرش است (Montgomery, 2005). در نمونه گیری برای پذیرش داده از

^۱ International Federation of Library Associations and Institutions

^۲ National Information Standards organization (NISO)

^۳ Statistical Quality Control

انباشته‌هایی با اندازه مشخص تشکیل می‌شود که پس از نمونه‌گیری از هر انباشته، در خصوص کیفیت آن انباشته قضاوت می‌شود. یکی از استانداردهای معرفی شده به منظور استفاده از نمونه‌گیری برای پذیرش، استاندارد Mil-STD-105E است.

بازرسی برای پذیرش معمولاً به یکی از سه روش نمودار کنترل، بازرسی صد در صد و یا نمونه‌گیری برای پذیرش انجام می‌شود. زمانی که از نمودارهای کنترل استفاده می‌شود، قضاوت بر اساس نتایج حاصل از نمودارهایی صورت می‌گیرد که در مراحل مختلف تولید و پردازش داده، تهیه شده است. در حالی که در روش بازرسی صد در صد همه‌ی داده‌های موجود در یک انباشته انباشته مورد بازرسی قرار می‌گیرند. منظور از انباشته، جامعه‌ای است که قرار است بررسی روی آن انجام شود. چنانچه داده از انباشته‌هایی با اندازه مشخص تشکیل شود و پس از نمونه‌گیری از هر انباشته و بازرسی نمونه‌ها، در خصوص کیفیت انباشته قضاوت گردد، نمونه‌گیری برای پذیرش گفته می‌شود. نمونه‌گیری برای پذیرش به یکی از شیوه‌های انباشته به انباشته، پیوسته و یا متوالی صورت می‌گیرد. نمونه‌گیری برای پذیرش انباشته به انباشته بیشترین کاربرد را در میان روش‌های نمونه‌گیری برای پذیرش دارد.

در این روش داده‌ها، در انباشته‌هایی که تعداد محدودی دارند، دسته‌بندی می‌گردند. پس از آن نمونه‌ای تصادفی با اندازه‌ی مشخص از هر یک از این انباشته‌ها گرفته می‌شود و همه‌ی قطعات نمونه، مورد بازرسی قرار می‌گیرند (Montgomery, 2005).

روشهای نمونه‌گیری برای پذیرش روشهایی برای ارزیابی انباشته هستند نه برآورد یا ایجاد کیفیت انباشته. همچنین چون فقط به بازرسی می‌پردازند، کیفیت ایجاد نمی‌کنند. به عبارت دیگر، این روش‌ها فقط به عنوان ابزار ممیزی برای کسب اطمینان از انطباق خروجی فرایند با خواسته‌های مورد نظر استفاده می‌شوند. از مهم‌ترین استانداردهایی که در این زمینه مورد استفاده قرار می‌گیرد، استاندارد MIL STD 105E است، که با توجه به سطح کیفیت قابل قبول (AQL- Acceptance Quality Level) طراحی شده است (Schilling and Neubauer, 2009).

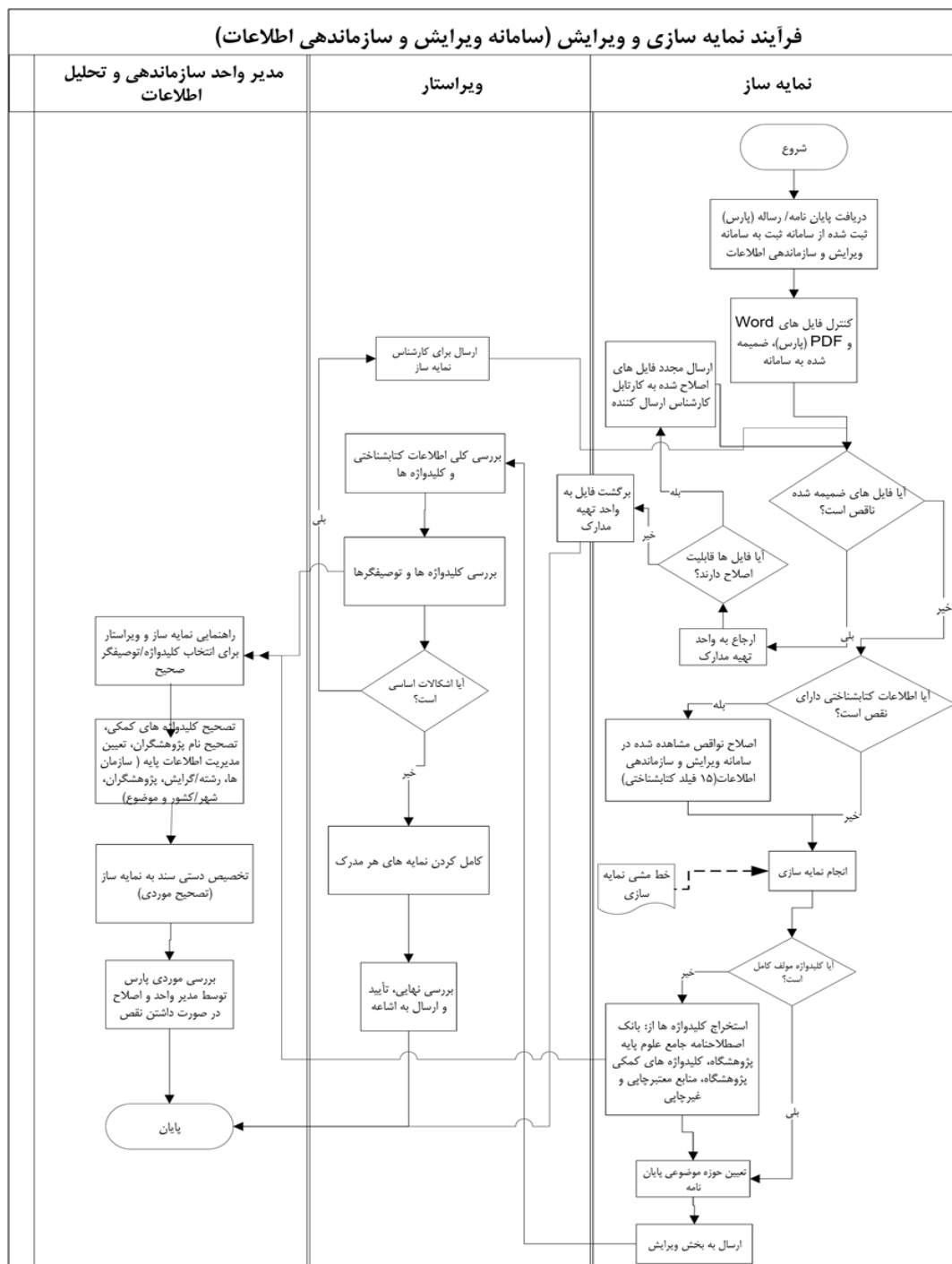
روش های مختلفی برای تقسیم بندی طرح های نمونه گیری برای پذیرش وجود دارد. یکی از انواع این تقسیم بندیها شامل طرحهای یک بار، دوبار و چندبار نمونه گیری است.

در طرح های یک بار نمونه گیری معیار پذیرش انباشته، نتایج فقط بر اساس یک نمونه ی تصادفی با اندازه ی مشخص می باشد، در حالی که طرح های دوبار یا چندبار نمونه گیری می تواند براساس بازرسی دو یا چند نمونه صورت پذیرد. در واقع یک طرح یک بار، روشی برای ارزیابی کیفیت انباشته است که با اندازه نمونه ی n تایی از انباشته و عددپذیرش C مشخص می شود (Schilling and Neubauer, 2009). با استفاده از این روش ابتدا یک نمونه ی n تایی به صورت تصادفی انتخاب می شود اگر به تعداد عدد پذیرش C یا کم تر، محصول معیوب در نمونه مشاهده شد انباشته پذیرفته و اگر بیش از عدد C مشاهده شد، انباشته رد می شود.

سایر طرحهای نمونه برداری به دلیل پیچیدگی اجرا در این طرح پژوهشی مورد توجه قرار نمی گیرند.

پایگاه گنج، گنجینه ای با ارزش از اطلاعات علمی و فراداده های اساتید، دانشجویان و محققان ایرانی است. این پایگاه علی رغم 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 ایرانداک در حوزه پارساهای داخل کشور بهتر محقق شود. بخشی از نابسامانی ها در این سامانه ناشی از خطاهای انسانی است که به هنگام نمایه سازی و ورود اطلاعات در بخش سازماندهی اطلاعات صورت گرفته است (ارشادی و همکاران، ۱۳۹۵). بخشی دیگر از ایرادها منشاء رایانه ای و سیستمی دارد که به مرور زمان و با تغییر نرم افزارها و سخت افزارها و کاراکترها و در هنگام تبدیلهای مختلف و ورود ماشینی اطلاعات در پایگاه به وجود آمده است. فرایند فراهم آوری، سازماندهی و اشاعه اطلاعات علمی که منجر به تولید پایگاه گنج می شود، فرایندی جاری، پیوسته و در حال انجام است (ارشادی و ایاسی، ۱۳۹۶). موفقیت یا شکست پایگاههای اطلاعاتی تا حد زیادی با کیفیت داده های موجود به عنوان مبنایی برای برنامه های کاربردی آن پایگاهها در ارتباط است. بنابراین یک بخش جدایی ناپذیر از هر پایگاه، ادغام داده های حاصل از سیستم های عامل می باشد. قبل از شروع فرآیند ادغام یک سیستم منبع، تحلیل غنی داده های منبع مورد نیاز است.

ارشادی و احترامی (۱۳۹۵) در پژوهش خود فرایند نمایه سازی و ویرایش را که یکی از مهمترین مراحل در چرخه تولید و اشاعه پارساها به حساب می آید مستند کردند (شکل ۳)



شکل ۱-۲ فرایند نمایه سازی و ویرایش پارسا

همانگونه که در بخش ۴ از سند حاضر اشاره شده است و در شکل ۱-۲ نیز جزییات آن به نمایش در آمده است کنترل های مختلفی در فرایند نمایه سازی و ویرایش بر روی پارساها انجام می گیرد، که در وضعیت کنونی برای نمونه برداری های صورت گرفته و همچنین اقدامات بعدی آن رویه مشخصی وجود ندارد. نداشتن خط مشی مشخص در نمونه برداری، کنترل های انجام شده و تصمیم گیری های بعدی به این منجر خواهد شد که کارایی کنترل ها کاهش پیدا کرده و در نتیجه برخی خطاها با وجود کنترل های پی در پی در سامانه گنج مشاهده شود. به عبارتی مسئولیت افراد واحد مدیریت و سازماندهی اطلاعات در کیفیت داده/فرا داده نهایی اشاعه داده شده در گنج مشخص نبوده و امکان ارائه بازخورد و اقدام اصلاحی مناسب در رابطه با ناهمخوان بودن داده های اشاعه داده شده به درستی فراهم نشود. همچنین مخاطره دیگری که نمونه برداری های ۱۰۰ درصد و پیاپی دارد اطمینان کارشناس مرحله قبلی به عملکرد مرحله بعدی و در نتیجه افت کارایی کنترل در مراحل قبلی خواهد بود. از این رو وجود طرح های نمونه برداری مشخص می تواند کارایی کنترل های هر مرحله را افزایش دهد.

برپایه موارد شرح داده شده در پژوهش جاری برای مرحله های ۲ تا ۳ شکل ۱، طرح های نمونه برداری طراحی خواهد شد. همچنین خروجی دیگر این پژوهش تعیین مسئول اصلی (کارشناس واحد فراهم آوری/نمایه ساز/ ویراستار) و پاسخگوی کیفیت در هر یک از اقلام اطلاعاتی داده/فرا داده پارسا است.

فصل سوم

ارزیابی داده / فراداده‌های ثبت شده در سامانه ثبت جدید

۳-۱ مقدمه

در فصل اول از این گزارش به ارائه اهمیت ارزیابی و کنترل فراداده‌ها در پایگاه‌های اطلاعاتی علمی پرداخته شد و در فصل دوم نیز مهم‌ترین پژوهش‌های پیشین این حوزه معرفی شد. همانگونه که در فصل اشاره شد، نمایه‌سازی و در ادامه آن ویراستاری مدارک در جهت اطمینان از کیفیت مدرک اشاعه داده شده بسیار حیاتی است. این کار در صورتی که مدارک ثبت شده از کیفیت مناسبی برخوردار باشد و همچنین تعداد مدارک ثبت شده زیاد بوده و کنترل اولیه‌ای بر روی آنها صورت پذیرفته باشد، می‌تواند در چارچوب یک طرح نمونه‌برداری صورت پذیرد. از آنجا که هدف اصلی این پژوهش ارائه طرح نمونه‌برداری از مدارک در فرایند نمایه‌سازی و ویراستاری است لازم است پیش از هرگونه ارائه طرح نمونه‌برداری از کیفیت داده‌های ثبت شده در سامانه ثبت جدید که در واحد ثبت و فراهم‌آوری اطلاعات نیز کنترل بر روی آنها صورت پذیرفته است ارزیابی اولیه‌ای ارائه نمود. از این رو در این فصل به ارائه نتایج ارزیابی از کیفیت داده‌های ثبت شده در سامانه ثبت جدید خواهیم پرداخت.

ارائه این توضیح ضروری است که همه اطلاعات ارائه شده در این فصل و چارچوب بیان شده در ادامه و همچنین تعداد نمونه انتخاب شده برای ارزیابی کیفیت داده‌ها در فرایندهای نمایه‌سازی و

ویراستاری به تأیید کارگروه اجرایی شامل معاون گرامی اطلاعات علم و فناوری، مدیر گرامی سازماندهی و تحلیل اطلاعات و مجری رسیده است.

در ادامه این فصل در بخش ۲-۳ اقلام اطلاعاتی فرایندهای نمایه سازی و ویراستاری که توسط نمایه سازان و ویراستاران کنترل می شود، معرفی خواهد شد. سپس، در بخش های ۳-۳ و ۴-۳ چارچوب ارزیابی کیفیت داده ها و تعیین ناهمخوانی ها به ترتیب در فرایندهای نمایه سازی و ویراستاری بیان خواهد شد. همچنین در بخش های ۵-۳ و ۶-۳ نتیجه ارزیابی ها برپایه چارچوب معرفی شده ارائه خواهد شد.

۳-۲ معرفی اقلام اطلاعاتی در فرایند نمایه سازی و ویراستاری

پیش از آغاز فرایند ارزیابی کیفیت در فرایندهای نمایه سازی و ویراستاری مدارک لازم است اقلام اطلاعاتی که در این دو بخش اصلی فرایند مورد ارزیابی قرار می گیرند در این بخش معرفی شوند. براین اساس در ادامه این بخش در قالب دو زیربخش مختلف اقلام اطلاعاتی مربوط به دو فرایند نمایه سازی و ویراستاری مدارک معرفی خواهند شد.

۳-۲-۱ اقلام اطلاعاتی فرایند نمایه سازی

در فرایند نمایه سازی فراداده ثبت شده در سامانه ثبت جدید و کنترل شده در فرایند کنترل (در واحد ثبت و فراهم آوری اطلاعات) مورد بررسی قرار می گیرد. فایل های ورد و پی.دی.اف به فراداده ضمیمه شده و در کارتابل کارشناس نمایه سازی قرار می گیرد. قلم های اطلاعاتی کنترل شده در این واحد عبارتند از:

- ردیف
- کد رهگیری
- عنوان فارسی

- عنوان لاتین
- پدیدآوران (دانشگاه، دانشکده، دانشجو، استاد راهنما، استاد مشاور)
- تاریخ دفاع
- اطلاعات تخصصی پایان نامه شامل: مقطع تحصیلی، گروه، رشته و گرایش
- مشخصات ظاهری مدرک
- چکیده فارسی و چکیده انگلیسی
- فهرست‌ها (مندرجات، منابع فارسی، منابع لاتین)
- حوزه موضوعی

در مورد فهرست اقلام اطلاعاتی بالا بیان چند نکته زیر ضروری است.

- از میان همه قلم‌های اطلاعاتی بالا تعداد محدودی قلم اطلاعاتی توسط نمایه‌ساز به قلم‌های اطلاعاتی مدرک اضافه می‌شود. حوزه موضوعی، مشخصات ظاهری مدرک و کلیدواژه مواردی هستند که توسط نمایه‌ساز اضافه می‌شوند.
- حوزه موضوعی در دو قلم اطلاعاتی حوزه موضوعی سطح اول و حوزه موضوعی سطح دوم، از یک فهرست از پیش تعیین شده توسط نمایه‌ساز انتخاب می‌شود. این دو حوزه نوع مدرک (هنر و علوم انسانی، علوم اجتماعی، و علوم و علوم کاربردی) را در سطح اول مشخص می‌کند و سپس در سطح دوم پس از انتخاب سطح اول فهرستی از موضوع‌های سطح دوم ارائه می‌کند که نمایه‌ساز از میان آنها یکی را انتخاب می‌کند. این دو سطح در سامانه گنج قابلیت جستجو و نمایش دارد.
- بخش اطلاعات عمومی دارای دو فیلد است، زبان‌های انتخابی که شامل زبان نوشتاری مدرک است و مشخصات ظاهری که مدرک شامل جدول، تصویر، نمودار، نقشه، واژه‌نامه،

پرسشنامه، سایر ضمائم و کتابنامه است که توسط نمایه ساز برپایه محتوای مدرک تیک زده می شود.

- علاوه بر کلیدواژه هایی که کاربر انتخاب کرده است (در زمان ثبت مدرک)، نمایه ساز نیز برپایه اصطلاحنامه ها و تزاروس ها تعدادی را انتخاب خواهد کرد. کلیدواژه های انتخابی در سامانه گنج با رنگ متمایز نمایش داده خواهد شد و به بازیابی مناسب مدرک کمک شایانی خواهد کرد.

۲-۲-۳ اقسام اطلاعاتی فرایند ویراستاری

در فرایند ویراستاری اطلاعات تازه ای به مدرک ثبت و نمایه شده افزوده نخواهد شد. بلکه قلم های اطلاعاتی که در فرایند نمایه سازی کنترل شده و همچنین به ویژه سه قلم اطلاعاتی که توسط نمایه ساز به مدرک اضافه شده است مورد بررسی دوباره قرار خواهد گرفت و اشتباهات اصلاح و پس از تأیید، مدرک اشاعه داده خواهد شد.

۳-۳ چارچوب ارزیابی کیفیت مدارک ثبت شده در فرایند نمایه سازی

در این بخش باتوجه به این که گوناگونی در ناهمخوانی ها در هر قلم اطلاعاتی قابل توجه بود، برای هر قلم اطلاعاتی اشاره شده در بخش قبلی چارچوبی برای دسته بندی ناهمخوانی ها ارائه شد. در ادامه در جدول ۳-۱ به صورت جداگانه به ازای هر قلم اطلاعاتی دسته بندی ناهمخوانی ها ارائه شده است.

جدول ۳-۱ ناهمخوانی های بالقوه در فرایند نمایه سازی پارسا

ردیف	نام قلم اطلاعاتی	دسته ناهمخوانی	چگونگی ثبت و کنترل ناهمخوانی
۱	عنوان فارسی	مشکل ویرایشی / نگارشی	ناهمخوانی ویرایشی و نگارشی: مانند نیم فاصله، د به جای ذ، ر به جای ز، فاصله اضافه و ترکیب شیمیایی، فرمول ریاضی (توان و علائم ریاضی) در صورت تأیید درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"

ردیف	نام قلم اطلاعاتی	دسته ناهمخوانی	چگونگی ثبت و کنترل ناهمخوانی
		مشکل اصلی (صحت/ ناهمخوانی با متن و ...)	ایراد اصلی مانند ناهمخوانی با متن پایان نامه و ... در صورت تأیید درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"
		مشکل ویرایشی/ نگارشی	در صورت تأیید درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"
۲	عنوان لاتین	مشکل اصلی (صحت/ ناهمخوانی با متن و ...)	در صورت تأیید درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"
		دانشگاه	
		دانشکده	
		دانشجو	
		اساتید راهنما	
		اساتید مشاور	
۳	پدیدآوران		در صورت تأیید درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"
۴	تاریخ دفاع	-	در صورت همخوانی با صفحه عنوان، درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"
۵	رشته	-	در صورت همخوانی با صفحه عنوان، درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"
۶	گرایش	-	در صورت همخوانی با صفحه عنوان، درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"
		ناهمخوانی ویرایشی	ناهمخوانی ویرایشی و نگارشی: مانند نیم فاصله، د به جای ذ، ر به جای ز، فاصله اضافه و
		ناهمخوانی نگارشی	در صورت تأیید درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"
		وجود کلمات اضافی	در صورت مشاهده کلمات اضافی (مانند چکیده، کلیدواژه، اطلاعات دانشجو و غیره) در صورت تأیید درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"
		ناهمخوانی با متن پایان نامه	در صورت تأیید درج عبارت "همخوانی دارد" و در صورت ناهمخوانی درج عبارت "همخوانی ندارد"
		اشکال در اعشار/فرمولها و روابط ریاضی/	در صورت مشاهده اشکال درج عبارت "مشکل دارد" و در صورت ناهمخوانی درج عبارت "مشکل ندارد"
۷	چکیده فارسی		

ردیف	نام قلم اطلاعاتی	دسته ناهمخوانی	چگونگی ثبت و کنترل ناهمخوانی
		شیمی	
۸	چکیده لاتین	ناهمخوانی ویرایشی	ناهمخوانی ویرایشی و نگارشی: مانند نیم فاصله، د به جای ذ، ر به جای ز، فاصله اضافه و
		ناهمخوانی نگارشی	در صورت تأیید درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"
		وجود کلمات اضافی	در صورت مشاهده کلمات اضافی (مانند چکیده، کلیدواژه، اطلاعات دانشجو و غیره)
			در صورت تأیید درج عبارت "مشکل ندارد" و در صورت ناهمخوانی درج عبارت "مشکل دارد"
		ناهمخوانی با متن پایان نامه	در صورت تأیید درج عبارت "همخوانی دارد" و در صورت ناهمخوانی درج عبارت "همخوانی ندارد"
		اشکال در اعشار/فرمولها و روابط ریاضی / شیمی	در صورت مشاهده اشکال درج عبارت "مشکل دارد" و در صورت همخوانی درج عبارت "مشکل ندارد"
۹	فهرست مندرجات	-	در صورت همخوانی با متن و همچنین مشاهده نشدن مشکل عبارت "مشکل ندارد" درج شود. در صورت خالی بودن قلم
۱۰	فهرست منابع فارسی	-	اطلاعاتی درج عبارت "موجود نیست". در صورت ناقص بودن قلم
۱۱	فهرست منابع لاتین	-	اطلاعاتی در مقایسه با متن پارسا درج عبارت "ناقص است" و همچنین در صورت به هم چسبیده بودن درج عبارت "به هم چسبیده" درج شود.
			عبارت "جابجایی فهرست فارسی و لاتین" در صورت مشاهده ناهمخوانی جابجایی میان فهرست‌های فارسی و لاتین

۳-۴ چارچوب ارزیابی کیفیت مدارک ثبت شده در فرایند ویراستاری

از آنجاکه مدارک پارسا پس از نمایه‌سازی توسط نمایه‌سازان برای بررسی نهایی برای ویراستاران ارسال می‌شود، پایه و اساس بررسی‌ها در فرایند ویراستاری اقلام اطلاعاتی درج شده در جدول ۳-۱ خواهد بود. به بیان دیگر، همه اقلام اطلاعاتی جدول ۳-۱ دوباره در فرایند ویراستاری کنترل خواهد شد. علاوه بر این اقلام حوزه موضوعی، مشخصات ظاهری مدرک و کلیدواژه که توسط نمایه‌ساز به

مدرك اضافه شده است توسط ویراستار مورد بررسی قرار خواهد گرفت. در بخش ۳-۲-۱ چگونگی اضافه شدن این سه دسته قلم اطلاعاتی توسط نمایه ساز به هر مدرک توضیح داده شده است.

در ادامه این بخش و در بخش ۳-۵ نتایج ارزیابی کیفیت مدارک ثبت شده در سامانه ثبت (توسط دانشجویان تحصیلات تکمیلی) پس از کنترل هایی که توسط کارشناسان اداره ثبت و فراهم آوری اطلاعات صورت پذیرفته است ارائه خواهد شد.

۳-۵ نتایج ارزیابی کیفیت مدارک ثبت شده در فرایند نمایه سازی

در این بخش با همکاری نمایه سازان گرامی در واحد مدیریت سازماندهی و تحلیل اطلاعات و همچنین برپایه چارچوب ارائه شده در بخش ۳-۳ داده ها و فراداده های پارساهای ثبت شده در سامانه ثبت جدید ارزیابی و به صورت همزمان نمایه سازی شد. نتایج این ارزیابی در ادامه این بخش و در قالب سه زیربخش شامل ارزیابی کیفیت اقلام صفحه عنوان، ارزیابی کیفیت چکیده های فارسی و لاتین و همچنین ارزیابی کیفیت فهرست ها ارائه خواهد شد.

۳-۵-۱ ارزیابی کیفیت اقلام صفحه عنوان

جداول ۲-۳ تا ۴-۳ نتایج تحلیل بر روی کنترل های انجام شده بر روی مدارک پارسای ثبت شده در سامانه ثبت را توسط نمایه سازان نشان می دهد. یادآوری این نکته ضروری است که همانگونه که پیش از این نیز اشاره شد این مدارک پس از تأیید همکاران در اداره ثبت و فراهم آوری اطلاعات وارد سامانه ویرایش شده اند. به بیان دیگر، کنترل کیفیت اولیه ای بر روی اقلام مختلف سامانه ثبت پارسا صورت پذیرفته و اکنون در مرحله نمایه سازی (و پیش از انجام نمایه سازی) ارزیابی کیفیت آنها انجام شده است.

جدول ۲-۳ نتایج ارزیابی کیفیت اقلام اطلاعاتی صفحه عنوان پارساها در نمایه سازی را نشان می دهد. همانگونه که از نتایج ارائه شده مشاهده می شود نام دانشکده با میزان ۶۵٫۵ درصد بیشترین ناهمخوانی را در میان اقلام اطلاعاتی دارد. پس از آن ویرایش و نگارش عنوان فارسی با ۵۸٫۲۴

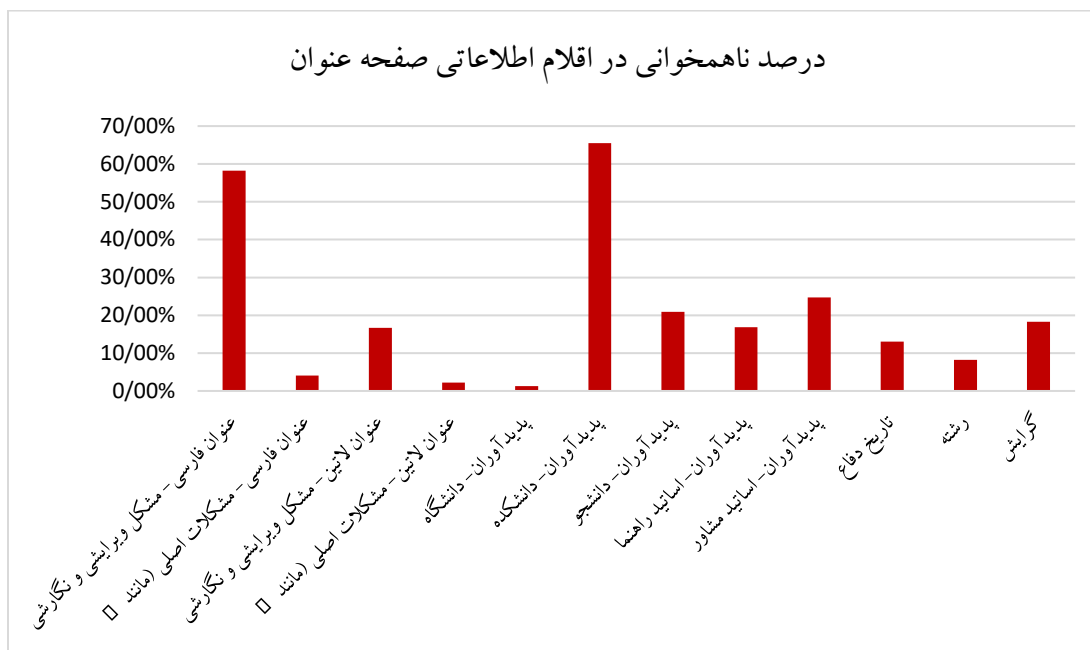
ارائه طرح نمونه‌برداری پذیرش در فرایند ویراستاری □ ۳۲

درصد دارای بیشترین ناهمخوانی است. نام دانشگاه و ناهمخوانی‌های اصلی عنوان لاتین کمترین ناهمخوانی‌ها را در میان سایر اقلام اطلاعاتی صفحه عنوان دارند.

نمودار ارائه شده در شکل ۳-۱ به صورت گرافیکی تحلیلی از وضعیت میزان ناهمخوانی‌ها در اقلام اطلاعاتی صفحه عنوان پارساها به نمایش می‌گذارد.

جدول ۲-۳ نتایج ارزیابی کیفیت صفحه عنوان پارساها در نمایه‌سازی

ردیف	نام مشخصه کیفی ارزیابی شده در نمایه‌سازی	تعداد ناهمخوانی	تعداد موارد همخوان	درصد ناهمخوانی
۱	عنوان فارسی - مشکل ویرایشی و نگارشی	318	228	58/24%
۲	عنوان فارسی - مشکلات اصلی (مانند ناهمخوانی با متن پایان نامه و غیره)	22	523	4/04%
۳	عنوان لاتین - مشکل ویرایشی و نگارشی	91	454	16/70%
۴	عنوان لاتین - مشکلات اصلی (مانند ناهمخوانی با متن پایان نامه و غیره)	12	533	2/20%
۵	پدیدآوران - دانشگاه	7	538	1/28%
۶	پدیدآوران - دانشکده	357	188	65/50%
۷	پدیدآوران - دانشجو	114	432	20/88%
۸	پدیدآوران - اساتید راهنما	92	454	16/85%
۹	پدیدآوران - اساتید مشاور	135	411	24/73%
۱۰	تاریخ دفاع	71	475	13/00%
۱۱	رشته	45	501	8/24%
۱۲	گرایش	100	446	18/32%



شکل ۳-۱ نمودار مقایسه کیفیت اقلام صفحه عنوان پارساها در نمایه سازی

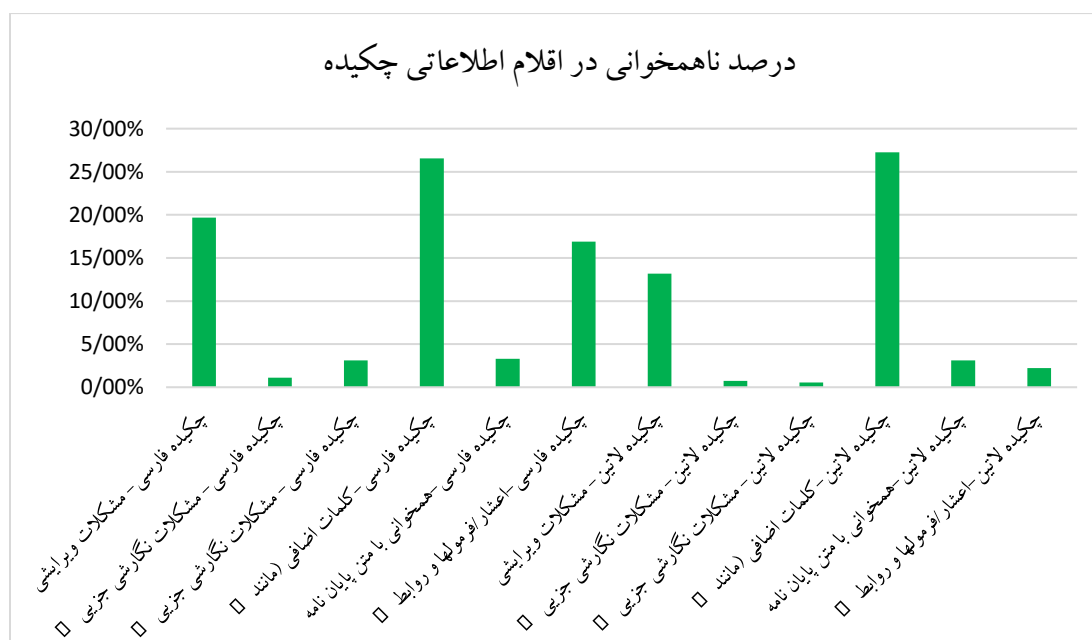
۲-۵-۳ ارزیابی کیفیت چکیده های فارسی و لاتین

جدول ۳-۳ نتایج ارزیابی کیفیت اقلام اطلاعاتی چکیده های فارسی و لاتین پارساها در فرایند نمایه سازی را نشان می دهد. هماهنگونه که از نتایج ارائه شده مشاهده می شود چکیده لاتین - کلمات اضافی (مانند چکیده، کلیدواژه، اطلاعات دانشجو) با میزان ۲۶,۲۶ درصد بیشترین ناهمخوانی را در میان اقلام اطلاعاتی دارد. پس از آن چکیده لاتین - کلمات اضافی (مانند چکیده، کلیدواژه، اطلاعات دانشجو) با ۲۶,۵۶ درصد دارای بیشترین ناهمخوانی است. چکیده لاتین - مشکلات نگارشی (۳ و بیشتر از ۳ مورد) و همچنین چکیده لاتین - مشکلات نگارشی (کمتر از ۳ مورد) کمترین ناهمخوانی ها را در میان سایر اقلام اطلاعاتی چکیده های فارسی و لاتین دارند.

نمودار ارائه شده در شکل ۳-۲ به صورت گرافیکی تحلیلی از وضعیت میزان ناهمخوانی ها در اقلام چکیده های فارسی و لاتین پارساها را به نمایش می گذارد.

جدول ۳-۳ نتایج ارزیابی کیفیت چکیده‌های فارسی و لاتین پارساها در نمایه‌سازی

ردیف	نام مشخصه کیفی ارزیابی شده در نمایه‌سازی	تعداد ناهمخوانی	تعداد موارد همخوان	درصد ناهمخوانی
۱	چکیده فارسی - مشکلات ویرایشی	107	437	19/67%
۲	چکیده فارسی - مشکلات نگارشی جزئی (کمتر از ۳ مورد)	6	522	1/10%
۳	چکیده فارسی - مشکلات نگارشی (۳ و بیشتر از ۳ مورد)	17		3/12%
۴	چکیده فارسی - کلمات اضافی (مانند چکیده، کلیدواژه، اطلاعات دانشجو)	145	401	26/56%
۵	چکیده فارسی - همخوانی با متن پایان نامه	18	527	3/30%
۶	چکیده فارسی - اعشار/فرمولها و روابط ریاضی / شیمی	92	453	16/88%
۷	چکیده لاتین - مشکلات ویرایشی	70	461	13/18%
۸	چکیده لاتین - مشکلات نگارشی جزئی (کمتر از ۳ مورد)	4	536	0/74%
۹	چکیده لاتین - مشکلات نگارشی (۳ و بیشتر از ۳ مورد)	3		0/55%
۱۰	چکیده لاتین - کلمات اضافی (مانند چکیده، کلیدواژه، اطلاعات دانشجو)	148	395	27/26%
۱۱	چکیده لاتین - همخوانی با متن پایان نامه	17	526	3/13%
۱۲	چکیده لاتین - اعشار/فرمولها و روابط ریاضی / شیمی	12	531	2/21%



شکل ۳-۲ نمودار مقایسه کیفیت اقلام چکیده‌های فارسی و لاتین پارساها در نمایه‌سازی

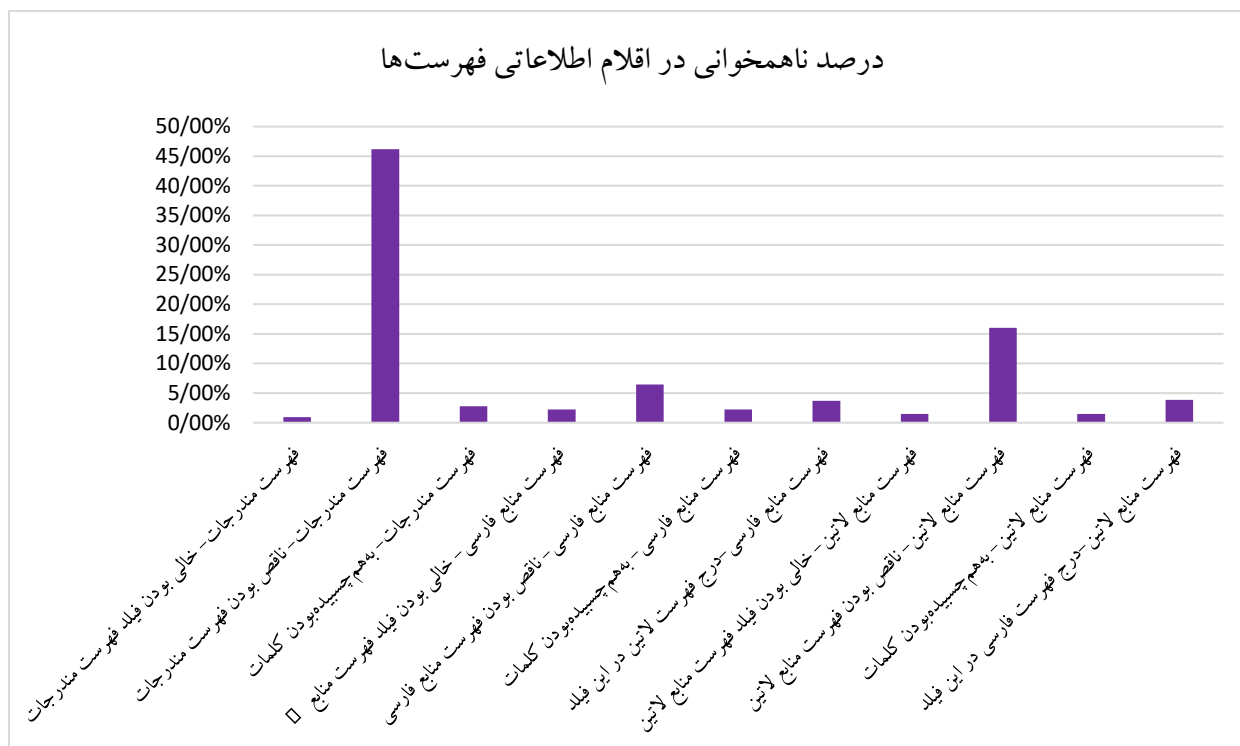
۳-۵-۳ ارزیابی کیفیت فهرست‌ها

جدول ۳-۴ نتایج ارزیابی کیفیت اقلام اطلاعاتی فهرست‌های پارساها را در فرایند نمایه‌سازی نشان می‌دهد. فهرست مندرجات، فهرست منابع فارسی و نیز فهرست منابع لاتین سه دسته اساسی این اقلام هستند. ناقص بودن فهرست مندرجات با ۴۶,۱۵ درصد بیشترین ناهمخوانی را در میان اقلام اطلاعاتی فهرست‌ها داراست. پس از آن ناقص بودن فهرست منابع لاتین با ۱۵,۹۹ درصد دارای بیشترین ناهمخوانی است. خالی بودن فهرست منابع لاتین و به‌هم‌چسبیده‌بودن کلمات در فهرست لاتین کمترین ناهمخوانی‌ها را در میان سایر اقلام اطلاعاتی فهرست‌ها دارند.

نمودار ارائه شده در شکل ۳-۳ به‌صورت گرافیکی تحلیلی از وضعیت میزان ناهمخوانی‌ها را در اقلام فهرست‌های فارسی و لاتین پارساها به نمایش می‌گذارد.

جدول ۳-۴ نتایج ارزیابی کیفیت فهرست‌های پارساها در نمایه‌سازی

ردیف	نام مشخصه کیفی ارزیابی شده در نمایه‌سازی	تعداد ناهمخوانی	تعداد موارد همخوان	درصد ناهمخوانی
۱	فهرست مندرجات - خالی بودن فیلد فهرست مندرجات	5	274	0/92%
۲	فهرست مندرجات - ناقص بودن فهرست مندرجات	252		46/15%
۳	فهرست مندرجات - به‌هم‌چسبیده‌بودن کلمات	15		2/75%
۴	فهرست منابع فارسی - خالی بودن فیلد فهرست منابع فارسی	12	464	2/21%
۵	فهرست منابع فارسی - ناقص بودن فهرست منابع فارسی	35		6/45%
۶	فهرست منابع فارسی - به‌هم‌چسبیده‌بودن کلمات	12		2/21%
۷	فهرست منابع فارسی - درج فهرست لاتین در این فیلد	20		3/68%
۸	فهرست منابع لاتین - خالی بودن فیلد فهرست منابع لاتین	8	420	1/47%
۹	فهرست منابع لاتین - ناقص بودن فهرست منابع لاتین	87		15/99%
۱۰	فهرست منابع لاتین - به‌هم‌چسبیده‌بودن کلمات	8		1/47%
۱۱	فهرست منابع لاتین - درج فهرست فارسی در این فیلد	21		3/86%



شکل ۳-۳ نمودار مقایسه کیفیت اقلام فهرست‌های پارساها در نمایه‌سازی

در ادامه و در بخش ۳-۶ نتایج ارزیابی کیفیت مدارک نمایه‌سازی شده خواهیم پرداخت.

۳-۶ نتایج ارزیابی کیفیت مدارک ثبت شده در فرایند ویراستاری

در این بخش به تحلیل کیفیت اقلام اطلاعاتی پارسا پس از نمایه‌سازی خواهیم پرداخت. در فصل اول (به‌ویژه در شکل ۱-۱) توضیح داده شد که در فرایند ویراستاری، آخرین مرحله پیش از اشاعه مدارک، کنترل نهایی برروی فراداده‌های پارساها انجام خواهد شد و ناهمخوانی‌های جزئی در این بخش برطرف می‌شوند.

جداول ۳-۵ تا ۳-۸ نتایج تحلیل برروی کنترل‌های انجام شده برروی مدارک پارسای نمایه‌سازی شده در سامانه ویرایش را توسط ویراستاران نشان می‌دهد. این کنترل‌ها پس از نمایه‌سازی و در قالب فرایند ویراستاری انجام شده است. از آنجا که مدارک این مرحله دو مرحله کنترل (یکی در اداره ثبت و فراهم‌آوری اطلاعات و دیگری در نمایه‌سازی) گذرانده‌اند، انتظار می‌رود کیفیت مطلوب‌تری

داشته باشند. افزون بر این، ارزیابی کیفیت سایر اقلام اضافه شده در نمایه سازی نیز در این مرحله صورت می پذیرد که نتایج آن در انتها ارائه خواهد شد.

۱-۶-۳ ارزیابی کیفیت اقلام صفحه عنوان

جدول ۳-۵ نتایج ارزیابی کیفیت اقلام اطلاعاتی صفحه عنوان پارساها در مرحله ویراستاری را نشان می دهد. همانگونه که از نتایج ارائه شده مشاهده می شود به جز چهار ناهمخوانی که به میزان بسیار جزئی در اقلام اطلاعاتی صفحه عنوان پارساها مشاهده می شود، سایر اقلام اطلاعاتی صحیح می باشند.

برپایه کنترل های انجام شده در مرحله ویراستاری بر روی مدارکی که نمایه سازی شده اند می توان نتیجه گرفت که تنها در قلم های اطلاعاتی عنوان فارسی، عنوان لاتین، دانشکده، و گرایش تنها یک مورد از ۵۰۰ مدرک مشاهده شده دارای ناهمخوانی بوده است. در عنوان فارسی یک مورد ناهمخوانی مشاهده شده مربوط به مشکل ویرایشی و نگارشی بوده و در عنوان لاتین مشکل اصلی (مانند ناهمخوانی با متن پایان نامه و غیره) مشاهده شده است.

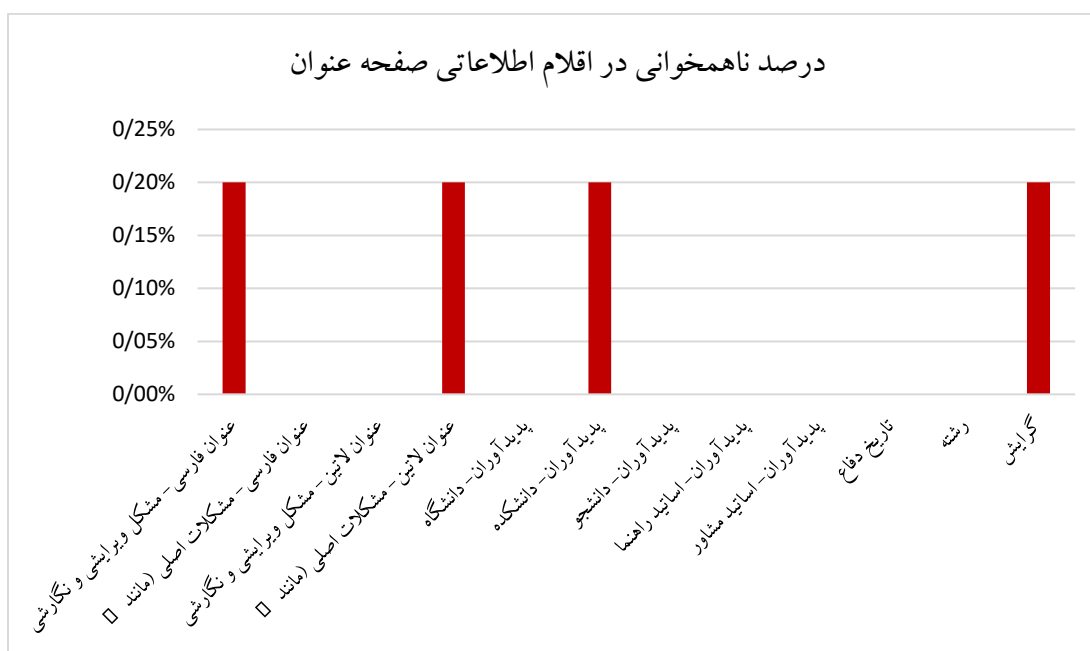
در شکل ۳-۴ نمودار مقایسه ای ناهمخوانی ها برپایه نتایج جدول ۳-۵ ارائه شده است.

جدول ۳-۵ نتایج ارزیابی اقلام صفحه عنوان پارساها در ویراستاری

ردیف	نام مشخصه کیفی ارزیابی شده در نمایه سازی	تعداد ناهمخوانی	تعداد موارد همخوان	درصد ناهمخوانی
۱	عنوان فارسی - مشکل ویرایشی و نگارشی	1	499	0/20%
۲	عنوان فارسی - مشکلات اصلی (مانند ناهمخوانی با متن پایان نامه و غیره)	0	500	0/00%
۳	عنوان لاتین - مشکل ویرایشی و نگارشی	0	500	0/00%
۴	عنوان لاتین - مشکلات اصلی (مانند ناهمخوانی با متن پایان نامه و غیره)	1	499	0/20%
۵	پدیدآوران - دانشگاه	0	500	0/00%
۶	پدیدآوران - دانشکده	1	499	0/20%
۷	پدیدآوران - دانشجو	0	500	0/00%
۸	پدیدآوران - اساتید راهنما	0	500	0/00%
۹	پدیدآوران - اساتید مشاور	0	500	0/00%

ارائه طرح نمونه برداری پذیرش در فرایند ویراستاری □ ۳۸

۱۰	تاریخ دفاع	0	500	0/00%
۱۱	رشته	0	500	0/00%
۱۲	گرایش	1	499	0/20%



شکل ۳-۴ نمودار مقایسه کیفیت اقلام صفحه عنوان پارساها در ویراستاری

در ادامه و در بخش ۳-۶-۲ نتایج ارزیابی کیفیت چکیده‌های فارسی و لاتین در مرحله ویراستاری ارائه می‌شود.

۲-۶-۳ ارزیابی کیفیت چکیده‌های فارسی و لاتین

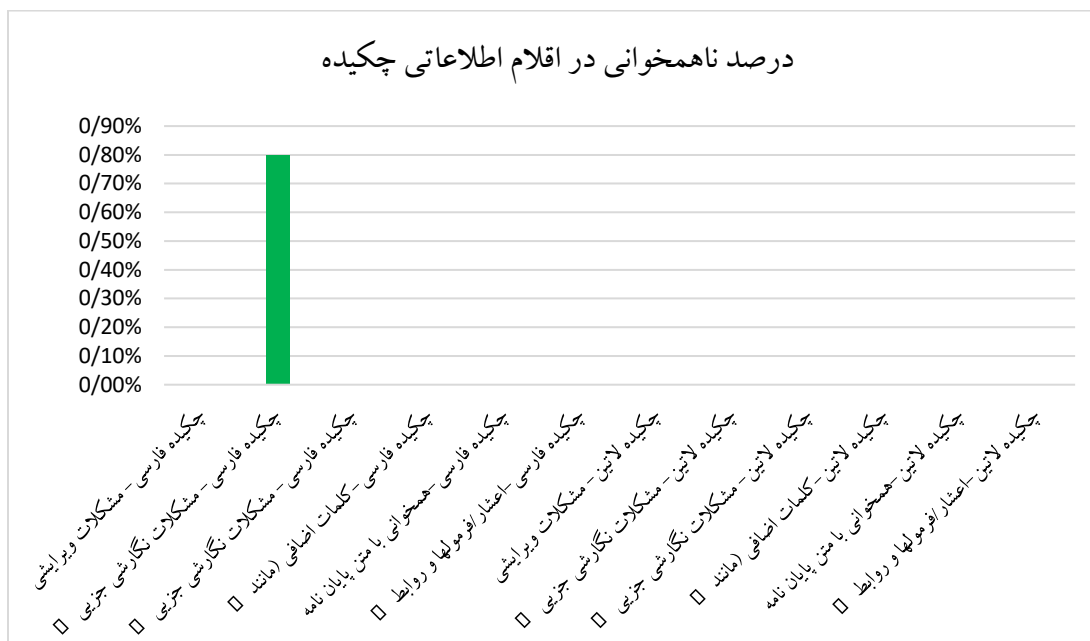
جدول ۳-۶ نتایج ارزیابی کیفیت اقلام اطلاعاتی چکیده‌های فارسی و لاتین پارساها در مرحله ویراستاری را نشان می‌دهد. همانگونه که از نتایج ارائه شده مشاهده می‌شود به جز چهارمورد ناهمخوانی که به میزان بسیار جزئی در اقلام اطلاعاتی چکیده‌های پارساها مشاهده می‌شود سایر اقلام اطلاعاتی نقصی نداشته‌اند.

برپایه کنترل‌های انجام شده در مرحله ویراستاری بر روی مدارکی که نمایه‌سازی شده‌اند می‌توان نتیجه گرفت که تنها در قلم اطلاعاتی چکیده فارسی تنها چهار مورد از ۵۰۰ مدرک مشاهده شده دارای ناهمخوانی بوده است. به بیان دیگر، ۴ مدرک دارای مشکلات نگارشی (کمتر از ۳ مورد) داشته و در بقیه مدرک‌ها هیچ ناهمخوانی مشاهده نشده است.

در شکل ۳-۵ نمودار مقایسه‌ای ناهمخوانی‌ها برپایه نتایج جدول ۳-۶ ارائه شده است.

جدول ۳-۶ نتایج ارزیابی اقلام چکیده‌های فارسی و لاتین پارساها در ویراستاری

ردیف	نام مشخصه کیفی ارزیابی شده در نمایه‌سازی	تعداد ناهمخوانی	تعداد موارد همخوان	درصد ناهمخوانی
۱	چکیده فارسی - مشکلات ویرایشی	۰	۵۰۰	۰/۰۰%
۲	چکیده فارسی - مشکلات نگارشی جزئی (کمتر از ۳ مورد)	۴	۴۹۶	۰/۸۰%
۳	چکیده فارسی - مشکلات نگارشی جزئی (۳ و بیشتر از ۳ مورد)	۰		۰/۰۰%
۴	چکیده فارسی - کلمات اضافی (مانند چکیده، کلیدواژه، اطلاعات دانشجو)	۰	۵۰۰	۰/۰۰%
۵	چکیده فارسی - همخوانی با متن پایان نامه	۰	۵۰۰	۰/۰۰%
۶	چکیده فارسی - اعشار/فرمولها و روابط ریاضی/ شیمی	۰	۵۰۰	۰/۰۰%
۷	چکیده لاتین - مشکلات ویرایشی	۰	۵۰۰	۰/۰۰%
۸	چکیده لاتین - مشکلات نگارشی جزئی (کمتر از ۳ مورد)	۰	۵۰۰	۰/۰۰%
۹	چکیده لاتین - مشکلات نگارشی جزئی (۳ و بیشتر از ۳ مورد)	۰		۰/۰۰%
۱۰	چکیده لاتین - کلمات اضافی (مانند چکیده، کلیدواژه، اطلاعات دانشجو)	۰	۵۰۰	۰/۰۰%
۱۱	چکیده لاتین - همخوانی با متن پایان نامه	۰	۵۰۰	۰/۰۰%
۱۲	چکیده لاتین - اعشار/فرمولها و روابط ریاضی/ شیمی	۰	۵۰۰	۰/۰۰%



شکل ۳-۵ نمودار مقایسه کیفیت اقلام اطلاعات چکیده پارساها در ویراستاری

در ادامه و در بخش ۳-۶-۳ نتایج ارزیابی کیفیت فهرست‌ها در مرحله ویراستاری ارائه می‌شود.

۳-۶-۳ ارزیابی کیفیت فهرست‌ها

جدول ۳-۷ نتایج ارزیابی کیفیت اقلام اطلاعاتی فهرست‌های مختلف پارساها در مرحله ویراستاری را نشان می‌دهد. همانگونه که از نتایج ارائه شده مشاهده می‌شود به جز یک مورد ناهمخوانی که به میزان بسیار جزئی در اقلام اطلاعاتی فهرست‌ها مشاهده می‌شود سایر اقلام اطلاعاتی نقصی نداشته‌اند.

برپایه کنترل‌های انجام شده در مرحله ویراستاری بر روی مدارکی که نمایه‌سازی شده‌اند می‌توان نتیجه گرفت که تنها در قلم اطلاعاتی فهرست مندرجات تنها یک مورد از ۵۰۰ مدرک مشاهده شده دارای ناهمخوانی بوده است. به بیان دیگر، تنها یک مدرک ناقص بوده و در بقیه مدرک‌ها هیچ ناهمخوانی مشاهده نشده است.

در شکل ۳-۶ نمودار مقایسه‌ای ناهمخوانی‌ها برپایه نتایج جدول ۳-۷ ارائه شده است.

جدول ۳-۷ نتایج ارزیابی کیفیت فهرست‌های پارساها در ویراستاری

ردیف	نام مشخصه کیفی ارزیابی شده در نمایه‌سازی	تعداد ناهمخوانی	تعداد موارد همخوان	درصد ناهمخوانی
۱	فهرست مندرجات - خالی بودن فیلد فهرست مندرجات	0	499	0/00%
۲	فهرست مندرجات - ناقص بودن فهرست مندرجات	1		0/20%
۳	فهرست مندرجات - به هم چسبیده بودن کلمات	0		0/00%
۴	فهرست منابع فارسی - خالی بودن فیلد فهرست منابع فارسی	0	500	0/00%
۵	فهرست منابع فارسی - ناقص بودن فهرست منابع فارسی	0		0/00%
۶	فهرست منابع فارسی - به هم چسبیده بودن کلمات	0		0/00%
۷	فهرست منابع فارسی - درج فهرست لاتین در این فیلد	0		0/00%
۸	فهرست منابع لاتین - خالی بودن فیلد فهرست منابع لاتین	0	500	0/00%
۹	فهرست منابع لاتین - ناقص بودن فهرست منابع لاتین	0		0/00%
۱۰	فهرست منابع لاتین - به هم چسبیده بودن کلمات	0		0/00%
۱۱	فهرست منابع لاتین - درج فهرست فارسی در این فیلد	0		0/00%



شکل ۳-۶ نمودار مقایسه کیفیت اقلام فهرست‌های پارساها در ویراستاری

همانگونه که در بخش در بخش ۳-۲-۱ مورد بحث قرار گرفت، در بخش نمایه‌سازی سه قلم اطلاعاتی تازه به مدارک پارسا اضافه می‌شوند. این سه قلم عبارتند از مشخصات ظاهری، حوزه موضوعی اول و حوزه موضوعی دوم.

کیفیت اقلام اضافه شده در نمایه‌سازی در مرحله ویراستاری مورد بررسی قرار می‌گیرد. برای این کار در قالب ۵۰۰ مدرک اشاره شده قبل، ارزیابی توسط کارشناسان واحد ویراستاری صورت پذیرفت که نتیجه آن در بخش بعدی ارائه می‌شود.

۴-۶-۳ ارزیابی کیفیت اقلام اضافه شده در نمایه‌سازی

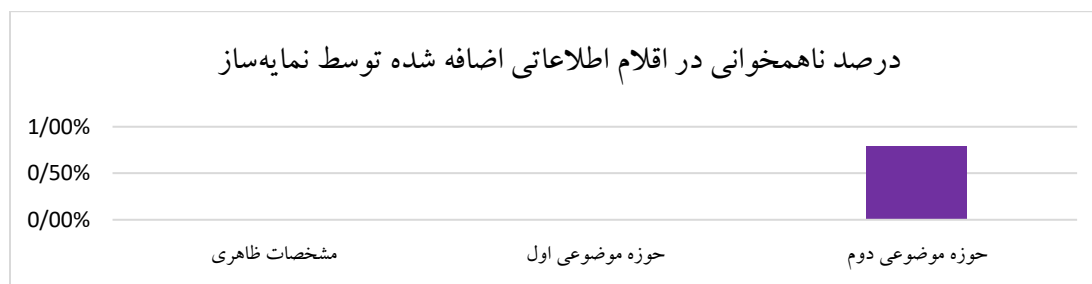
جدول ۳-۸ نتایج ارزیابی کیفیت اقلام اطلاعاتی اضافه شده در مرحله نمایه‌سازی را نشان می‌دهد. همانگونه که از نتایج ارائه شده برمی‌آید به جز یک قلم اطلاعاتی در سایر اقلام اطلاعاتی ناهمخوانی مشاهده نمی‌شود.

ارائه طرح نمونه برداری پذیرش در فرایند ویراستاری □ ۴۳

برپایه کنترل‌های انجام شده در مرحله ویراستاری بر روی مدارکی که نمایه‌سازی شده‌اند می‌توان نتیجه گرفت که تنها در قلم اطلاعاتی حوزه موضوعی دوم تنها چهار مورد از ۵۰۰ مدرک مشاهده شده دارای ناهمخوانی بوده است. به بیان دیگر، تنها چهار مدرک ناقص بوده و در بقیه مدرک‌ها هیچ ناهمخوانی مشاهده نشده است. در شکل ۳-۷ نمودار مقایسه‌ای ناهمخوانی‌ها برپایه نتایج جدول ۳-۸ ارائه شده است.

جدول ۳-۸ نتایج ارزیابی کیفیت اقلام اضافه شده توسط نمایه‌ساز در ویراستاری

ردیف	نام مشخصه کیفی ارزیابی شده در نمایه‌سازی	تعداد ناهمخوانی	تعداد موارد همخوان	درصد ناهمخوانی
۱	مشخصات ظاهری	۰	۵۰۰	۰/۰۰%
۲	حوزه موضوعی اول	۰	۵۰۰	۰/۰۰%
۳	حوزه موضوعی دوم	۴	۴۹۶	۰/۷۹%



شکل ۳-۷ نمودار مقایسه کیفیت اقلام اضافه شده توسط نمایه‌ساز در ویراستاری

در ادامه و در بخش ۳-۷ به جمع‌بندی و تحلیل نهایی اطلاعات جمع‌آوری شده از کیفیت نمایه‌سازی و ویراستاری مدارک پارسال خواهیم پرداخت.

۳-۷ جمع بندی و تحلیل نهایی

در این بخش نتایج بررسی میزان ایجاد بهبود در کیفیت مدارک پارسا پس از نمایه سازی و ویراستاری ارائه خواهد شد. همانگونه که در جدول ۳-۹ مشاهده می شود تمام مشخصه های کیفی در مدارک پارسا که نتایج تحلیل آنها جداگانه در بخش پیشین از فصل سوم ارائه شده بود در یک ستون آورده شده است. میزان بهبود صورت گرفته در ستون آخر نشان داده شده است. این عدد عبارتست از ستون درصد ناهمخوانی شناسایی شده در فرایند نمایه سازی منهای درصد ناهمخوانی در فرایند ویراستاری. همانطور که مشاهده می شود فرایند نمایه سازی توانسته است میزان قابل توجهی کیفیت مدارک پارسا را بهبود بخشد. همچنین، با تقریب بسیار خوبی می توان گفت کیفیت مدارک پس از نمایه سازی تا سطح نزدیک به ۱۰۰ درصد ارتقاء یافته اند. تنها در چند مشخصه ناهمخوانی اندکی مشاهده می شود که در ادامه نامهای آنها آورده شده است.

- عنوان فارسی - مشکل ویرایشی و نگارشی
- عنوان لاتین - مشکلات اصلی (مانند ناهمخوانی با متن پایان نامه و غیره)
- پدیدآوران - دانشکده
- گرایش
- چکیده فارسی - مشکلات نگارشی جزئی (کمتر از ۳ مورد)
- فهرست مندرجات - ناقص بودن فهرست مندرجات

در سه ردیف آخر از آنجا که این سه قلم اطلاعاتی در مرحله نمایه سازی اضافه شده بودند امکان ارائه میزان بهبود وجود نداشت اما همانگونه که از نتایج مشخص است فقط در حوزه موضوعی دوم مقدار ناچیزی ناهمخوانی دیده می شود.

جدول ۳-۹ بررسی میزان بهبود در اقسام اطلاعاتی مدارک پارسا

ردیف	نام مشخصه کیفی در مدارک پارسا	نمایه‌سازی			ویراستاری			میزان بهبود
		تعداد ناهمخوانی	تعداد موارد همخوان	درصد ناهمخوانی	تعداد موارد همخوان	تعداد ناهمخوانی	درصد ناهمخوانی	
۱	عنوان فارسی - مشکل ویرایشی و نگارشی	318	228	58/24%	1	499	0/20%	58/04%
۲	عنوان فارسی - مشکلات اصلی (مانند ناهمخوانی با متن پایان نامه و غیره)	22	523	4/04%	0	500	0/00%	4/04%
۳	عنوان لاتین - مشکل ویرایشی و نگارشی	91	454	16/70%	0	500	0/00%	16/70%
۴	عنوان لاتین - مشکلات اصلی (مانند ناهمخوانی با متن پایان نامه و غیره)	12	533	2/20%	1	499	0/20%	2/00%
۵	پدیدآوران - دانشگاه	7	538	1/28%	0	500	0/00%	1/28%
۶	پدیدآوران - دانشکده	357	188	65/50%	1	499	0/20%	65/30%
۷	پدیدآوران - دانشجو	114	432	20/88%	0	500	0/00%	20/88%
۸	پدیدآوران - اساتید راهنما	92	454	16/85%	0	500	0/00%	16/85%
۹	پدیدآوران - اساتید مشاور	135	411	24/73%	0	500	0/00%	24/73%
۱۰	تاریخ دفاع	71	475	13/00%	0	500	0/00%	13/00%
۱۱	رشته	45	501	8/24%	0	500	0/00%	8/24%
۱۲	گرایش	100	446	18/32%	1	499	0/20%	18/12%
۱۳	چکیده فارسی - مشکلات ویرایشی	107	437	19/67%	0	500	0/00%	19/67%
۱۴	چکیده فارسی - مشکلات نگارشی جزئی (کمتر از ۳ مورد)	6	522	1/10%	4	496	0/80%	0/3%
۱۵	چکیده فارسی - مشکلات نگارشی (۳ و بیشتر از ۳ مورد)	17	522	3/12%	0		0/00%	3/12%
۱۶	چکیده فارسی - کلمات اضافی (مانند چکیده، کلیدواژه، اطلاعات دانشجو)	145	401	26/56%	0	500	0/00%	26/56%
۱۷	چکیده فارسی - همخوانی با متن پایان نامه	18	527	3/30%	0	500	0/00%	3/30%
۱۸	چکیده فارسی - اعراس/فرمولها و روابط ریاضی/شیمی	92	453	16/88%	0	500	0/00%	16/88%

ردیف	نام مشخصه کیفی در مدارک پارسا	نمایه‌سازی			ویراستاری			میزان بهبود
		تعداد ناهمخوانی	تعداد موارد همخوان	درصد ناهمخوانی	تعداد موارد همخوان	تعداد ناهمخوانی	درصد ناهمخوانی	
۱۹	چکیده لاتین - مشکلات ویرایشی	70	461	13/18%	0	500	0/00%	13/18%
۲۰	چکیده لاتین - مشکلات نگارشی جزئی (کمتر از ۳ مورد)	4	536	0/74%	0	500	0/00%	0/74%
۲۱	چکیده لاتین - مشکلات نگارشی (۳ و بیشتر از ۳ مورد)	3		0/55%	0		0/00%	0/55%
۲۲	چکیده لاتین - کلمات اضافی (مانند چکیده، کلیدواژه، اطلاعات دانشجو)	148	395	27/26%	0	500	0/00%	27/26%
۲۳	چکیده لاتین - همخوانی با متن پایان نامه	17	526	3/13%	0	500	0/00%	3/13%
۲۴	چکیده لاتین - اعشار/فرمولها و روابط ریاضی/شیمی	12	531	2/21%	0	500	0/00%	2/21%
۲۵	فهرست مندرجات - خالی بودن فیلد فهرست مندرجات	5	274	0/92%	0	499	0/00%	0/92%
۲۶	فهرست مندرجات - ناقص بودن فهرست مندرجات	252		46/15%	1		0/20%	45/95%
۲۷	فهرست مندرجات - به هم چسبیده بودن کلمات	15		2/75%	0		0/00%	2/75%
۲۸	فهرست منابع فارسی - خالی بودن فیلد فهرست منابع فارسی	12	464	2/21%	0	500	0/00%	2/21%
۲۹	فهرست منابع فارسی - ناقص بودن فهرست منابع فارسی	35		6/45%	0		0/00%	6/45%
۳۰	فهرست منابع فارسی - به هم چسبیده بودن کلمات	12		2/21%	0		0/00%	2/21%
۳۱	فهرست منابع فارسی - درج فهرست لاتین در این فیلد	20		3/68%	0		0/00%	3/68%
۳۲	فهرست منابع لاتین - خالی بودن فیلد فهرست منابع لاتین	8	420	1/47%	0	500	0/00%	1/47%
۳۳	فهرست منابع لاتین - ناقص بودن فهرست منابع لاتین	87		15/99%	0		0/00%	15/99%

ردیف	نام مشخصه کیفی در مدارک پارسا	نمایه سازی			ویراستاری			میزان بهبود
		تعداد ناهمخوانی	تعداد موارد همخوان	درصد ناهمخوانی	تعداد موارد همخوان	تعداد ناهمخوانی	درصد ناهمخوانی	
۳۴	فهرست منابع لاتین - به هم چسبیده بودن کلمات	8		1/47%		0	0/00%	1/47%
۳۵	فهرست منابع لاتین - درج فهرست فارسی در این فیلد	21		3/86%		0	0/00%	3/86%
۳۶	مشخصات ظاهری	کاربرد ندارد			0	500	0/00%	-
۳۷	حوزه موضوعی اول				0	500	0/00%	-
۳۸	حوزه موضوعی دوم				4	496	0/79%	-

فصل چهارم

معرفی روش های نمونه برداری

۴-۱ مقدمه

در فصل‌های گذشته در مورد اهمیت نمونه برداری در فرایندهای نمایه سازی و ویراستاری مدارک پارسا به تفصیل توضیح داده شد. از آنجا که در این فرایندها لازم است از مدارک تولید یا اصلاح شده در مرحله قبل عملیات نمونه برداری انجام شود و خطاهای نمونه برداری موضوعی است که به صورت ویژه در این کار دارای اهمیت است، بنابراین، لازم است این کار برپایه اصول و تکنیک‌های علمی صورت پذیرد. با وجود این که فرایند نمونه برداری امروزه یک علم به حساب می آید و جنبه‌های آماری و محاسباتی فراوانی دارد اما در این فصل به برخی مفاهیم ساده و کاربردی آن که در این پژوهش مورد استفاده قرار گرفته است خواهیم پرداخت.

• نمونه برداری

نمونه برداری برای پذیرش، بخش اصلی و ویژه‌ای از ماهیت کنترل کیفیت را به خود اختصاص می‌دهد. اگرچه معمولاً از نمونه گیری به منظور پذیرش مواد دریافتی (ورودی کالا به سیستم) استفاده می‌شود، ولی موارد دیگری نیز برای استفاده از آن وجود دارد. برای نمونه ممکن است یک تولیدکننده کالای نیمه ساخته‌ای را به مرحله بعدی تولید ارسال نماید، در حالی که محصولات نیمه کاره در این مرحله مردود شده، دوباره کاری روی آنها صورت گرفته و یا دور ریخته شوند. از این رو، نمونه برداری ابزاری مناسب برای کنترل کیفیت یا ارزیابی سطح کیفیت در مراحل گوناگون شکل گیری محصول است.

• بهر یا انباشته (lot & batch):

مجموعه ای از اقلام و یا واحدهای بررسی شده که ویژگی‌های یکسانی دارند.

- **نمونه تصادفی (random sample):**

یک یا چند قلم کالای انتخاب شده از یک بهر که اطلاعات خاصی را در جهت تصمیم گیری درباره بهر و یا جریان تولید در اختیار قرار می دهند.

- **حجم بهر lot size:**

به تعداد اقلام یا کالای موجود در یک بهر گفته می شود.

- **حجم نمونه sample size:**

تعداد اقلام تشکیل دهنده نمونه را حجم نمونه می نامند، یک نمونه ممکن است از یک یا چند قلم کالا تشکیل شده باشد.

نکات ذیل در مورد نمونه برداری به منظور پذیرش از اهمیت ویژه ای برخوردار می باشد:

۱. نمونه برداری یک ابزار بازرسی است که قابل قبول یا مردود بودن یک بهر را تعیین می کند. کنترل یا پایش تغییر کیفیت یک فرآیند از عهده این تکنیک خارج است.
۲. طرح های نمونه برداری به منظور پذیرش (روش های نمونه برداری)، به صورت سراسر است یک ابزار کنترل کیفیت نیستند. بلکه ابزاری برای سنجش دقت تولید محصول هستند.
۳. برای اطمینان از کیفیت محصول در فرآیند تولید می توان به نحو موثری از طرح های نمونه برداری به منظور پذیرش (به عنوان یک ابزار بازرسی) استفاده کرد.

- **رابطه تولید کننده و مصرف کننده**

وقتی از نمونه برداری به منظور پذیرش استفاده می شود، بین منافع مصرف کننده و منافع تولید کننده تضاد ایجاد می شود. از یک سو، تولید کننده می خواهد تمام بهرهای خوب پذیرفته شوند و از سوی

دیگر، مصرف کننده تمایل دارد تمام بهرهای بدرد شوند. در این زمینه اصطلاح ریسک تولید کننده و ریسک مصرف کننده تعریف می شود.

• ریسک مصرف کننده (β)

ریسک مصرف کننده احتمال پذیرش یک بهر بد یا غیر قابل قبول است. مقدار این ریسک معمولاً ۰٫۱ تعیین می شود. در رابطه با ریسک مصرف کننده یک تعریف عددی بهر معیوب وجود داد که LTPD یا درصد رواداری اقلام معیوب بهر نامیده می شود.

نکته: احتمال پذیرش بهر با LTPD درصد معیوب، β درصد خواهد بود.

• ریسک تولید کننده (α)

ریسک تولید کننده (α): احتمال رد شدن یک بهر خوب یا قابل قبول است که اغلب برای این ریسک مقدار ۰٫۰۵ در نظر گرفته می شود. در رابطه با ریسک تولید کننده یک تعریف عددی درصد اقلام معیوب بهر خوب که AQL نامیده می شود، وجود دارد. AQL حداکثر درصد اقلام معیوبی است که جهت نمونه گیری به منظور پذیرش می تواند رضایت بخش باشد. این مقدار یک نقطه مرجع بر روی منحنی OC است و برای القای قابل قبول بودن یک درصد اقلام معیوب به تولید کننده نیست. تنها راه ضمانت پذیرفته شدن بهرها برای تولید کننده آن است که بهرها صفر درصد اقلام معیوب داشته باشند. نکته مهم در مورد ریسک تولید کننده این است که محصولی که درصد اقلام معیوب آن AQL است، α درصد شانس رد شدن خواهد داشت.

۱-۱-۴ معرفی انواع طرح های نمونه گیری برای پذیرش

پیش از معرفی انواع طرح های نمونه برداری لازم است شناخت کافی در مورد انواع مشخصه ها صورت پذیرد. به صورت کلی انواع مشخصه ها به دو دسته کیفی و وصفی طبقه بندی می شوند. مشخصه های

کیفی متغیر مشخصه‌هایی هستند که در مقیاس عددی اندازه‌گیری می‌شوند و مشخصه‌های کیفی وصفی مشخصه‌هایی هستند که بر اساس روش برو - نرو بیان می‌شوند. بازرسی مواد اولیه محصولات نیمه تمام و یا محصولات نهایی بخش مهمی از یک نظام تضمین کیفیت را تشکیل می‌دهند. وقتی بازرسی به منظور پذیرش یا رد محصول با در نظر گرفتن استاندارد خاصی انجام می‌گیرد، روش استفاده شده را نمونه‌گیری جهت پذیرش می‌نامند. فلسفه کلی در اینجا این است که روش‌های نمونه‌گیری جهت پذیرش جایگزین مناسبی برای روش‌های کنترل فرآیند نخواهند بود. در حقیقت، استفاده موفقیت‌آمیز از فنون کنترل فرآیند در مراحل اولیه تولید، نظیر استقرار روش‌های کنترل آماری در تسهیلات تامین‌کننده یا توزیع‌کننده می‌تواند به میزان قابل توجهی نیاز به بازرسی نمونه‌گیری را کاهش و در بعضی موارد نیز از بین ببرد. به طور کلی سه روش برای تصمیم‌گیری در مورد انباشته وجود دارد.

۱. پذیرش بدون بازرسی

۲. بازرسی صد درصد یا به عبارت دیگر، بازرسی تک تک محصولات انباشته و جداسازی محصولات معیوب (محصولات معیوب ممکن است برای تامین‌کننده پس فرستاده شوند، دوباره کاری شوند، با محصولات سالم جایگزین شوند و یا دور ریخته شوند).

۳. نمونه‌گیری جهت پذیرش

پذیرش بدون بازرسی در مواقعی کاربرد دارد که فرآیند تامین‌کننده از کارایی بالایی برخوردار باشد به طوری که هیچ‌گاه محصولات معیوبی تولید نشود و یا اینکه توجه اقتصادی برای شناسایی محصولات معیوب وجود نداشته باشد. بازرسی صد درصد در مواقعی کاربرد دارد که قطعه مورد نظر یک قطعه خیلی بحرانی محسوب می‌شود به طوری که عدم شناسایی محصولات معیوب منجر به هزینه زیاد و غیر قابل قبول در مراحل بعدی می‌گردد.

روش‌های نمونه‌گیری جهت پذیرش در مراحل زیر مورد استفاده قرار می‌گیرند.

- وقتی آزمایش‌ها خراب هستند.
- وقتی که هزینه بازرسی صد درصد خیلی زیاد باشد.

- وقتی که بازرسی صد درصد از نظر فن آوری توجیحی ندارد یا آن قدر زمان نیاز دارد که برنامه تولید مختل می شود.
- وقتی که تعداد محصولاتی که باید بازرسی شوند خیلی زیاد است و میزان خطا نیز آن قدر بالاست که بازرسی صد درصد ممکن است باعث گردد تا تعداد محصولات معیوبی که از روش بازرسی صد درصد پذیرش می شوند بیشتر از محصولاتی باشند که از طریق طرح های نمونه گیری پذیرفته می شوند.
- وقتی که تامین کننده از سابقه عالی در زمینه کیفیت برخوردار است و می خواهیم میزان بازرسی را از صد درصد به سطح کمتری کاهش دهیم ولی نسبت به کارایی فرآیند تامین کننده آن قدر بزرگ نیست که بتوانیم روش بدون بازرسی را اعمال کنیم.
- وقتی که ریسک تعهد و مسئولیت در قبال محصول به طور بالقوه زیاد است و علی رغم رضایت بخش بودن فرآیند تامین کننده هنوز استفاده از یک برنامه نظارت مستمر برای محصول ضروری احساس می گردد.

• تشکیل انباشته

چگونگی تشکیل انباشته می تواند بر موثر بودن طرح نمونه گیری جهت پذیرش اثر داشته باشد چند نکته مهم در مورد تشکیل انباشته ها وجود دارد که باید به آن ها توجه شود. از میان آنها می توان به موارد زیر اشاره کرد:

۱. **انباشته ها باید همگن باشند.** اقلام یک انباشته باید به وسیله یک دستگاه، یک اپراتور، مواد اولیه یکسان و تقریباً به طور همزمان تولید شده باشند. اگر انباشته ها همگن نباشند (نظیر زمانی که خروجی دو خط تولید متفاوت با یکدیگر مخلوط شوند) آنگاه روش های نمونه گیری جهت پذیرش کارایی لازم را نخواهند داشت. انباشته های ناهمگن سبب می شوند تا برنامه ریزی برای تصمیمات اصلاحی جهت حذف منابع تولید محصولات معیوب با مشکل مواجه گردد.
۲. **انباشته های بزرگتر بر انباشته های کوچکتر ترجیح داده می شوند.** از لحاظ اقتصادی به طور معمول بازرسی انباشته های بزرگ بر انباشته های کوچکتر ارجحیت دارد.

۳. انباشته‌ها باید با سیستم‌های جا به جایی مواد در تسهیلات تولید کننده و

مصرف کننده سازگاری داشته باشند. به علاوه، بسته بندی اقلام انباشته ها باید به

گونه‌ای باشد که ریسک‌های مربوط به حمل و نقل و جا به جایی کمینه گردد و از طرف

دیگر انتخاب اقلام نمونه نیز ساده باشد.

۲-۱-۴ مقایسه میان طرح‌های یک، دو و چندبار نمونه برداری

روش‌های گوناگونی برای گروه بندی طرح‌های نمونه گیری جهت پذیرش وجود دارد. یکی از این

گروه بندی ها با توجه به وصفی و یا متغیر بودن مشخصه کیفی مورد نظر انجام می‌شود.

یک طرح یک بار نمونه گیری روشی برای ارزیابی انباشته است. در این روش یک نمونه تصادفی

n تایی انتخاب و بر اساس اطلاعات حاصل از نمونه تصمیم گرفته می شود که انباشته پذیرش یا رد

گردد. به عنوان مثال، یک طرح یک بار نمونه گیری برای مشخصه های کیفی وصفی دارای اندازه

نمونه n و عدد پذیرش c است. از طریق این روش ابتدا یک نمونه n تایی از انباشته به صورت

تصادفی انتخاب می‌شود. اگر تعداد c یا کمتر معیوب در نمونه مشاهده شد، انباشته پذیرش و اگر بیش

از c محصول معیوب در نمونه مشاهده شد، انباشته رد می شود (ایزدی، ۱۳۹۳).

روش دوم، طرح جفت (دو) نمونه گیری است که در شرایط خاصی نمونه دوم انتخاب و سپس تصمیم

گرفته می‌شود که آیا انباشته رد یا پذیرش گردد. در این روش ابتدا یک نمونه تصادفی به اندازه n_1 از

انباشته انتخاب و تعداد اقلام ناهمخوان در نمونه (d_1) شمارش می‌گردد. اگر $d_1 \leq c_1$ باشد، آنگاه

انباشته بر اساس نمونه اول پذیرش می‌شود. اگر $d_1 > c_1$ باشد، آنگاه انباشته بر اساس نمونه اول رد

می‌شود. اگر $c_1 < d_1 \leq c_2$ باشد، نمونه دوم به اندازه n_2 از انباشته انتخاب و تعداد اقلام معیوب

آن d_2 شمارش می‌شود. در این مرحله از مجموع اقلام ناهمخوان در هر نمونه $d_1 + d_2$ جهت نمونه

گیری در مورد انباشته استفاده می‌شود. اگر $d_1 + d_2 \leq c_2$ باشد، انباشته پذیرش و اگر $d_1 + d_2$

$> c_2$ باشد، انباشته رد می‌شود. برتری اصلی طرح جفت نمونه گیری نسبت به طرح یک بار

نمونه گیری امکان کاهش تعداد کل بازرسی های مورد نیاز است (ایزدی، ۱۳۹۳)

روش سوم، طرح چند بار نمونه گیری است. عملکرد این طرح به این صورت است که اگر در هر مرحله از نمونه گیری تعداد اقلام معیوب کمتر یا مساوی عدد پذیرش باشد، انباشته پذیرش و اگر تعداد اقلام معیوب مساوی یا بیشتر از عدد رد باشد، انباشته رد و در غیر این صورت نمونه دیگر انتخاب می شود. در عمل، نمونه گیری پی در پی از لحاظ نظری می تواند آنقدر ادامه یابد تا انباشته صد درصد بازرسی شود. همچنین در عمل فرآیند نمونه گیری، زمانی متوقف می شود که تعداد اقلام بازرسی شده به سه برابر اندازه نمونه ای برسد که در یک طرح یک بار نمونه گیری معادل می توانست استفاده گردد. اگر اندازه نمونه انتخاب شده در هر مرحله بیشتر از یک باشد، قرآیند نمونه گیری را نمونه گیری پی در پی گروهی می نامند. اگر اندازه نمونه ای که در هر مرحله بازرسی می شود برابر یک باشد، فرآیند نمونه گیری، نمونه گیری پی در پی قلم به قلم نامیده می شود (ایزدی، ۱۳۹۳)

برتری اصلی طرح های چند بار نمونه گیری کوچک بودن اندازه نمونه های هر مرحله نسبت به اندازه نمونه های مورد نیاز در طرح های یک بار و جفت نمونه گیری است.

۴-۲ بازرسی اصلاحی

در برنامه نمونه برداری جهت پذیرش وقتی انباشته ها رد می شوند، اقدامات اصلاحی یا چاره جویانه ضرورت پیدا می کند. اقداماتی که در چنین مواقعی انجام می گیرند عبارتند از: بازرسی صد درصد یا کنترل انباشته های رد شده. اگر در زمان انجام این اقدامات با اقلام معیوبی مواجه شدیم آنها را به منظور دوباره کاری یا ارجاع برای تامین کننده از سایر اقلام جدا می سازیم و یا با اقلام خوب جایگزین می کنیم. این برنامه های نمونه گیری را به علت تاثیری که فعالیت های بازرسی بر کیفیت نهایی محصول خروجی دارد، برنامه های بازرسی اصلاحی می نامند. برنامه های بازرسی اصلاحی در مواردی استفاده می شوند که تولید کننده مایل به تعیین سطح متوسط کیفیت در مراحل خاصی از عملیات تولید باشد. بنابراین، بازرسی اصلاحی یا در مرحله دریافت کالا یا مواد، یا در مرحله محصولات نیمه ساخته و یا در مرحله بازرسی نهایی محصولات ساخته شده مورد استفاده قرار

می گیرد. هدف از این روش در داخل سازمان کسب اطمینان از میزان متوسط کیفیت موادی است که در مرحله بعد فعالیت های ساخت استفاده می شود.

معیار متوسط کیفیت خروجی (AOQ^1) برای ارزیابی طرح های نمونه گیری اصلاحی مورد استفاده قرار می گیرد. متوسط کیفیت خروجی سطح کیفیت انباشته که در نتیجه استفاده از بازرسی اصلاحی حاصل می شود را نشان می دهد. با در نظر گرفتن تعریف فوق می توان به راحتی رابطه ای برای محاسبه متوسط کیفیت خروجی (AOQ) تهیه نمود.

فرض کنید اندازه انباشته N است و همه واحد های معیوب با سالم جایگزین می شوند. در این صورت (AOQ) در طرح یک بار نمونه گیری به صورت زیر به دست می آید:

۱. بعد از انجام بازرسی، n واحد که نمونه را تشکیل می دهد به علت جایگزین کردن اقلام معیوب با سالم فاقد معیوب خواهند بود.

۲. اگر انباشته رد شود، $N - n$ واحد فاقد معیوب خواهند بود.

۳. اگر انباشته پذیرش شود، $N - n$ واحد دارای میزان متوسط $P(N - n)$ خواهند بود که P نسبت اقلام معیوب انباشته را نشان می دهد.

بنابراین در مرحله خروجی فرآیند بازرسی تعداد متوسط اقلام معیوب برابر $P_a P(N - n)$ خواهند بود که می توان آن را به صورت متوسط نسبت اقلام معیوبی که توسط کیفیت خروجی نامیده می شود بیان کرد.

$$AOQ = \frac{P_a P(N - n)}{N}$$

که P_a احتمال پذیرش انباشته را نشان می دهد. اگر N نسبت به n بزرگ باشد، آن گاه می توان رابطه (AOQ) را به صورت زیر ارائه نمود:

$$AOQ \cong P_a P$$

¹ Average Outgoing Quality

بیشینه مقدار (AOQ)، بدترین میزان متوسط کیفیت که در نتیجه استفاده از بازرسی اصلاحی حاصل می شود را نشان می دهد. این میزان متوسط کیفیت را حد متوسط کیفیت خروجی (AOQL)^۱ می نامند.

۳-۴ روش نمونه برداری بیچ به بیچ از دیدگاه مبانی احتمال

امروزه با در نظر گرفتن مشخصه های گوناگون کمی و کیفی، انواع بازرسی های مربوطه روی داده ها، مواد اولیه، محصولات در جریان ساخت و محصولات ساخته شده انجام می شود. در این راستا بسیاری از کارشناسان کنترل کیفیت از جدول استاندارد بازرسی استاندارد مانند MIL-STD که توسط ارتش آمریکا تهیه شده و در کتاب های کنترل کیفیت آماری موجود می باشد، استفاده می کنند (Montgomery, 2009). از سوی دیگر متخصصان علم کنترل کیفیت برای پایش مستمر فرایندها روش های امروزی تری مانند استفاده از نمودارهای کنترل را به علت پویایی و توجه به رویکردهای تحلیل ریشه به کار می برند. بازرسی مشخصه های کیفی و کمی برپایه نمونه برداری در قالب طرح های زیر انجام می شود (حمصی و همکاران، ۱۳۸۵).

• طرح MIL-STD-105E

این طرح برای ارزیابی مشخصه های کیفی وصفی مورد استفاده قرار می گیرد. در این حالت، مقدار AQL برای بیشینه درصد ناهمخوانی مشخص می شود. تعیین این مقدار بستگی به سیاست گزار مدیر کیفیت سیستم دارد. پایین در نظر گرفتن مقدار AQL منجر به تایید بهرهایی با تعداد زیاد ناهمخوان شده و در نهایت به سود سازمان نخواهد بود. در خصوص سطح بازرسی که رابطه بین محصول و

^۱ Average Outgoing Quality Limit

حجم نمونه را تعیین می کند این استاندارد دارای سه سطح کلی بازرسی I، II و III می باشد که معمولاً از سطح بازرسی II استفاده می شود. وقتی با حجم نمونه نسبتاً کم سروکار داشته باشیم از سطوح بازرسی ویژه S_1, S_2, S_3, S_4 استفاده می شود. همه طرح های نمونه برداری در سطوح سه گانه ای مطرح بوده که هر کدام دارای جدول ویژه خود هستند. این سه سطح عبارتند از:

- بازرسی نرمال
- بازرسی فشرده
- بازرسی کاهش یافته

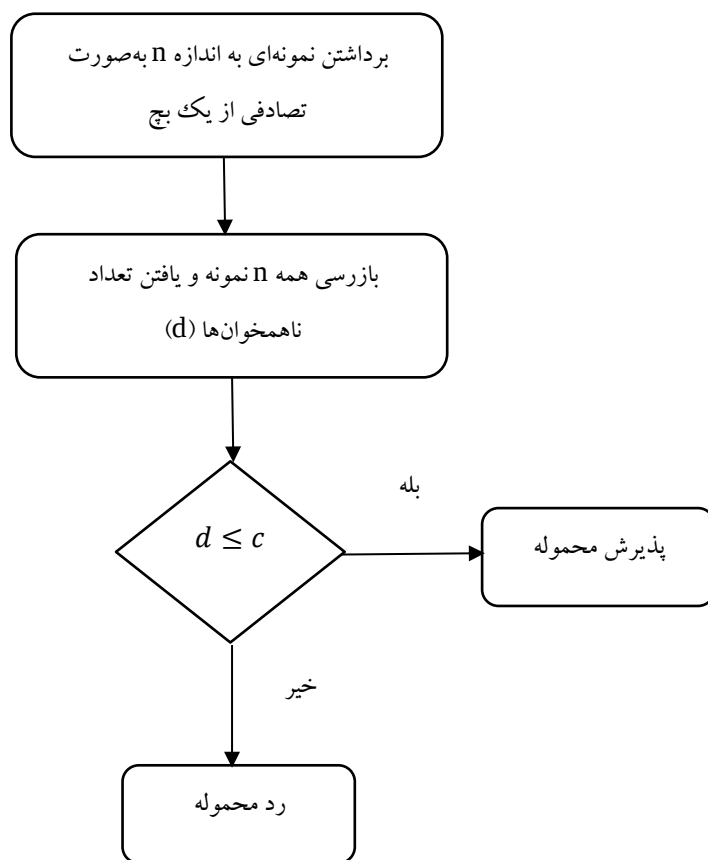
هر طرح نمونه برداری با روش بازرسی با شدت نرمال شروع می شود. اگر از پنج بچ پشت سرهم دو بچ رد شود، بازرسی نرمال جای خود را به بازرسی فشرده می دهد. اگر از 10 بچ بازرسی شده اول هیچ کدام رد نشود، بازرسی نرمال به بازرسی کاهش یافته تغییر می یابد. برای ارزیابی مشخصه های کیفی معمولاً از نمونه برداری یک باره استفاده می شود و در صورت تردید در تصمیم گیری در مورد قبول یا رد شدن بهر از نمونه برداری های دوباره و چندباره استفاده می شود.

• طرح MIL-STD-414

در این حالت مشخصه های کمی نمونه ها مانند وزن، جرم، طول، ضخامت، چگالی اندازه گیری و آزمون می شوند، در صورتیکه در طرح MIL-STD-105، نمونه ها فقط از جنبه همخوانی یا ناهمخوانی مورد بررسی قرار می گیرند. در این دسته از طرح ها، با توجه به اهداف کیفی سازمان، میزان رواداری و به بیان دیگر عدد AQL مشخص می شود. حجم نمونه به کمک جداول ویژه و با توجه به حجم بهر و سطح بازرسی، مشخص می شود. پنج سطح بازرسی I، II، III، IV و V وجود دارد که معمولاً از سطح بازرسی IV استفاده می شود. برای تصمیم گیری مناسب در مورد رد یا قبول بهر، دو روش انحراف معیار و دامنه مورد استفاده قرار می گیرد.

طرح های یک مرحله ای با یک سه تایی (N, n, C) شناخته می شوند. به بیان دیگر، از محموله N تایی تعداد n نمونه بر می داریم و اگر تعداد ناهمخوانها بیشینه برابر C باشد محموله را پذیرفته و اگر تعداد ناهمخوانها از C بیشتر باشد محموله را رد می کنیم.

شکل ۴-۱ موضوع یاد شده را به صورت تصویری نمایش می دهد.



شکل ۴-۱ نمای شماتیک از روش نمونه برداری یک مرحله‌ای

۴-۳-۱-۱ منحنی OC

تابع مشخصه عملکرد (OC') احتمال پذیرش را به عنوان تابعی از θ (در این جا نسبت ناهمخوانی جامعه) نشان می‌دهد. نمودار این تابع به ما کمک می‌کند که احتمال رد یا قبول محموله ای را که نسبت اقلام ناهمخوان خاصی دارد پیدا کنیم.

می‌دانیم اگر X تعداد اقلام معیوب نمونه n تایی باشد که از جامعه ای نامتناهی گرفته شده است (N را خیلی بزرگ در نظر می‌گیریم) آن گاه، احتمال رخداد X از توزیع بینم (دوجمله‌ای) پیروی خواهد کرد. اگر احتمال همخوان بودن یک نمونه را p و احتمال همخوان بودن آنرا q بنامیم، آنگاه احتمال مشاهده X ناهمخوان در نمونه n تایی به شرح زیر خواهد بود.

$$X \sim b(n, p)$$

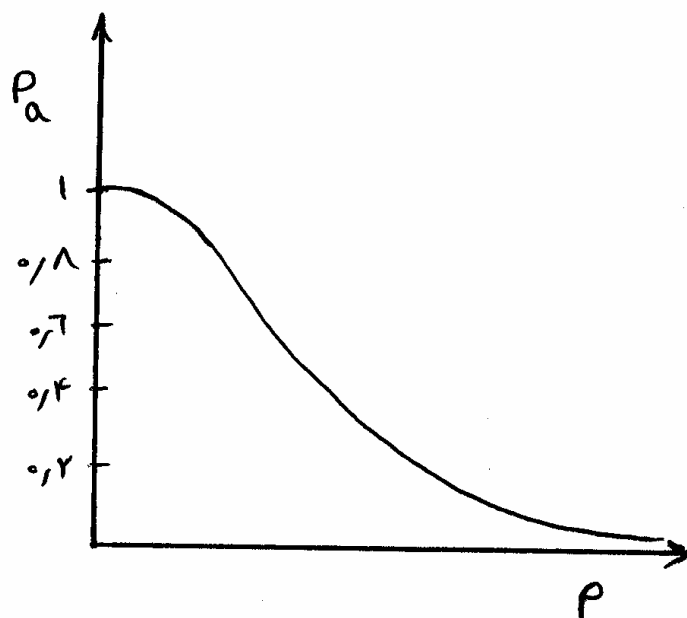
$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

در این صورت احتمال پذیرش عبارتست از:

$$P_a = P(x \leq c) = \sum_{x=0}^c \binom{n}{x} p^x q^{n-x}$$

حال با داشتن مقادیر احتمال پذیرش می‌توان منحنی OC را به‌ازای نسبت‌های ناهمخوانی گوناگون به‌دست آورد. این تابع را می‌توان به عنوان تابعی از p در نظر گرفت یعنی π و c را معلوم فرض نموده P_a را بر حسب p رسم کرد.

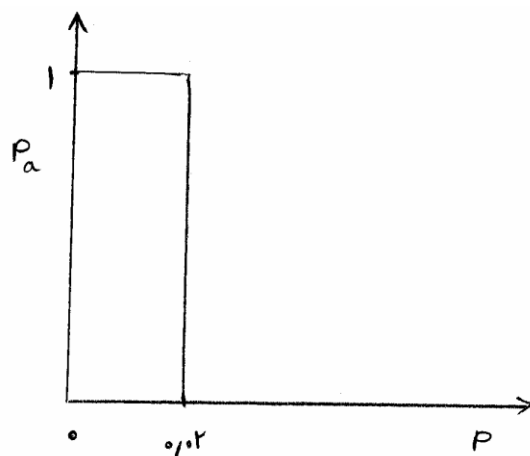
نمودار زیر به عنوان مثال یکی از نمودارهای OC را نشان می‌دهد. این نمودار به‌صورت شماتیک بوده و فقط یک چارچوب کلان از این نمودار را نمایش می‌دهد. همانگونه که در این شکل مشاهده می‌شود هر دو متغیر P و P_a در این نمودار بین صفر و یک بوده و این نمودار، یک نمودار نزولی است. به بیان دیگر، با افزایش مقدار نسبت ناهمخوانی (محور X) احتمال پذیرش کاهش می‌یابد.



شکل ۴-۲ نمای شماتیک از منحنی OC

ارائه طرح نمونه برداری پذیرش در فرایند ویراستاری □ ۶۱

سیستم کیفیت یک سازمان انتظار دارد (هرچند در شرایط واقعی امکان پذیر نیست) منحنی OC در حالت ایده‌ال رفتار کند. در این حالت، بچ‌های با کیفیت بهتر از AQL ۱۰۰ درصد پذیرفته می‌شوند و سایر بچ‌ها با احتمال صفر پذیرفته می‌شوند (با احتمال ۱۰۰ درصد رد می‌شوند). برای نمونه اگر AQL برابر ۰,۰۲ (۲ درصد) باشد شکل ۳-۴ یک منحنی OC ایده‌ال را برای این حالت ارائه می‌کند.



شکل ۳-۴ منحنی OC در حالت ایده‌ال

از سوی دیگر این نمودار فقط وقتی $n=N$ می‌تواند رخ دهد در این حالت بازرسی ۱۰۰٪ انجام می‌شود.

می‌توان نشان داد که هرچه n بزرگتر باشد شیب منحنی OC طوری افزایش می‌یابد که به حالت ایده‌ال بالا نزدیک‌تر می‌شود.

AQL یکی از مهم‌ترین پارامترهای این منحنی است که پایین‌ترین سطح کیفیت ارائه شده را که از دید مصرف‌کننده سطح متوسط کیفیت قابل قبول در نظر گرفته می‌شود نشان می‌دهد.

به AQL سطح کیفیت قابل قبول نیز گفته می‌شود. لازم به ذکر است:

الف) AQL مشخصه فنی و یا مقدار هدف در فرآیند نیست.

ب) AQL ربطی به طرح نمونه‌گیری ندارد و به فرآیند تولید بستگی دارد.

ج) AQL استاندارد است که برای ارزیابی بچ به کار می رود و امیدواریم سطح کیفیت خیلی بهتر از AQL باشد.

یک پارامتر مهم دیگر در نمودار OC، $LTPD^1$ پایین ترین سطح کیفیت را نشان می دهد که مصرف کننده در یک بچ مجاز و قابل قبول می داند. LTPD را نسبت به اقلام ناهمخوان مجاز محموله می نامند و آن را با RQL^2 به معنی سطح کیفیت قابل رد و LQL^3 به معنی سطح کیفیت حدی هم می دانند.

خواص الف و ب در مورد AQL و LTPD هم مناسب می باشد و در مورد خاصیت (ج) باید گفت امید داریم سطح معیوب های محموله از LTPD کمتر باشد.

• ریسک مصرف کننده و تولید کننده

در هر آزمون فرض H_0 و H_1 فرضیه هایی هستند که مدعیان آن ها در مقابل هم قرار دارند یعنی منافع آنها در تضاد است. در نمونه گیری برای پذیرش، هم تولید کننده می خواهد تمام محموله های خوب قبول شوند و مصرف کننده نیز می خواهد تمام محموله های بد را رد کند. تنها در صورتی که بازرسی صد درصد انجام می شود محموله بد را با احتمال ۱ رد و محموله های خوب را با احتمال یک پذیرش می کند.

برای تولید کننده و مصرف کننده دو مقدار شناخته شده و مهم وجود دارند که در ادامه به توضیح آنها خواهیم پرداخت.

الف) ریسک تولید کننده (Producer's risk)

ریسک تولید کننده (که با α هم نمایش داده می شود) احتمال رد کردن نادرست یک محموله است. یعنی یک محموله خوب چقدر احتمال دارد رد شود. این احتمال را با α نشان می دهند و معمولاً عددی بین ۰,۰۱ تا ۰,۱ و بیشتر اوقات ۰,۰۵ در نظر گرفته می شود. از آنجا که محور γ نمودار OC احتمال قبول محموله است، α را نمی توان نشان داد اما $1-\alpha$ را می توان نشان داد.

¹ Lot Tolerance Percent Defective

² Rejectable Quality Level

³ Limiting Quality Level

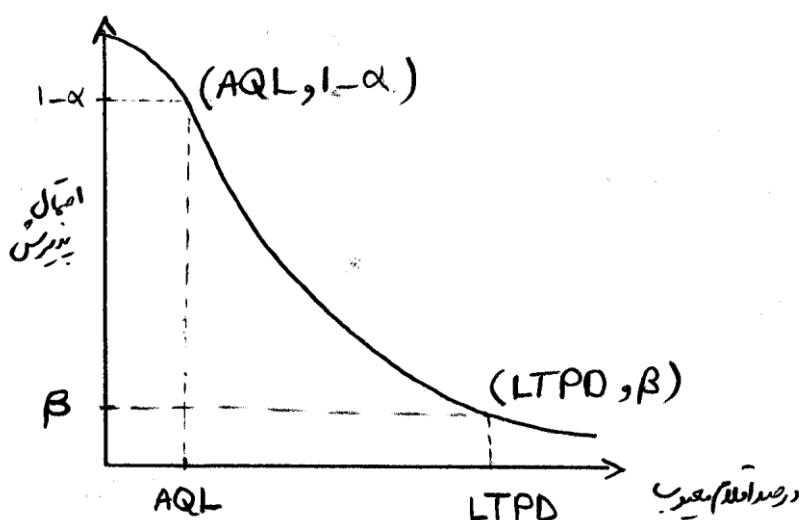
با توجه به تعریف AQL که حداکثر درصد اقلام معیوب رضایت بخش می باشد نقطه (AQL, α) یک نقطه بر روی نمودار OC می باشد.

ب) ریسک مصرف کننده (Consumer's risk)

ریسک مصرف کننده احتمال پذیرش نادرست یک محموله است. یعنی احتمال آن است که محموله ای را که نامناسب است قبول کنیم. ریسک مصرف کننده را با β نشان می دهند و معمولاً آن را با ۰٫۱ تعیین می کنند.

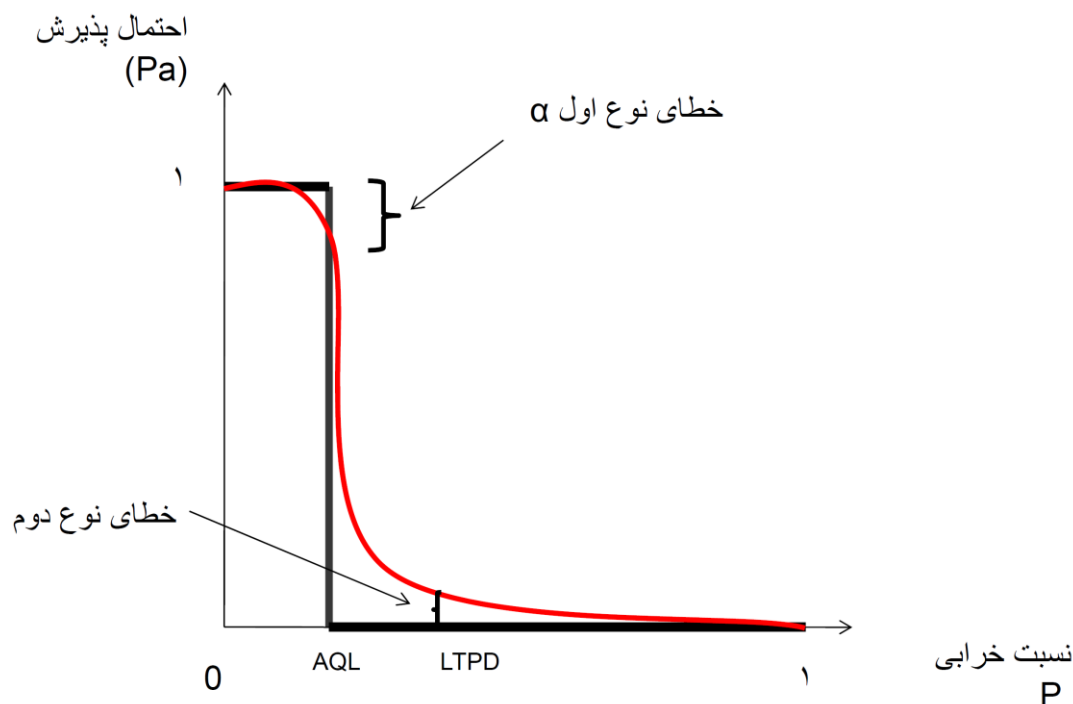
با توجه به تعریف LTPD درصد اقلام معیوب در محموله است که مصرف کننده می خواهد احتمال پذیرش آن کم باشد نقطه $(LTPD, \beta)$ یکی از نقاط نمودار OC است.

در واقع مصرف کننده به ناچار حاضر است با احتمال β محموله ای را که درصد معیوب آن LTPD است قبول کند به نمودار زیر در خصوص دو نقطه $(LTPD, \beta)$ ، $(AQL, 1-\alpha)$ توجه کنید.



شکل ۴-۴ جایگاه ریسکهای مصرف کننده و تولید کننده در منحنی OC

همانگونه که در شکل ۴-۴ قابل مشاهده است، احتمال پذیرش AQL به مراتب بیشتر از احتمال پذیرش LTPD است. در حقیقت، محموله هایی که سطح کیفیت آنها AQL باشند با احتمال $1-\alpha$ پذیرفته می شوند و محموله هایی که سطح کیفیت آن ها LTPD است با احتمال $1-\beta$ رد می شوند.



شکل ۴-۵ مقایسه دو حالت ایده‌آل و واقعی در منحنی OC

در شکل ۴-۵ نیز دو حالت ایده‌آل و واقعی را در یک تصویر به نمایش می‌گذارد. در این نمودار هرچه منحنی واقعی (قرمز رنگ) به حالت ایده‌آل (سیاه‌رنگ) نزدیک شود وضعیت مطلوب‌تری را خواهد داشت. همچنین در این نمودار خطای نوع اول (α) و خطای نوع دوم (β) نیز نمایش داده شده است.

در ادامه و در بخش ۴-۴ روش پیوسته را در نمونه‌برداری معرفی خواهیم کرد.

۴-۴ روش نمونه‌برداری پیوسته (CSP)

در بخش‌های قبلی در مورد بازرسی بچ‌به‌بچ توضیحاتی ارائه شد و مبانی این روش مورد توجه قرار گرفت. در بسیار اوقات واحدهای محصول (یا داده) در زمان‌های تاحدودی طولانی تولید می‌شوند که امکان صبر کردن تا تشکیل یک محموله برای تصمیم‌گیری کمتر فراهم می‌شود. برای نمونه در صنعت جوشکاری ممکن است یک جوشکار در روز تعداد محدودی قطعه را جوشکاری کند و اگر مدیر کیفیت بخواهد برای تصمیم‌گیری تشکیل یک بچ تولیدی را تصور کند برپایه روشهای بچ‌به‌بچ باید تمام موارد را بازرسی کند. در صورت انتخاب راهبرد صبر برای تولید بچ، این کار افزون بر

این که زمانبر است در حالتهایی که محصول باید فرایندهای دیگری را به صورت متوالی طی کند امکان پذیر نیست. اما روش های پیوسته این امکان را فراهم می کنند تا مدیر کیفیت از ابتدا همراه و هم گام با فرایند تولید محصول، گام به گام نمونه برداری را برپایه یک چارچوب مشخص پیش برد و به تولید کننده مشکلات کیفی را هم گوشزد کند. یک طرح نمونه برداری پیوسته (^۱CSP) چارچوبی مشخص در نمونه برداری است که یک متوسط کیفیت خروجی (AOQ) را در بهینه ترین سطح پیچیدگی اجرایی^۲ (که با شاخص میانگین کسر بازرسی شده AFI^3 نشان داده می شود) ارائه می کند (Bebbington, 2003). این روش اطمینان کافی از حداکثر کیفیت خروجی تولید را در یک فرایند پیوسته با حداقل هزینه فراهم می کند. فرض کنید که p (کیفیت ورودی) اقلام تولید شده ناهمخوان یا ناسازگار است. ویژگیهای کیفی اقلام تولید شده معمولاً مستقل فرض می شوند. سپس در این روش یک طرح نمونه برداری اجرا می شود که کیفیت اقلام منتقل شده را با آزمایش موارد و رفع نقص به $AOQ(p)$ (کیفیت متوسط خروجی) بهبود می بخشد. این کار با هزینه ای انجام می شود که توسط میانگین کسر بازرسی شده، AFI ، اندازه گیری می شود. اثربخشی این طرح را می توان با حد متوسط کیفیت خروجی، $AOQL$ ، اندازه گیری کرد. این موضوع در رابطه زیر نمایش داده شده است.

$$AOQL = \max_p AOQ(p)$$

برنامه های نمونه برداری پیوسته برای اولین بار توسط فردی به نام Dodge پیشنهاد شد، و برنامه اولیه CSP-1 نامیده میشود. با شروع این طرح، همه واحدها (محصولات) به صورت ۱۰۰٪ بازرسی می شوند. به محض این که i واحد متوالی محصول (پارامتر روش که از قبل مشخص شده است) عاری از نقص مشاهده شود، بازرسی ۱۰۰٪ متوقف می شود و فقط کسری از واحدها (پارامتر f) بررسی می شود. این واحدهای نمونه هر بار از جریان فرایند انتخاب می شوند. اگر یک واحد انتخاب شده معیوب باشد، بازرسی ۱۰۰٪ از سر گرفته می شود. واحدهای معیوب پیدا شده اصلاح یا دور انداخته می شوند.

وانگ و چانگ (۱۹۸۹) مدل Dodge را با جایگزینی بازرسی ۱۰۰٪ با کسری از نسبت یک به صورت $f_0 < 1$ اصلاح کردند. ایشان اعتقاد داشتند که بازرسی ۱۰۰٪ ممکن است در کارخانه با نرخ تولید بسیار بالا امکان پذیر نباشد. برای آزمایش های پرهزینه، یا مخرب، بازرسی ۱۰۰٪ نیز امکان پذیر نیست. طرح نمونه برداری اصلاح شده به ترتیب دارای دو سطح با کسرهای نمونه f_0 و f_1

^۱ Continous Sampling Plan

^۲ Effort

^۳ Average Fraction Inspected

می باشد. که در اینجا به ترتیب $f_1 < f_0 < 1$ هستند. برپایه پژوهش وانگ و چانگ (۱۹۸۹) می توان نمونه برداری را در روش CSP با کسری از تولیدات آغاز کرد (f_0) و به تدریج در صورت مشاهده وضعیت مطلوب به صورت پیوسته کسر نمونه برداری را کاهش داد (f_1).

ویژگی های کیفی واحدهای تولیدی (مستقل از همخوان یا ناهمخوان بودن آنها با استانداردها) معمولاً مستقل فرض می شوند، فرایند تولید از الگوی توزیع برنولی مستقل یکسان برخوردار است. از آنجا که مجموع چند متغیر دارای توزیع برنولی توزیع دوجمله ای خواهد داشت برای محاسبه احتمالات در این حوزه از این توزیع آماری استفاده خواهد شد.

برای هر طرح نمونه گیری پیوسته، تعیین درست AOQ و همچنین AOQL یک کار ضروری است. در یک رویکرد عمومی کاربرد روش CSP می توان به صورت کلی چارچوب زیر را مورد استفاده قرار داد.

(۱) در ابتدا کسری از واحدها (f_0) بررسی می شود. روند بازرسی ادامه دارد تا زمانی که i واحد پشت سرهم که نمونه برداری شده و بازرسی می شوند همه وضعیت مطلوبی را داشته باشند.

(۲) اگر روند مطلوب بودن نمونه های بازرسی شده تا زمان مشخصی (تعداد i واحد پشت سرهم) ادامه یابد، میزان کسر بازرسی به یک سطح پایین تر (f_1) می رود. به عبارت دیگر، کسر کمتری از تولیدهای انجام شده مورد بازرسی قرار می گیرند.

(۳) اگر یک واحد نمونه در بازرسی انجام شده، ناهمخوان ارزیابی شود، کسر بازرسی شده به کسر ابتدایی (f_0) باز خواهد گشت. این سطح باقی خواهد ماند تا دوباره i واحد پشت سرهم با کیفیت مناسب بازرسی شوند.

(۴) در فرایند نمونه برداری به روش CSP، واحدهای ناهمخوان اصلاح می شوند.

در هر دو مرحله (۱) و (۲) طرح CSP، برای انتخاب نمونه مناسب برای بازرسی، قانون نمونه گیری لازم است مورد توجه قرار گیرد. تصادفی بودن نمونه انتخاب شده (از جنبه های گوناگون مانند زمان، موقعیت مکانی یا زمانی) از مهم ترین بخش های این قانون است. در عمل، برای ساده سازی کاربرد این قانون نمونه برداری، فرکانس نمونه برداری معمولاً متقابل (وارونه) یک عدد صحیح است. علاوه بر این داج (۱۹۴۳) اشاره می کند که برای نمونه گیری های پیوسته، برای ایجاد اطمینان از برآورده شدن یک سطح کیفیت مناسب، نباید از کسر نمونه برداری کمتر از $\frac{1}{50}$ استفاده کرد. از این رو، مدیر کیفیت به برای یافتن کسر نمونه برداری به رابطه $f_1 = \frac{1}{n}$ محدود خواهد شد، جایی که $n \in \{2, 3, \dots, 50\}$ باشد. همچنین، از سوی دیگر از آنجا که در طرح های دو مرحله ای نمونه برداری با

کسر f_0 آغاز می شود نیاز است با توجه به شناخت کافی از وضعیت کیفیت کسر درستی انتخاب شود تا پایداری و قابلیت اطمینان روش نمونه برداری تا حد امکان افزایش یابد.

۴-۵ بررسی ویژگی ها و تمایزهای روش های نمونه برداری

همانگونه که در بخش های قبلی اشاره شد دو روش پایه ای که نمونه برداری وجود دارد و همیشه مورد توجه بوده اند روش های نمونه برداری بچ به بچ و نمونه برداری پیوسته هستند. در ادامه به ویژگی های هر یک خواهیم پرداخت.

- روش پیوسته (CSP) زمانی استفاده می شود که بازرسی به سرعت انجام می شود و نتایج در زمان اندکی حاصل می شود (Banzhaf, R. A., & Brugger, 1975).
- هیچ آزمون مخربی^۱ صورت نمی پذیرد. به بیان دیگر نمونه در حین فرایند نمونه برداری آسیب نمی بیند و یا حتی دچار تغییر نمی شود.
- کیفیت فرایند تا حدود زیادی پایدار و با ثبات است.
- نمونه ها به صورت یک به یک قابل تمایز^۲ هستند.

همانگونه که در ابتدای این فصل توضیح داده شود در روش پیوسته بازرسی در ابتدا بصورت ۱۰۰ درصد خواهد بود و به تدریج در صورت با ثبات و مناسب بودن کیفیت نمونه ها کسری از آنها بررسی خواهد شد. با ادامه فرایند نمونه برداری اگر ثبات در کیفیت ادامه پیدا کند کسر بزرگ و بزرگ تر خواهد شد و در صورت مشاهده نشانه هایی از افت کیفیت بازرسی به روش ۱۰۰٪ باز خواهد گشت. روش دیگری که بسیار بیشتر از CSP مورد توجه است روش نمونه برداری به صورت بچ به بچ است. در این روش براساس وضعیت نمونه برداشته شده کل محموله مورد قضاوت قرار میگیرد. استاندارد ANSI/ASQ z1.4-2003 یکی از استانداردهایی است که برپایه این روش شکل گرفته است. مزایای این روش به شرح زیر است.

- هزینه کمتر به علت بازرسی کمتر
- خسارت کمتر به علت جا به جایی کمتر
- قابل استفاده در آزمایشهای مخرب
- دخالت تعداد کمتری پرسنل در فعالیتهای بازرسی

^۱ Destructive test

^۲ Identical

- کاهش قابل توجهی در میزان خطای بازرسی
- ایجاد انگیزه برای تأمین کننده جهت بهبود کیفیت یا رد کل انباشته به جای فقط ارجاع محصولات

برخی معایب این روش را می توان به صورت زیر برتوصیف کرد (Muñiz Jiménez, 2017).

- وجود ریسک برای پذیرش انباشته های بد و رد انباشته های خوب
- کسب اطلاعات کمتر در مورد محصول و یا فرایندی که آنرا تولید کرده
- نیاز به برنامه ریزی و مکتوب کردن دستورعمل های نمونه گیری جهت پذیرش در مقایسه با عدم نیاز به اینگونه فعالیت ها در بازرسی ۱۰۰ درصد

بررسی مزایا و معایب گفته شده در بالا برای نمونه برداری های بچ به بچ در مقایسه با روش پیوسته از یک سو و همچنین پیچیدگی های اجرایی فراوانی که در روش پیوسته وجود دارد از سوی دیگر لزوم گرایش به روش بچ به بچ را در فرایندهای نمایه سازی و ویراستاری نمایان تر می سازد. همچنین از آنجا که در هر روز کاری تعداد مشخصی (بچ) از مدارک پارسا توسط یک نمایه ساز و یا ویراستار مورد بررسی قرار می گیرد و یکی از اهداف نمونه برداری بررسی عملکرد هر فرد (نمایه ساز/ ویراستار) است لزوم استفاده از روش بچ به بچ نمایان تر خواهد شد.

در ادامه و در فصل پنجم با توجه به اصول روش نمونه برداری که در این فصل معرفی شد یک روش مناسب برای نمونه برداری در فرایندهای نمایه سازی و ویراستاری طراحی شده و نتایج ارائه می شود.

ارائه طرح نمونه برداری پذیرش در فرایند ویراستاری □ ۶۹

فصل پنجم

طراحی روش نمونه برداری برای فرایندهای نمایه سازی و ویراستاری

۵-۱ مقدمه

در فصل چهارم در مورد اصول و مبانی انواع روش های نمونه برداری و شاخص های ارزیابی روش های گوناگون توضیح داده شد. همچنین، در انتهای این فصل روش بچ به بچ به عنوان یک روش مناسب برای ارزیابی داده های پارسا مورد توجه قرار گرفت. پارامترهای اصلی این روش n (تعداد نمونه)، c (عدد پذیرش) و r (عدد رد) معرفی شده و شاخص های خطای نوع I و خطای نوع II برای ارزیابی توانایی هر روش نمونه برداری طراحی شده توصیف گردید. همچنین AQL و LTPD به عنوان پیش فرض های یک روش نمونه برداری تعریف شدند. منحنی OC نیز که یک ابزار کاربردی در طراحی روش های گوناگون نمونه برداری است ارائه شد و بیان شد که هرچه این منحنی برای یک روش به وضعیت ایده آل و مطلوب نزدیکتر باشد طرح نمونه برداری طراحی شده از کیفیت بالاتری برخوردار است. محورهای افقی و عمودی در این روش به ترتیب نسبت ناهمخوانی (عددی بین صفر و یک) و احتمال پذیرش (عددی بین صفر و یک) هستند.

در این فصل بنا داریم برپایه اصول علمی، یک روش نمونه برداری برای داده های مدارک پارسا در فرایندهای نمایه سازی و ویراستاری طراحی کرده و نتایج آن را ارائه کنیم. در ادامه، پس از ارائه چارچوب کلان شیوه دستیابی به یک طرح نمونه برداری مناسب، الگوریتم یافتن طرح نمونه برداری را معرفی خواهیم کرد. سپس، الگوریتم معرفی شده را برای مشخصه های کیفی مختلفی که در فرایندهای نمایه سازی و ویراستاری کنترل می شوند بومی سازی کرده و برای هریک از این مشخصه ها به کار خواهیم بست و نتایج را ارائه خواهیم کرد. از سوی دیگر، از آنجاکه یکی از اهداف این پژوهش تعیین مسئول و پاسخگوی اصلی برای هر قلم اطلاعاتی ثبت و پردازش شده در سامانه ثبت است در یک بخش جداگانه مسئول اصلی هر قلم اطلاعاتی را معرفی خواهیم کرد.

از آنجاکه این پژوهش کاربردی بوده و نتایج آن در واحد مدیریت سازماندهی و تحلیل به کار برده خواهد شد برای ارزیابی میزان کاربردی بودن نتایج کارگروه کیفیت پارساها تشکیل شد و نتایج در آن کارگروه ارائه شد و پس از دریافت بازخوردها در گزارش نهایی ثبت شد. اعضای این کارگروه شامل مدیر واحد سازماندهی و تحلیل اطلاعات، مدیر واحد ثبت و فراهم آوری و همچنین معاون اطلاعات علم و فناوری ایران است. در این کارگروه برپایه نتایج تحلیل، تعداد نمونه (n) و عدد پذیرش (c) برای هر قلم اطلاعاتی و نیز مسئول اصلی کیفیت (کارشناس واحد فراهم آوری/نمایه ساز/ویراستار) تعیین شد.

در ادامه این فصل، و در بخش بعدی چارچوب کلان مراحل یافتن طرح نمونه برداری مناسب ارائه خواهد شد.

۵-۲ چارچوب کلان مراحل یافتن طرح نمونه برداری مناسب

در فصل چهارم نشان داده شد که برای یافتن یک طرح نمونه برداری مناسب پیش از یافتن متغیرهای بهینه لازم است برخی پارامترهای کلیدی مشخص شوند. در ادامه به این پارامترها اشاره خواهیم کرد.

- پارامتر سطح کیفیت قابل قبول (AQL): سطحی از نسبت ناهمخوانی که برای یک مشخصه کیفی قابل قبول است. برای نمونه برای قلم اطلاعاتی نام پدیدآور، ناهمخوانی با فایل پارسا، یک ناهمخوانی اصلی است و سطح AQL کمی قابل قبول است. برای نمونه عدد ۰,۰۰۵ نشان دهنده این است که در یک حجم پارسای ثبت شده، از هر ۱۰۰۰ نمونه بیشینه ۵ نمونه می تواند دارای این ناهمخوانی باشد و تمایل نداریم بیش از این ناهمخوانی را مشاهده کنیم.
- پارامتر سطح کیفیت قابل رد (LTPD): سطحی از نسبت ناهمخوانی که در صورتیکه این نسبت ناهمخوانی را در یک نمونه مشاهده کنیم تمایل داریم آن را به مرحله قبل بازگشت دهیم. برای نمونه عدد ۰,۱ نشان دهنده این است که در یک نمونه کنترل شده از پارساها اگر ۱۰ درصد ناهمخوانی مشاهده شود آن نمونه باید برگشت داده شود.
- خطای نوع I: در چارچوب ارائه شده در این بخش این خطا میزان احتمال بازگرداندن انباشته ای از پارسا است که دارای سطح AQL است. در منابع علمی خطای نوع I را برابر ۰,۰۵ در نظر می گیرند (Montgomery, 2009).
- خطای نوع II: در چارچوب ارائه شده در این بخش این خطا میزان احتمال پذیرش انباشته ای از پارسا است که دارای سطح LTPD است. در منابع علمی خطای نوع II را برابر ۰,۱ در نظر می گیرند (Montgomery, 2009).

با در نظر گرفتن این پارامترها و مقداردهی به آنها می توان چارچوب کلان مراحل یافتن یک طرح نمونه برداری را ارائه کرد. در این بخش و در قالب ۷ گام مراحل یادشده معرفی خواهند شد (شکل ۵-۱).



شکل ۱-۵ گام های اصلی یافتن طرح نمونه برداری در فرایندهای نمایه سازی و ویراستاری

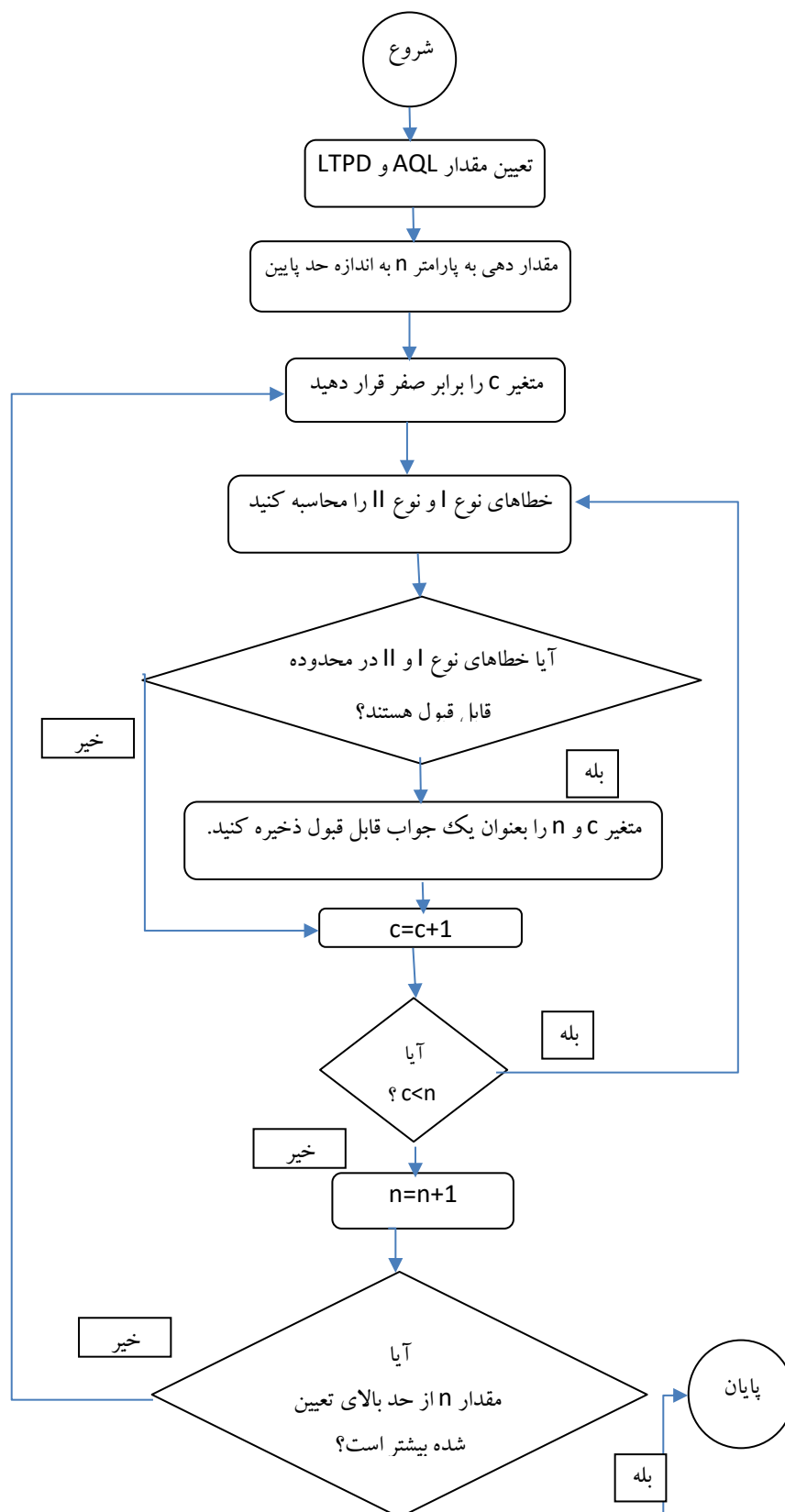
در ادامه و در بخش ۳-۵ الگوریتم اشاره شده در گام ششم از شکل ۱-۵ معرفی خواهد شد.

۳-۵ طراحی الگوریتم دستیابی به پارامترهای بهینه طرح نمونه برداری

در شکل ۲-۵ به گام‌های الگوریتمی اشاره شده است که به تعیین پارامترهای بهینه طرح نمونه برداری می‌انجامد. گامهای این الگوریتم برپایه روش مک‌وليامز^۱ و همکاران (۲۰۰۱) توسعه داده شده است.

همانگونه که پیش از این در بخش ۲-۵ توضیح داده شد پیش از اجرای این الگوریتم لازم است برخی پارامترها مقداردهی شوند که این کار برپایه بازخوردهای دریافت شده و گزارش‌های تهیه شده در فصل سوم و برپایه سطح کیفیت مورد انتظار صورت خواهد پذیرفت. به بیان دیگر، پس از مشخص شده AQL و LTPD از یک سو و مقداردهی خطای نوع I و خطای نوع II به ترتیب ۰,۰۵ و ۰,۱ از سوی دیگر شروع به اجرای الگوریتم خواهیم کرد. لازم به ذکر است همانگونه که در شکل ۲-۵ مشاهده می‌شود باید در هر مرحله پس از تنظیم مقادیر n و c خطاهای نوع I و نوع II حساب شوند برای این کار برپایه توضیحات ارائه شده در فصل قبل از توزیع بینم استفاده خواهد شد.

¹ McWilliams



شکل ۵-۲ الگوریتم تعیین پارامترهای بهینه طرح نمونه‌برداری

پیش از اجرای الگوریتم ارائه شده در شکل ۵-۲ موارد زیر به عنوان پارامترهای مساله تعیین تعداد نمونه باید مقداردهی شوند.

- **مقادیر پارامترهای AQL و LTPD:** این پارامترها برپایه نوع مشخصه کیفی (A، B یا C) در جدول ۵-۱ مقداردهی خواهند شد.
- **خطاهای نوع I و II:** مقادیر این خطاها همانگونه که پیش از این توضیح داده شد به ترتیب ۰,۰۵ و ۰,۱ تعیین شدند.
- **تعیین حدود قابل قبول برای پارامتر n (تعداد نمونه):** این حدود برپایه تعداد کار انجام شده (مدرک کنترل شده) در هر بخش تعیین خواهد شد (بخش ۴-۵).

پس از مقداردهی به پارامترهای بالا می توان الگوریتم های تعیین تعداد نمونه را اجرا نمود. این نکته گفتنی است که در پایان اجرای این الگوریتم جواب های گوناگونی به دست خواهد آمد. در نهایت پاسخی انتخاب خواهد شد که نزدیک ترین فاصله با خطاهای نوع اول و دوم از پیش تنظیم شده را داشته باشد.

۵-۴ ارائه طرح نمونه برداری برای فرایندهای نمایه سازی و ویراستاری

همانگونه که در بخش ۴-۴ توضیح داده شد، برای تعیین پارامترهای روش نمونه برداری افزون بر مقداردهی به پارامتر خطای نوع I و خطای نوع II لازم است مقادیر AQL و LTPD و همچنین حدود پارامتر n مشخص شود. براین اساس در جداول ۵-۱ و ۵-۲ به ترتیب مقادیر AQL و LTPD و همچنین حدود پارامتر n ارائه شده اند.

در جدول ۵-۱ پارامترهای AQL و LTPD بر مبنای میزان اهمیت هر ناهمخوانی تعیین شدند. مقادیر A، B و C به ترتیب تعاریف زیر را دارند.

- **مشخصه نوع A**
 - این ناهمخوانی بحرانی بوده و به صحت پاراسا مرتبط است. همچنین ممکن است در نمایش یک داده بی کیفیت که تاثیر گذاری شدید بر دیدگاه کاربران دارد در سامانه گنج منجر شود.

• مشخصه نوع B

○ این ناهمخوانی مهم بوده و به دقت پارسا مرتبط است. همچنین ممکن است در نمایش یک داده بی کیفیت که تاثیر گذاری شدید بر دیدگاه کاربران دارد در سامانه گنج منجر شود.

• مشخصه نوع C

○ این ناهمخوانی جزئی است و یا تاثیر زیادی در رضایت کاربران در سامانه گنج ندارد.

جدول ۵-۱ مقاردهی به پارامترهای AQL و LTPD بر پایه نوع مشخصه کیفی

ردیف	نام قلم اطلاعاتی	دسته ناهمخوانی	نوع مشخصه کیفی	AQL	LTPD
۱	عنوان فارسی	مشکل نگارشی	A	0.005	0.05
		مشکل اصلی (صحت) ناهمخوانی با متن و ...)	A	0.005	0.05
۲	عنوان لاتین	مشکل ویرایشی / نگارشی	B	0.01	0.08
		مشکل اصلی (صحت) ناهمخوانی با متن و ...)	B	0.01	0.08
۳	پدیدآوران	دانشگاه	A	0.005	0.05
		دانشکده	B	0.01	0.08
		دانشجو	A	0.005	0.05
		اساتید راهنما	A	0.005	0.05
		اساتید مشاور	A	0.005	0.05
۴	تاریخ دفاع	-	C	0.03	0.1
۵	رشته	-	B	0.01	0.08
۶	گرایش	-	B	0.01	0.08

ارائه طرح نمونه برداری پذیرش در فرایند ویراستاری □ ۷۸

ردیف	نام قلم اطلاعاتی	دسته ناهمخوانی	نوع مشخصه کیفی	AQL	LTPD
۷	چکیده فارسی	وجود کلمات اضافی (مانند کلمه چکیده در ابتدا)	A	0.005	0.05
		ناهمخوانی با متن پایان نامه	A	0.005	0.05
۸	چکیده لاتین	وجود کلمات اضافی (مانند کلمه چکیده در ابتدا)	B	0.01	0.08
		ناهمخوانی با متن پایان نامه	B	0.01	0.08
۹	فهرست مندرجات	–	C	0.03	0.1
۱۰	فهرست منابع فارسی	–	C	0.03	0.1
۱۱	فهرست منابع لاتین	–	C	0.03	0.1

در ادامه حدود قابل قبول برای پارامتر n به‌ازای هریک از مراحل که در آن نمونه‌برداری رخ می‌دهد، تعیین خواهد شد.

جدول ۵-۲ تعیین حدود قابل قبول برای تعداد نمونه

ردیف	نام مرحله	خط مشی عملکرد مرحله	مرحله کنترل‌کننده	بازه زمانی عملکرد (میزان انباشته برای نمونه‌برداری)	حدود n
۱	ثبت و فراهم‌آوری اطلاعات	روزانه ۷۰ مدرک	نمایه‌سازی	۱۰ روز کاری (دو هفته)	از ۱ تا ۷۰۰
۲	نمایه‌سازی	روزانه ۱۶,۱	ویراستاری	۱۵ روز کاری (سه)	از ۱ تا ۲۴۲

ردیف	نام مرحله	خط مشی عملکرد مرحله	مرحله کنترل کننده	بازه زمانی عملکرد (میزان انباشته برای نمونه‌برداری)	حدود n
		مدرک		هفته	
۳	ویراستاری	روزانه ۷۰ مدرک	مدیر واحد سازماندهی و تحلیل اطلاعات	۱۰ روز کاری (یک هفته)	از ۱ تا ۷۰۰

اکنون با داشتن همه پارامترهای پیشین یک طرح نمونه‌برداری می‌توان الگوریتم شکل ۵-۲ را پیاده‌سازی کرد. به بیان دیگر، در هریک از مراحل سه‌گانه جدول ۵-۲ و به‌ازای هر نوع مشخصه کیفی یک طرح نمونه‌برداری به‌دست خواهد آمد. در ادامه و در جدول ۵-۳ نتایج طرح نمونه‌برداری قابل مشاهده خواهد بود.

جدول ۵-۳ طرح نمونه‌برداری بهینه برای مراحل گوناگون ثبت و سازماندهی اطلاعات

ردیف	نام مرحله	مرحله کنترل کننده	نوع مشخصه	AQL	LTPD	n	c	r
۱	ثبت و فراهم‌آوری اطلاعات	نمایه‌سازی	A	0.005	0.05	105	2	3
			B	0.01	0.08	65	2	3
			C	0.03	0.1	116	7	8
۲	نمایه‌سازی	ویراستاری	A	0.005	0.05	105	2	3
			B	0.01	0.08	65	2	3
			C	0.03	0.1	116	7	8
۳	ویراستاری	مدیر واحد سازماندهی و تحلیل اطلاعات	A	0.005	0.05	105	2	3
			B	0.01	0.08	65	2	3
			C	0.03	0.1	116	7	8

همانگونه که از جدول ۵-۳ قابل مشاهده است در هریک از مراحل گوناگون به‌ازای هر گروه از مشخصه‌های کیفی طرح نمونه‌برداری تعیین شده است. برای نمونه، در مرحله نمایه‌سازی، کارشناس نمایه‌ساز از ۱۰ روز کاری پیشین یک کارشناس ثبت و فراهم‌آوری اطلاعات ۱۰۵ نمونه انتخاب خواهد کرد. در صورت مشاهده بیشینه ۲ ناهمخوانی در هر مشخصه نوع A، حجم کار انجام شده تایید خواهد شد و در صورت مشاهده ۳ یا بیش از ۳ ناهمخوانی کار مورد تایید نخواهد بود. عدددهای

ارائه طرح نمونه برداری پذیرش در فرایند ویراستاری □ ۸۰

پذیرش و رد در مشخصه های نوع B، به ترتیب ۲ و ۳ و در مشخصه های نوع C، به ترتیب ۷ و ۸ خواهد بود.

در ادامه و در بخش ۵-۵ مسئول اصلی (ثبت یا نمایه سازی) در کنترل کیفیت اقلام سامانه ثبت پارسا ارائه خواهد شد.

۵-۵ تعیین مسئول اصلی در فرایند کنترل اقلام اطلاعاتی پارسا

در این بخش برپایه نشست های برگزار شده توسط کارگروه کیفیت پارسا، قلم های اطلاعاتی گوناگون سامانه ثبت مورد بررسی و تحلیل قرار گرفت و مسئول اصلی در اقلام اطلاعاتی گوناگون تعیین شد. نتیجه این نشست ها در جدول ۴-۵ مشاهده می شود.

جدول ۴-۵ شناسایی مسئول اصلی اقلام اطلاعاتی سامانه ثبت

ردیف	نام قلم اطلاعاتی	مسئول اصلی		کنترل ها در واحد مدیریت ثبت و فراهم آوری اطلاعات	کنترل ها در واحد نمایه سازی
		کارشناس مدیریت ثبت	نمایه ساز		
۱	نام	■		• ناهمخوانی در مواردی مانند ال، حروف اضافه در نام نویسنده • حروف نام نویسنده به درستی درج نشده است • ناهمخوانی مواردی مانند سید، سیده، میر در ابتدای نام نویسنده با صفحه عنوان • درج شدن نام به صورت لاتین	• کنترل نام به صورت نمونه برداری (تصادفی)
۲	نام خانوادگی	■		• ناهمخوانی در مواردی مانند ال، حروف اضافه در نام خانوادگی نویسنده • حروف نام خانوادگی نویسنده به درستی درج نشده است • ناهمخوانی مواردی مانند سید، سیده، میر در ابتدای نام خانوادگی نویسنده با صفحه عنوان • درج شدن نام خانوادگی به صورت	• کنترل نام خانوادگی به صورت نمونه برداری (تصادفی)

ردیف	نام قلم اطلاعاتی	مسئول اصلی		کنترل‌ها در واحد مدیریت ثبت و فراهم آوری اطلاعات	کنترل‌ها در واحد نمایه سازی
		کارشناس مدیریت ثبت	نمایه ساز		
				لاتین • رعایت نکردن نیم فاصله، در صورتیکه پسوند معنادار نباشد (مانند پور، زاده و ...) و فاصله داشتن اگر معنادار باشد.	
۳	مقطع تحصیلی	■		• ناهمخوانی مقطع تحصیلی ثبت شده در سامانه ثبت با صفحه عنوان	• کنترل مقطع تحصیلی به صورت نمونه برداری (تصادفی)
۴	وابستگی سازمانی	■		• ناهمخوانی وابستگی سازمانی ثبت شده در سامانه ثبت با صفحه عنوان	• کنترل وابستگی سازمانی به صورت نمونه برداری (تصادفی)
۵	موسسه آموزشی/پژوهشی	■		• "دانشگاههای وابسته به حوزه علمیه" انتخاب شده است در حالیکه پایان نامه مربوط به حوزه علمیه نیست. • ناهمخوانی میان نام موسسه انتخاب شده با صفحه عنوان. • در دانشگاههای پیام نور، استان و مرکز مشخص نشده است. • در دانشگاههای پیام نور، استان و مرکز مشخص نشده است. • در دانشگاههای آزاد، واحد مشخص نشده است.	• به روزرسانی پایگاه داده اصلی موسسه های آموزشی/ پژوهشی • کنترل موسسه آموزشی/ پژوهشی به صورت تصادفی/ نمونه برداری
۶	دانشکده/پژوهشکده	■		• نام دانشکده/ پژوهشکده بصورت دستی توسط دانشجو در سامانه ثبت وارد شده است، در حالیکه در لیست کرکره ای موجود است. • ناهمخوانی میان نام دانشکده/ پژوهشکده و صفحه عنوان پایان نامه/ رساله	• به روزرسانی پایگاه داده اصلی فهرست دانشکده/ پژوهشکده • کنترل دانشکده/ پژوهشکده به صورت تصادفی/ نمونه برداری
۷	زبان اصلی پایان نامه	■		• هرگونه ناهمخوانی میان زبان انتخاب شده و متن نوشته شده در پایان نامه/ رساله	-
۸	عنوان فارسی (صحت عنوان)	■		• هرگونه ناهمخوانی (کلمات عنوان) میان فیلد عنوان و صفحه عنوان فارسی پایان نامه • وجود غلط های املایی در متن عنوان	• کنترل موارد نگارشی به صورت نمونه برداری/ تصادفی

ردیف	نام قلم اطلاعاتی	مسئول اصلی		کنترل‌ها در واحد مدیریت ثبت و فراهم‌آوری اطلاعات	کنترل‌ها در واحد نمایه‌سازی
		کارشناس مدیریت ثبت	نمایه‌ساز		
۹	عنوان فارسی (دقیق بودن عنوان)	■		-	• اصلاح موارد ویرایشی
۱۰	عنوان انگلیسی (صحت عنوان)	■		• هرگونه ناهمخوانی (کلمات عنوان) میان فیلد عنوان و صفحه عنوان انگلیسی پایان‌نامه • وجود غلط‌های املائی در متن عنوان	• کنترل موارد نگارشی به‌صورت نمونه‌برداری/ تصادفی
۱۱	عنوان انگلیسی (دقیق بودن عنوان)	■		-	• اصلاح موارد ویرایشی
۱۲	رشته / گرایش	■		• ناهمخوانی با نام رشته / گرایش مندرج در صفحه عنوان پایان‌نامه / رساله • رشته بصورت دستی توسط دانشجو در سامانه ثبت وارد شده است، درحالی‌که در لیست کرکره‌ای موجود است.	• به‌روزرسانی پایگاه داده اصلی فهرست رشته / گرایشها • کنترل رشته / گرایش به‌صورت تصادفی/ نمونه‌برداری
۱۳	تاریخ دفاع	■		• ناهمخوانی با صورتجلسه دفاع (در صورت موجود بودن)	-
۱۴	نام استادراهنما/ استاد مشاور	■		• حروف نام استادراهنما/ مشاور به درستی درج نشده است. (درستی حروف، مواردی مانند ال و حروف اضافه) • ناهمخوانی مواردی مانند سید، سیده، میر در ابتدای نام استادراهنما/ استاد مشاور با صفحه عنوان • درج مواردی مانند مهندس، دکتر، حجه الاسلام • درج نام به صورت لاتین	• کنترل و اصلاح موارد ویرایشی/ نگارشی نام استادراهنما/ مشاور به‌صورت تصادفی/ نمونه‌برداری • به‌روزرسانی پایگاه داده اصلی فهرست نام‌ها
۱۵	نام خانوادگی استادراهنما/ استاد مشاور	■		• درست بودن «نام خانوادگی استاد راهنما/ استاد مشاور» در سامانه (پیشوند/ پسوند/ نیم‌فاصله/ اشتباه در تایپ). • درج نشدن نام خانوادگی به صورت لاتین. • درست بودن نیم‌فاصله در صورتیکه پسوند معنادار نباشد (مانند پور، زاده و ...) و فاصله داشتن اگر معنادار باشد. • نبود مواردی مانند درج استادراهنمای ۲،	• کنترل و اصلاح موارد ویرایشی/ نگارشی نام خانوادگی استادراهنما/ مشاور به‌صورت تصادفی/ نمونه‌برداری

ردیف	نام قلم اطلاعاتی	مسئول اصلی		کنترل‌ها در واحد مدیریت ثبت و فراهم‌آوری اطلاعات	کنترل‌ها در واحد نمایه‌سازی
		کارشناس مدیریت ثبت	نمایه‌ساز		
				استاد مشاور ۲ در فراداده	
۱۶	نام/ نام خانوادگی انگلیسی	■		- درست نوشته شدن نام/ نام خانوادگی انگلیسی (پیشوند/ پسوند/ نیم‌فاصله/ اشتباه در تایپ). - فارسی نوشته نشدن نام/ نام خانوادگی انگلیسی	کنترل و اصلاح موارد ویرایشی/ نگارشی نام/خانوادگی انگلیسی به صورت تصادفی/ نمونه برداری
۱۷	چکیده فارسی (کنترل موارد نگارشی صوری)	■		- درج نشدن کلماتی نظیر چکیده و خلاصه به اول چکیده یا کلید واژه‌ها در انتهای چکیده - به هم چسبیده نبودن کلمات - فارسی بودن زبان چکیده - درج نشدن ایمیل یا شماره تلفن	کنترل و اصلاح موارد صوری چکیده فارسی به صورت تصادفی/ نمونه برداری
۱۸	چکیده فارسی (کنترل محتوا)	■		همخوانی حجم چکیده در فراداده با متن پارسا	کنترل ناهمخوانی‌های نگارشی در چکیده
۱۹	چکیده انگلیسی	■		- درج نشدن کلماتی نظیر چکیده و خلاصه به اول چکیده یا کلید واژه‌ها در انتهای چکیده - به هم چسبیده نبودن کلمات - فارسی بودن زبان چکیده - درج نشدن ایمیل یا شماره تلفن	- همخوانی با عنوان فارسی - کنترل و اصلاح موارد صوری چکیده انگلیسی به صورت تصادفی/ نمونه برداری
۲۰	فهرست محتوا/ فهرست منابع فارسی/ فهرست منابع لاتین	■		- به هم ریخته نبودن ظاهر و ساختار - همخوانی صوری حجم در فراداده با محتوا	کنترل و اصلاح موارد صوری فهرستها به صورت تصادفی/ نمونه برداری
۲۱	کلیدواژه‌ها (کنترل موارد نگارشی صوری)	■		- ناهمخوانی با کلیدواژه موجود در متن - تکرار شدن یک کلید واژه در چند فیلد - ناهمخوانی کلید واژه فارسی و لاتین - جابه‌جا وارد شدن کلید واژه (انگلیسی به جای فارسی یا بالعکس) - فینگیلیش نوشته نشدن کلید واژه انگلیسی بر مبنای کلیدواژه فارسی - در یک فیلد بیش از یک کلید واژه وارد شده باشد.	کنترل و اصلاح موارد صوری کلیدواژه‌ها به صورت تصادفی/ نمونه برداری
۲۲	کلیدواژه‌ها (کنترل و اصلاح محتوا)	■		-	افزودن کلیدواژه از اصطلاحنامه

ردیف	نام قلم اطلاعاتی	مسئول اصلی		کنترل‌ها در واحد مدیریت ثبت و فراهم‌آوری اطلاعات	کنترل‌ها در واحد نمایه‌سازی
		کارشناس مدیریت ثبت	نمایه‌ساز		
				- پیشنهاد کلیدواژه کمکی - درج منبع برای کلیدواژه کمکی	
۲۳	حوزه موضوعی		▪	-	این قلم اطلاعاتی در سامانه ویرایش توسط نمایه‌ساز درج می‌شود.
۲۴	فایل‌های پی.دی.اف و ورد پارسا		▪	کنترل فایل‌های پی.دی.اف و ورد برپایه طرح کیفیت پارسا	-

۵-۶ منحنی OC برای طرح‌های نمونه برداری

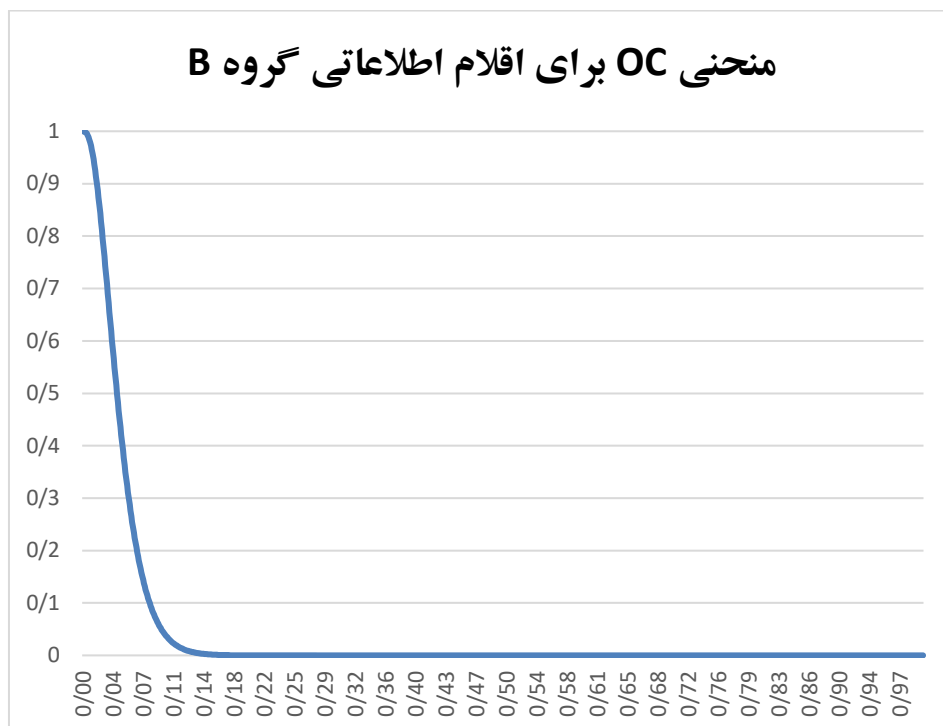
در این بخش منحنی OC برای سه دسته اصلی A، B و C طراحی شده در فرایند نمونه برداری ارائه شده است. در ادامه و به ترتیب در شکل‌های ۵-۳ تا ۵-۵ منحنی OC را برای گروه‌های مشخصه A تا C می‌توان مشاهده کرد.

همانگونه که در شکل‌های ۵-۳ تا ۵-۵ دیده می‌شود منحنی OC شیب مطلوبی دارد یعنی در فاصله بین AQL تا LTPD این منحنی با شیب مناسبی و به سرعت به پایین حرکت می‌کند و این نشان می‌دهد که طرح نمونه برداری ارائه شده می‌تواند بچ‌های مناسب را از بچ‌های نامناسب تمیز دهد.



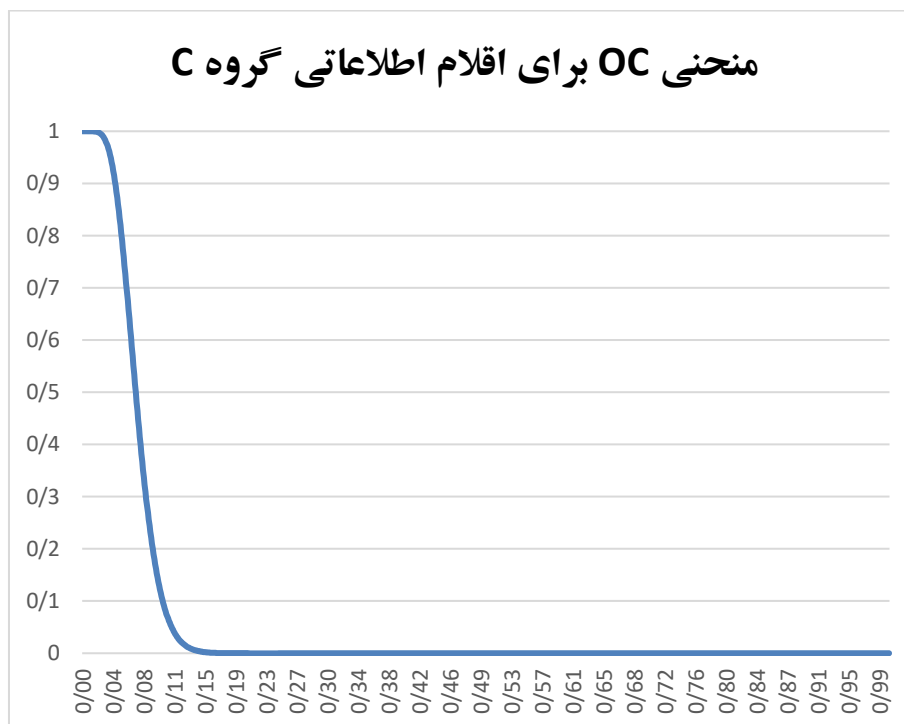
شکل ۳-۵ منحنی OC برای طرح نمونه برداری مشخصه های نوع A

در این بخش عنوان شد که مقادیر AQL و LTPD برای مشخصه های کیفی گروه A به ترتیب عبارتند از ۰,۰۰۵ و ۰,۰۵. همانگونه که در منحنی OC بالا می بینید برای مقادیر کمتر از ۰,۰۰۵ احتمال پذیرش مدارک ارائه شده تا میزان قابل ملاحظه ای بالاست. از سوی دیگر، برای مقادیر بیش از ۰,۰۵ احتمال پذیرش به صفر میل می کند. این نشان می دهد که طرح نمونه برداری ارائه شده برای اقلام اطلاعاتی و مشخصه های کیفی گروه A از اعتبار لازم برخوردار است.



شکل ۵-۴ منحنی OC برای طرح نمونه برداری مشخصه های نوع B

همچنین مقادیر AQL و LTPD برای مشخصه های کیفی گروه B به ترتیب عبارتند از ۰,۰۱ و ۰,۰۸. همانگونه که در منحنی OC بالا می بینید برای مقادیر کمتر از ۰,۰۱ احتمال پذیرش مدارک ارائه شده تا میزان قابل ملاحظه ای بالاست. از سوی دیگر، برای مقادیر بیش از ۰,۰۸ احتمال پذیرش به صفر میل می کند. این نشان می دهد که طرح نمونه برداری ارائه شده برای اقلام اطلاعاتی و مشخصه های کیفی گروه B از اعتبار لازم برخوردار است.



شکل ۵-۵ منحنی OC برای طرح نمونه برداری مشخصه های نوع C

همچنین مقادیر AQL و LTPD برای مشخصه های کیفی گروه C به ترتیب عبارتند از ۰,۳ و ۰,۱. همانگونه که در منحنی OC بالا می بینید برای مقادیر کمتر از ۰,۳ احتمال پذیرش مدارک ارائه شده تا میزان قابل ملاحظه ای بالاست. از سوی دیگر، برای مقادیر بیش از ۰,۱ احتمال پذیرش به صفر میل می کند. این نشان می دهد که طرح نمونه برداری ارائه شده برای اقلام اطلاعاتی و مشخصه های کیفی گروه C از اعتبار لازم برخوردار است.

فصل ششم

نتیجه گیری و ارائه رهنمودهای اجرایی

۶-۱ نتایج کلیدی پژوهش

همانگونه که در فصل‌های گذشته مورد بحث قرار گرفت، هدف اصلی این پژوهش ارائه یک چارچوب مناسب برای نمونه برداری از مدارک پارسا در فرایندهای نمایه سازی و ویراستاری است. برای رسیدن به این هدف پیش از هر کاری لازم است کنترل‌های صورت گرفته در فرایندهای یادشده را شناسایی و دسته بندی نمود. این کار در قالب یک گروه تمرکزی شامل خبرگان این حوزه صورت پذیرفت. در این راستا همه کنترل‌های صورت گرفته در بخش‌های گوناگون نمایه سازی و ویراستاری مستند و دسته بندی شد. سپس تحلیل آماری گسترده ای بر روی وضعیت موجود مدارک پارسا در هریک از اقلام اطلاعاتی صورت پذیرفت. همچنین پس از شناسایی وضعیت موجود کنترل‌ها و میزان ناهمخوانی‌های شناسایی شده در هریک از بخش‌های نمایه سازی و ویراستاری روش نمونه برداری در هریک از این مراحل ارائه شد. برخی نتایج اصلی این پژوهش در ادامه ارائه می شود.

- نتایج به دست آمده از تحلیل‌های آماری ناهمخوانی‌ها در دو مرحله نمایه سازی و ویراستاری نشان داد که فرایند نمایه سازی می تواند میزان قابل توجهی کیفیت مدارک پارسا را بهبود بخشد. به بیان دیگر کیفیت مدارک پارسا، پیش و پس از نمایه سازی برپایه ارزیابی انجام شده بهبود قابل توجهی را نشان می دهد.
- با تقریب بسیار خوبی می توان گفت کیفیت مدارک پس از نمایه سازی تا سطح نزدیک به ۱۰۰ درصد (وضعیت مطلوب) ارتقاء یافته اند. گفتنی است، تنها در چند مشخصه زیر ناهمخوانی بسیار اندکی مشاهده می شود
 - عنوان فارسی - مشکل ویرایشی و نگارشی
 - عنوان لاتین - مشکلات اصلی (مانند ناهمخوانی با متن پایان نامه و غیره)
 - پدیدآوران - دانشکده
 - گرایش
 - چکیده فارسی - مشکلات نگارشی جزئی (کمتر از ۳ مورد)
 - فهرست مندرجات - ناقص بودن فهرست مندرجات
- در فرایند نمایه سازی، فعالیت‌ها در دو دسته زیر قرار می گیرند.

○ فعالیت‌های ارزش‌افزا (به داده‌های پارسا): مانند تخصیص کلیدواژه‌ها یا تعیین حوزه

موضوعی

○ فعالیت‌های مربوط به کنترل کیفی اقلام اطلاعاتی: مانند کنترل‌های ویرایش و

نگارش عنوان، نام پدیدآور و اصلاح آنها

نکته: فعالیت‌های دسته دوم با فعالیت‌های فرایند کنترل در اداره ثبت و فراهم‌آوری اطلاعات (در سامانه ثبت جدید) هم‌پوشانی فراوانی دارد.

- نتایج تحلیل کیفیت مدارک، پیش و پس از ویراستاری نشان می‌دهد پیش از ویراستاری بسیار از اقلام اطلاعاتی کیفیت مطلوب داشته‌اند و نیازی به کنترل دوباره نبوده است. همچنین، در برخی اقلام اطلاعاتی، پس از نمایه‌سازی، تعداد کمی ناهمخوانی وجود داشته است. براین اساس می‌توان این‌گونه نتیجه‌گیری کرد که می‌توان با بازنگری در فرایندهای ویراستاری کارایی واحد سازماندهی و تحلیل اطلاعات را به میزان قابل توجهی افزایش داد.

۶-۲ ارائه رهنمودهای کاربردی

در این پژوهش برپایه روش‌ها و تحلیل‌های آماری، یک چارچوب نمونه‌برداری در فرایندهای نمایه‌سازی و ویراستاری ارائه شده است. همانگونه که در توضیح‌های ارائه شده در بخش قبل نیز مورد توجه قرار گرفت باتوجه به سطح کیفیت قابل قبولی که مدارک پارسا پس از ثبت و نیز نمایه‌سازی دارند وجود یک طرح نمونه‌برداری می‌تواند کارایی فرایند را به سطح مطلوب نزدیک‌تر کند. براین اساس، طرح نمونه‌برداری به‌گونه‌ای ارائه شد که خطاهای نمونه‌برداری را کاهش داده و اطمینان مناسب از کیفیت مطلوب مدارک را فراهم سازد. در زمان پیاده‌سازی طرح نمونه‌برداری ارائه شده نکات زیر بایسته است مورد توجه قرار گیرد.

- در طراحی روش نمونه‌برداری سه سطح اهمیت برای ناهمخوانی‌ها در نظر گرفته شد. این سه سطح، از مشخصه‌های کیفی بحرانی (A) گرفته تا مشخصه‌های اصلی (B) و ناهمخوانی‌های جزئی (C) در کارگروه کیفیت پارسا و برپایه نیازمندی‌های شناسایی شده دسته‌بندی شد.

دسته بندی ارائه شده ممکن است در گذر زمان توسط کارگروه کیفیت مورد بازبینی قرار گرفته و در مشخصه های هردسته جابه جایی صورت گیرد. پیشنهاد می شود این کار سالیانه انجام شود.

- میزان AQL و LTPD تعیین شده به هردسته از مشخصه های کیفی در کارگروه کیفیت پارسا و برپایه اهمیت هر دسته در کیفیت نهایی مدارک اشاعه داده شده تعیین شد. از آنجا که مقادیر تعیین شده در تعداد نمونه نهایی، عدد پذیرش و همچنین عدد رد نقش به سزایی دارد، در صورت بازبینی دسته بندی های A تا C میزان مقادیر AQL و LTPD متناظر با آنها نیز باید مورد بررسی دوباره صورت گیرد.
- از آنجا که هرگونه بازبینی در مقادیر AQL و LTPD در خطاهای نوع I و II اثرگذار است، لازم است الگوریتم های توسعه داده شده در این پژوهش برای مقادیر تازه AQL و LTPD نیز اجرا شود. بدیهی است نتایج اجرای الگوریتم مانند تعداد نمونه، عدد پذیرش و نیز عدد رد در این فرایند به روزرسانی خواهند شد.

ارائه طرح نمونه برداری پذیرش در فرایند ویراستاری □ ۹۲

مراجع

فهرست منابع فارسی

- اثنی عشری، حمیده و اسدی غلامحسین. (۲۰۱۶). معیارهای سنجش کیفیت اطلاعات حسابداری و بازده اضافی مطلق. دانش حسابداری مالی، ۲(۴)، ۴۷-۷۰.
- ارشادی، محمدجواد، و روزبه ایاسی. ۱۳۹۶. طراحی مدل تضمین کیفیت برای سامانه‌های اطلاعاتی گردآوری و ثبت، سازمان‌دهی و اشاعه اطلاعات پایان‌نامه‌ها/رساله‌های دانش‌آموختگان داخل کشور. فصلنامه علمی-پژوهشی مدیریت فناوری اطلاعات ۹ (۲): ۱۶۷-۱۹۰.
- ارشادی، محمدجواد، تقی رجبی، فرهاد شیرانی، و نسا رضایی. ۱۳۹۵. کاربرد تکنیک تحلیل ریشه در حل مشکلات کیفی سامانه‌های اطلاعاتی تحقیقاتی: مطالعه موردی سامانه اشاعه اطلاعات پایان‌نامه‌ها/رساله‌های دانش‌آموختگان داخل کشور (گنج. (مدیریت اطلاعات ۱) ۱ و ۲): ۷۵-۸۹.
- ایزدی، اعظم سادات. ۱۳۹۳. برخی از طرحهای اقتصادی در کنترل کیفیت آماری. (پایان‌نامه کارشناسی ارشد). دانشگاه فردوسی مشهد.
- حمصی سید امیرهومن، نیکومرام هاشم و سیدمحمدی نیما. ۱۳۸۵. بررسی کاربرد فرایند کنترل آماری کیفیت در شرکت سیما چوب. نشریه علوم کشاورزی، ۱۲ (۲)، ۴۶۳-۴۷۸.

فهرست منابع لاتین

- Banzhaf, R. A., & Brugger, R. M. (1975). MIL-STD-1235 (ORD), single-and multi-level continuous sampling procedures and tables for inspection by attributes. *Journal of Quality Technology*, 7(sup1), 77-89.
- Barton, J., Currier, S., & Hey, J. M. (2003, September). Building quality assurance into metadata creation: an analysis based on the learning objects and e-prints communities of practice. In *International Conference on Dublin Core and Metadata Applications* (pp. 39-48).
- Batini, C., & Scannapieco, M. (2016). *Data and information quality*. Cham, Switzerland: Springer International Publishing. Google Scholar, 43.
- Bebbington, M., Lai, C. D., & Govindaraju, K. (2003). Continuous sampling plans for Markov-dependent production processes under limited inspection capacity. *Mathematical and computer modelling*, 38(11-13), 1137-1145.

- Bruce, T. R., & Hillmann, D. I. (2004). The continuum of metadata quality: defining, expressing, exploiting. ALA editions.
- Cárdenas-García, J. F., De Mesa, B. S., & Castro, D. R. (2019). Understanding Globalized Digital Labor in the Information Age. *Perspectives on Global Development and Technology*, 18(3), 308-326.
- David, R. H., & Thomas, D. (2015). Assessing metadata and controlling quality in scholarly ebooks. *Cataloging & Classification Quarterly*, 53(7), 801-824.
- Day, M., Guy, M., & Powell, A. (2004). Improving the quality of metadata in eprint archives. *Ariadne*, 38.
- Dodge, H. F. (1943). A sampling inspection plan for continuous production. *The Annals of mathematical statistics*, 14(3), 264-279.
- Hillmann, D. I. (2008). Metadata quality: From evaluation to augmentation. *Cataloging & Classification Quarterly*, 46(1), 65-80.
- Keiser, B. E. (1990, May). Quality assurance: A management challenge for the 1990s. In National online meeting (pp. 1-3).
- Lau, A., & Moere, A. V. (2007, July). Towards a model of information aesthetics in information visualization. In 2007 11th International Conference Information Visualization (IV'07) (pp. 87-92). IEEE.
- Marchand, D. (1990). Managing information quality. In: Wormell I.(ed.) *Information Quality: Definitions and Dimensions*. Taylor Graham, Londres.
- McWilliams, T. P., Saniga, E. M., & Davis, D. J. (2001). On the design of single sample acceptance sampling plans. *Stochastics and Quality Control*, 16(2), 193-198.
- Montgomery, D.C. (2009). *Design and analysis of experiments*. 7th ed. John Wiley & Sons, Inc., New York.
- Muñiz Jiménez, C. J. (2017). Reduce Customer Complaints on Packaging Counting Discrepancy Units. *Manufacturing Engineering*;
- Palavitsinis, N., Manouselis, N., & Sanchez-Alonso, S. (2014). Metadata quality in learning object repositories: a case study. *The Electronic Library*.
- Park, J. R. (2009). Metadata quality in digital repositories: A survey of the current state of the art. *Cataloging & classification quarterly*, 47(3-4), 213-228.
- Park, J. R., & Tosaka, Y. (2010). Metadata quality control in digital repositories and collections: criteria, semantics, and mechanisms. *Cataloging & classification quarterly*, 48(8), 696-715.
- Park, J. R., Tosaka, Y., Maszaros, S., & Lu, C. (2010). From metadata creation to metadata quality control: Continuing education needs among cataloging and metadata professionals. *Journal of education for library and information science*, 158-176.
- Price, R., & Shanks, G. (2016). A semiotic information quality framework: development and comparative analysis. In *Enacting Research Methods in Information Systems* (pp. 219-250). Palgrave Macmillan, Cham.
- Ravit H. David & Dana Thomas (2015) Assessing Metadata and Controlling Quality in Scholarly Ebooks, *Cataloging & Classification Quarterly*, 53:7, 801-824, DOI: 10.1080/01639374.2015.1018397

Russell-Rose, T., Chamberlain, J., & Azzopardi, L. (2018). Information retrieval in the workplace: A comparison of professional search practices. *Information Processing & Management*, 54(6), 1042-1057.

Schilling, E. G., & Neubauer, D. V. (2009). *Acceptance sampling in quality control*. Chapman and Hall/CRC.

Stvilia, B., & Gasser, L. (2008). Value-based metadata quality assessment. *Library & Information Science Research*, 30(1), 67-74.

Stvilia, B., Gasser, L., & Twidale, M. B. (2007). Metadata quality problems in federated collections. In *Challenges of Managing Information Quality in Service Organizations* (pp. 154-186). IGI Global.

Tennant, R. (2004). *Specifications for metadata processing tools*. California Digital Library.

Tenopir, C. (1992). A Day in the Life of a Database Producer. *Database*, 15(3), 15-17.

Wang, R. C., & Chang, S. L. (1989). The design of a continuous sampling plan under the conditioned inspection capacity. *Engineering Costs and Production Economics*, 16(4), 269-279.

Wilkinson, L. (2006). Revising the Pareto chart. *The American Statistician*, 60(4), 332-334.

به نام خداوند دانا



پژوهشگاه علوم و فناوری اطلاعات ایران

(ایرانداک)

ایرانداک در نیم قرن دوم کار خود، همچنان جوان و کوشا؛ همراه هر پژوهش، پژوهشگر، و سیاستگذار

(۱۳۹۹-۱۳۴۷)

**راهنمای فرایند نمونه برداری از اقلام اطلاعاتی مدارک
پایان نامه / رساله ها در سامانه ویرایش**

تهران - خیابان انقلاب اسلامی - چهارراه فلسطین - شماره ۱۰۹۰

تلفن: ۶۶۴۹۴۸۵۵ - دورنگار: ۶۶۴۶۲۲۵۴

IRANDOC.AC.IR

INFO@IRANDOC.AC.IR

زمستان ۱۳۹۹

۱. مقدمه

پس از ثبت مدرک توسط دانشجو دو مجموعه اقدام کلیدی بر روی مدرک ثبت شده صورت می‌پذیرد. یکی کنترل مدرک بلافاصله پس از ثبت مدرک توسط همکاران مدیریت و سازماندهی اطلاعات در قالب سامانه ثبت پایان‌نامه‌ها/رساله‌ها و پیشنهادها که تمرکز اصلی آن اطمینان از صحت فراداده‌های ثبت شده در سامانه و کیفیت مطلوب فایل‌های پی.دی.اف و ورد است. مجموعه اقدام دوم نمایه‌سازی و ویراستاری مدرک ثبت شده است که با هدف ارتقاء سطح کیفیت مدرک، افزایش قابلیت بازیابی آن و همچنین ارائه خدمات ارزش افزوده به پژوهشگران مانند فهرست واژگان، نام‌ها و همچنین آمایش سرزمین است. مجموعه اقدام دوم توسط نمایه‌سازان و ویراستاران در قالب سامانه ویرایش و در واحد سازماندهی و تحلیل اطلاعات انجام می‌پذیرد. براین اساس، می‌توان مسئولیت‌ها در حوزه کیفیت اقلام اطلاعاتی را بین این دو واحد تقسیم‌بندی کرد. در مواردی که مسئول اصلی اقلام اطلاعاتی همکاران واحد مدیریت ثبت و فراهم‌آوری اطلاعات هستند، همکاران واحد سازماندهی و تحلیل اطلاعات می‌توانند متناسب با سطح کیفیت تعیین شده و ظرفیت مورد انتظار نمایه‌سازی به‌صورت نمونه‌برداری به بررسی و ویرایش مدرک ثبت شده بپردازند. در سایر موارد (مانند اضافه کردن کلیدواژه، ویرایش کلیدواژه و تعیین حوزه موضوعی) مسئولیت به‌صورت کامل برعهده همکاران مدیریت سازماندهی و تحلیل اطلاعات خواهد بود.

۲. معرفی واژگان و اصطلاحات کلیدی

در این راهنما برخی واژگان تخصصی نمونه‌برداری استفاده شده اند که در ادامه به معرفی آنها خواهیم پرداخت.

حدود قابل پذیرش (AQL^۱): سطح کیفیت قابل قبول برای سازمان که معمولاً به درصد بیان می‌شود. در این راهنما مراد از AQL میزان درصدی از پارساهاست که انتظار داریم درصد ناهمخوانی‌ها (نسبت مدارک ناهمخوان به کل مدارک) برای یک فرد از این میزان بیشتر نشود.

حدود قابل رد (LTPD^۲): سطحی از ناهمخوانی‌ها که غیرقابل قبول بوده و انتظار داریم در فرایند نمونه‌برداری این میزان از ناهمخوانی در مدارک خروجی یک فرد به راحتی کشف شود.

میزان انباشته: کل مدارک پردازش شده توسط یک فرد در بازه زمانی مشخص. برای نمونه اگر بازه زمانی عملکرد یک فرد نمایه‌ساز ۱۰ روز کاری در نظر گرفته شود و این فرد در هر روز ۱۶٫۱ مدرک را نمایه‌سازی کرده باشد میزان انباشته ۱۶۱ مدرک خواهد بود.

۳. دسته‌بندی ناهمخوانی‌ها برپایه میزان اهمیت

در صورتی که کنترل کیفیت اقلام اطلاعاتی سامانه ثبت برپایه نمونه‌برداری انجام پذیرد لازم است ناهمخوانی‌ها به سه صورت زیر دسته‌بندی شوند.

- مشخصه نوع A

○ این ناهمخوانی بحرانی بوده و به صحت پارسا مرتبط است. همچنین، ممکن است در نمایش یک داده بی‌کیفیت که تاثیرگذاری شدیدی بر دیدگاه کاربران دارد در سامانه گنج منجر شود.

- مشخصه نوع B

○ این ناهمخوانی مهم بوده و به دقت پارسا مرتبط است. همچنین، ممکن است به نمایش یک داده بی‌کیفیت که تاثیرگذاری شدید بر دیدگاه کاربران دارد در سامانه گنج منجر شود.

^۱ Acceptable Quality Level

^۲ Lot Tolerance Percent Defective

• مشخصه نوع C

○ این ناهمخوانی جزئی است یا تاثیر زیادی در رضایت کاربران در سامانه گنج ندارد.

در فرایند نمونه برداری لازم است حدود قابل پذیرش (AQL) و حدود قابل رد (LTPD) برپایه اهمیت هر ناهمخوانی تعیین شوند. در جدول ۱ پارامترهای AQL و LTPD بر مبنای میزان اهمیت هر ناهمخوانی تعیین شدند. همچنین نوع ناهمخوانی نیز در همین جدول تعیین شده‌اند.

جدول ۱ تعیین میزان اهمیت ناهمخوانی‌ها در سامانه ثبت

ردیف	نام قلم اطلاعاتی	دسته ناهمخوانی	نوع مشخصه کیفی	AQL	LTPD
۱	عنوان فارسی	مشکل نگارشی	A	0.005	0.05
		مشکل اصلی (صحت) ناهمخوانی با متن و ...)	A	0.005	0.05
۲	عنوان لاتین	مشکل ویرایشی / نگارشی	B	0.01	0.08
		مشکل اصلی (صحت) ناهمخوانی با متن و ...)	B	0.01	0.08
۳	پدیدآوران	دانشگاه	A	0.005	0.05
		دانشکده	B	0.01	0.08

ردیف	نام قلم اطلاعاتی	دسته ناهمخوانی	نوع مشخصه کیفی	AQL	LTPD
		دانشجو	A	0.005	0.05
		اساتید راهنما	A	0.005	0.05
		اساتید مشاور	A	0.005	0.05
۴	تاریخ دفاع	–	C	0.03	0.1
۵	رشته	–	B	0.01	0.08
۶	گرایش	–	B	0.01	0.08
۷	چکیده فارسی	وجود کلمات اضافی (مانند کلمه چکیده در ابتدا)	A	0.005	0.05
		ناهمخوانی با متن پایان نامه	A	0.005	0.05
۸	چکیده لاتین	وجود کلمات اضافی (مانند کلمه چکیده در ابتدا)	B	0.01	0.08
		ناهمخوانی با متن پایان نامه	B	0.01	0.08
۹	فهرست مندرجات	–	C	0.03	0.1
۱۰	فهرست منابع فارسی	–	C	0.03	0.1

ردیف	نام قلم اطلاعاتی	دسته ناهمخوانی	نوع مشخصه کیفی	AQL	LTPD
۱۱	فهرست منابع لاتین	-	C	0.03	0.1

۴. طرح نمونه برداری

در هر طرح نمونه برداری سه متغیر اصلی باید تعیین شود. اولین متغیر تعداد نمونه (n) است. دومین متغیر عدد پذیرش (c) است و سومین متغیر عدد رد (r) است. به بیان دیگر باید تعیین کنیم که کارشناس کنترل کننده هر بار چند مدرک را باید انتخاب کند (n) و پس از انتخاب تصادفی این n مدرک در صورت مشاهده بیشینه چند مدرک ناهمخوان (c) کیفیت کار را تایید کند و در صورت مشاهده کمینه چند مدرک ناهمخوان (r) کیفیت را رد کند. از سوی دیگر باید مشخص شود میزان انباشته چقدر است و به بیان دیگر، نمونه برداری از کار چندروز همکاران انجام خواهد پذیرفت. در ۶ ستون اول جدول ۲ این مقادیر مشخص شده اند.

اکنون با داشتن همه پارامترهای پیشین، یک طرح نمونه برداری را ارائه خواهیم کرد. به بیان دیگر، در هریک از مراحل سه گانه ثبت، نمایه سازی و ویرایش و به ازای هر نوع مشخصه کیفی یک طرح نمونه برداری به دست خواهد آمد. در ادامه و در جدول ۲ طرح نمونه برداری قابل مشاهده خواهد بود.

جدول ۲ طرح نمونه برداری بهینه برای مراحل گوناگون ثبت و سازماندهی اطلاعات

ردیف	نام مرحله	مرحله کنترل کننده	خط مشی عملکرد مرحله	بازه زمانی عملکرد	حدود n	نوع مشخصه	AQL	LTPD	n	c	r
۱	ثبت و فراهم آوری اطلاعات	نمایه سازی	روزانه ۷۰ مدرک	۱۰ روز کاری (دو هفته) (	A	از ۱ تا ۷۰۰	0.005	0.05	105	2	3
					B		0.01	0.08	65	2	3
					C		0.03	0.1	116	7	8

ردیف	نام مرحله	مرحله کنترل کننده	خط مشی عملکرد مرحله	بازه زمانی عملکرد*	حدود n	نوع مشخصه	AQL	LTPD	n	c	r
۲	نمایه سازی	ویراستاری	روزانه ۱۶,۱ مدرک	۱۵ روز کاری (سه هفته)	از ۱ تا ۲۴۲	A	0.005	0.05	105	2	3
						B	0.01	0.08	65	2	3
						C	0.03	0.1	116	7	8
۳	ویراستاری	مدیر واحد سازماندهی و تحلیل اطلاعات	روزانه ۷۰ مدرک	۱۰ روز کاری (یک هفته)	از ۱ تا ۷۰۰	A	0.005	0.05	105	2	3
						B	0.01	0.08	65	2	3
						C	0.03	0.1	116	7	8

* میزان انباشته برای نمونه برداری

همانگونه که از جدول ۲ قابل مشاهده است در هریک از مراحل گوناگون به ازای هر گروه از مشخصه های کیفی، طرح نمونه برداری تعیین شده است. برای نمونه، در مرحله نمایه سازی، کارشناس نمایه ساز از ۱۰ روز کاری پیشین یک کارشناس ثبت و فراهم آوری اطلاعات ۱۰۵ نمونه انتخاب خواهد کرد. در صورت مشاهده بیشینه ۲ ناهمخوانی در هر مشخصه نوع A، حجم کار انجام شده تایید خواهد شد و در صورت مشاهده ۳ یا بیش از ۳ ناهمخوانی کار مورد تایید نخواهد بود. عدهای پذیرش و رد در مشخصه های نوع B، به ترتیب ۲ و ۳ و در مشخصه های نوع C، به ترتیب ۷ و ۸ خواهد بود.

مراحل اجرایی یک فرایند نمونه برداری در قالب یک مثال در ادامه ارائه شده است. فرض کنید ویراستار بخواهد عملکرد نمایه ساز را در ناهمخوانی نوع B برپایه ردیف ۲ جدول بالا کنترل کند. گامهای اجرایی زیر را باید طی کند.

(۱) مشاهده مدارک نمایه سازی شده ۱۵ روز کاری نمایه ساز

(۲) انتخاب ۶۵ نمونه به صورت تصادفی

(۳) شمارش تعداد مدارک ناهمخوان

۴) پذیرش عملکرد نمایه ساز در صورتی که تعداد مدارک ناهمخوان ۲ یا کمتر از ۲ است.

۵) عدم پذیرش عملکرد نمایه ساز در صورتی که تعداد مدارک ناهمخوان ۳ یا بیشتر از ۳ است

۶) پایان نمونه برداری

Abstract

In the thesis / dissertation (Parsa) documents of the graduates of the whole country, indexing is a key process that, if done correctly, provides the possibility of proper retrieval in the treasure system. The current status of registration, indexing, and editing of Parsa documents is such that after the registration of documents by graduate students in the final registration and control system by the colleagues of the supply unit, the approved documents are automatically entered into a system called "Editing and organizing scientific information". In this system, in addition to re-checking all the thesis information items/ dissertations, indexing is done and then sent to the editorial staff. In this section, the final control is performed on the indexes made and if there is a problem, it is returned, otherwise it is sent to the treasure system for dissemination. In the current situation, in addition to the previous controls, all documents are re-examined by the head of the unit in order to evaluate the quality of the work done. Hence, as can be seen, some items may have multiple controls. The importance of this research will increase when in the new registration system, since some information items (such as university name, field and orientation, etc.) are selected as a pre-determined list by the user, there is no possibility of error. Therefore, there is no need to control these items in the information editing and organizing system. In this research, a framework for sampling in indexing and editing processes is provided. In this research, in the stage of presenting the sampling method in addition to using the focus group method, in order to statistically design an optimization model to minimize type I and II errors has been used. The results of this study showed that, considering the acceptable quality of many pious information items after indexing, sampling is a vital task in improving the efficiency of the information organization and analysis unit. The classification of information items into three categories is critical, main and partial, and determining the number and method of sampling for each category is another result of this research.

Keywords: indexing; sampling; Quality Control; Iran Institute of Information Science and Technology (Irاندک); Organize and analyze information



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

Research Project

Providing an Acceptance Sampling Plan for Thesis/ Dissertations in the Indexing and Editing Processes

By:

Mohammad Javad Ershadi

Co-Author:

Zahra Dehsaraie

November 2020