



ارزیابی کیفیت داده‌های ثبتی یا اداری برای استفاده در سرشماری ثبتی مبنا

مجری

سید نعمت‌اله میرفلاح نصیری

همکاران

کاوه کیانی

علی اکبر محزون

شکوفه قصوری

محمد شیرینی

اشکان شباک

مریم پوررضانور

الهام فتحی

حسام خدامرادی

آسیه عباسی

بهناز سرخیل

سعید فیاض

گروه پژوهشی پردازش داده‌ها و اطلاع‌رسانی

پاییز ۱۳۹۸

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

کد شناسه: RP-9940

عنوان: ارزیابی کیفیت داده‌های ثبتی یا اداری برای استفاده در سرشماری ثبتی منبأ

مجری: سید نعمت‌اله میرفلاح نصیری

همکاران: کاوه کیانی، علی‌اکبر محزون، شکوفه قصوری، محمد شیرینی، اشکان شباک، مریم پوررضانور، الهام فتحی، حسام خدامرادی، آسیه عباسی، بهناز سرخیل، سعید فیاض

گروه پژوهشی: پردازش داده‌ها و اطلاع‌رسانی

تاریخ انتشار: پاییز ۱۳۹۸

نوبت انتشار: اول

طراح جلد: ساناز مهندسی

صفحه‌آرا: نجمه ناظریان

❖ حق مالکیت معنوی این طرح پژوهشی متعلق به پژوهشکده‌ی آمار است و نقل مطالب فقط با ذکر مأخذ مجاز است.



تهران، خیابان دکتر فاطمی، خیابان باباطاهر، خیابان سرتیپ فکوری،

شماره‌ی ۱۴۵

۰۲۱ ۸۸۶۳۰۴۴۰ -۳

www.srtc.ac.ir

پیش‌گفتار

دستیابی به برآوردهای قابل اعتماد از جمعیت، ترکیب خانوارها، ویژگی‌های آن‌ها و مشخصات محل سکونت آن‌ها هدف انجام سرشماری‌ها است. در کنار اهمیت و مطلوبیت تغییر روش سرشماری از سرشماری سنتی به سرشماری ثبتی مبنای برای سازمان‌های آماری، «ارزیابی کیفیت» از مهم‌ترین و پایه‌ای‌ترین جنبه‌های بهره‌برداری از داده‌های این سرشماری است. اندازه‌گیری درستی و صحت داده‌ها باید با روش‌ها و تحت یک مدیریت ارزیابی کیفیت منحصر به خود انجام شود و لزوماً نمی‌توان از معیارهای ارزیابی کیفیت آماری‌های سنتی بهره گرفت و تنها در صورتی که داده‌های اداری و آمارهای ثبتی دارای کیفیت کافی بوده و داده‌های در دسترس با مفاهیم مورد نیاز، مطابقت داشته باشند می‌توان از آن‌ها به جای آماری مستقیم استفاده کرد.

مفهوم ارزیابی کیفیت و به عبارت دیگر مفهوم ذات کیفیت متفاوت خواهد بود. لذا لازم است با نگاه به روش‌شناسی تهیه و تولید آمارهای ثبتی مبنای چارچوبی مطابق با سرشماری ثبتی مبنای طراحی شود. کیفیت داده‌های ورودی به قدری اهمیت دارد که می‌تواند کیفیت کل فرایند سرشماری ثبتی مبنای را تحت شعاع قرار دهد و به عبارت دیگر، ارزیابی کیفیت از ضروریات اجرای سرشماری ثبتی-مبنای است که با ارزیابی کیفیت سرشماری‌های سنتی و آماری‌های نمونه‌ای متفاوت است. زمانی که منابع ترکیب می‌شوند، ارزیابی سازگاری وظیفه‌ی جدید منحصر به آماری‌های ثبتی است. ارزیابی کیفیت کاری نظام‌مند جهت تشخیص خطاهای مشخص و محتمل است. بنا بر این، ویرایش، تضمین کیفیت، و طراحی یک سرشماری ثبتی-مبنای برای این کار مباحثی در ارتباط مستقیم با یکدیگر هستند که در این تحقیق به آن پرداخته می‌شود. علاوه بر آن فراهم بودن زیرساخت‌های مرتبط در حوزه IT و تهیه بستر سیستمی استاندارد، متناسب با بانک‌های اطلاعاتی دریافتی و همچنین فراهم نمودن زیرساخت‌های قانونی به منظور دریافت ریزداده‌های مورد نیاز از سازمان‌ها و استفاده از داده‌های اداری امری ضروری است.

این گزارش، در گام اول به دنبال طراحی «چارچوب ارزیابی کیفیت» برای ارزیابی داده‌های ورودی جهت استفاده در سرشماری ثبتی مبنای است. چارچوب ارزیابی کیفیت مورد هدف این طرح پژوهشی، به نحوی طراحی شده است که ابعاد مختلف کیفیت داده‌های ثبتی را مورد توجه قرار داده است.

این طرح با همکاری جناب آقای سید نعمت‌الله میرفلاح نصیری به عنوان مجری طرح و و همکاری کاوه کیانی، علی اکبر محزون، شکوفه قصوری، محمد شیرینی، اشکان شباک، مریم پوررضانور، الهام فتحی، حسام خدامرادی، آسیه عباسی، بهناز سرخیل و سعید فیاض به انجام رسیده است، که بدین وسیله از همه عزیزان صمیمانه تشکر و قدردانی می‌شود.

گروه پژوهشی پردازش داده‌ها و اطلاع‌رسانی

ژوئیه ۱۳۹۵

فهرست مطالب

کلیات	۱
۱-۱- شرح مسئله	۱
۲-۱- ضرورت تحقیق	۲
۳-۱- سابقه‌ی علمی مسئله	۳
۴-۱- اهداف تحقیق	۳
۵-۱- مراحل انجام تحقیق	۴
مطالعات تطبیقی، توصیه‌ها و استانداردهای بین‌المللی موجود در خصوص ارزیابی کیفیت داده‌های ثبتی	۵
۱-۲- مقدمه	۵
۲-۲- مطالعات تطبیقی کشورها	۵
۱-۲-۲- اتریش	۵
۲-۲-۲- ایرلند	۶
۳-۲-۲- استونی	۹
۴-۲-۲- لهستان	۱۲
۵-۲-۲- انگلستان و ولز	۱۶
۶-۲-۲- ایتالیا	۲۲
۷-۲-۲- مطالعه موردی آلمان	۲۶
۳-۲- نتیجه‌گیری	۳۰
تدوین چارچوب ارزیابی کیفیت داده‌های اداری مورد استفاده در سرشماری ثبتی مبنا با توجه به زیر ساخت‌های آماری و ثبتی کشور	۳۳
۱-۳- مقدمه	۳۳
۲-۳- تعاریف ابعاد کیفیت	۳۴
۳-۳- ارزیابی کیفیت داده‌های ثبتی به منظور استفاده در سرشماری ثبتی مبنا	۳۶
۴-۳- کیفیت داده ثبتی	۳۷
۵-۳- مفهوم کیفیت آمارهای رسمی	۳۸

- ۳-۶- نگاه‌های بر برخی از شاخص‌های مورد استفاده در دیگر کشورها ۴۱
- ۳-۶-۱- ارزیابی کیفیت داده‌های اداری به عنوان ورودی برای آمار رسمی ۴۱
- ۳-۶-۲- فهرست روش‌های موجود ۴۳
- ۳-۶-۳- پیشنهاد یک روش تجدید نظر شده ۴۴
- ۳-۶-۴- تست روش تجدید نظر شده ۴۵
- ۳-۶-۵- چارچوب کیفیت منابع داده‌های اداری در سرشماری ۲۰۱۱ اتریش ۴۹
- ۳-۶-۶- ارزیابی کیفیت در هلند ۵۷
- ۳-۶-۷- نگاه‌های بر سیستم تولید آمار ۶۰
- ۳-۶-۸- ارزیابی کیفیت خروجی‌ها ۶۱
- ۳-۶-۹- ارزیابی کیفیت ورودی‌ها ۶۲
- ۳-۶-۱۰- اطلاعات مورد نظر از ساختار اداری مسئول در ثبت اداری ۶۶
- ۳-۶-۱۱- یکپارچه‌سازی منابع داده‌ای ثبتی با ثبت‌های پایه (ثبت‌های چهارگانه) ۶۸
- ۳-۶-۱۲- یکپارچه‌سازی با طرح‌های آماری با متغیرهای مشابه ۷۰
- ۳-۶-۱۳- شاخص‌ها ۷۴
- ۳-۶-۱۴- خصوصیت شاخص‌ها ۷۵
- ۳-۶-۱۵- شیوه ارزیابی ۷۶
- ۳-۶-۱۶- نتیجه‌گیری ۷۷
- تدوین چارچوب ارزیابی کیفیت داده‌های اداری مورد استفاده در ابزارهای ارزیابی کیفیت ۷۹
- ۴-۱- مقدمه ۷۹
- ۴-۲- فهرست شاخص‌های ارزیابی کیفیت داده‌های ثبتی یا اداری مورد استفاده در سرشماری ثبتی مبنا ۸۰
- ۴-۳- راهنمای اندازه‌گیری شاخص‌های ارزیابی کیفیت داده‌های ثبتی یا اداری مورد استفاده در سرشماری ثبتی مبنا ۸۲
- ۴-۴- ارزیابی تفصیلی (پس از دریافت ریزداده‌ها و انجام بررسی‌های عمیق) ۸۸
- پایلوت ارزیابی کیفیت بانک اطلاعاتی اقلام مربوط به وضع تحصیل برای استفاده در سرشماری ثبتی مبنا ۹۷
- ۵-۱- مقدمه ۹۷
- ۵-۲- مروری بر اقلام ضروری برای اجرای سرشماری ثبتی مبنا ۹۸
- ۵-۳- پیاده‌سازی روش‌های ارزیابی بر روی پایگاه آموزش و پرورش ۱۰۲
- ۵-۴- نتیجه‌گیری ۱۱۱
- مرجع‌ها ۱۱۳

فهرست جدول‌ها

جدول ۱-۳	روش‌ها برای ارزیابی کیفیت داده‌های ورودی برای آمار ثبتي مبنا	۴۴
جدول ۲-۳	نتایج اتریش	۴۶
جدول ۳-۳	نتایج دانمارک	۴۷
جدول ۴-۳	سؤالات امتیازدهی - مستندات HD	۵۳
جدول ۵-۳	محک‌زنی سرشماری خرد-وضع زناشویی	۵۴
جدول ۶-۳	مقایسه توزیع حاشیه‌ای-وضع زناشویی	۵۴
جدول ۷-۳	پیوند CDB - ویژگی‌های منحصر به فرد	۵۵
جدول ۸-۳	پیوند CDB - ویژگی‌های چندگانه	۵۶
جدول ۹-۳	ابعاد، شاخص‌های کیفیت و روش‌های ارزیابی منبع	۵۷
جدول ۱۰-۳	ابعاد، شاخص‌های کیفیت و روش‌های ارزیابی متا داده	۵۸
جدول ۱۱-۳	ابعاد، شاخص‌های کیفیت و روش‌های ارزیابی کیفیت داده‌ها	۵۹
جدول ۱۲-۳	نتایج ارزیابی بدست آمده از ارزیابی ثبت‌های اداری	۵۹
جدول ۱۳-۳	نتایج ارزیابی بدست آمده از ارزیابی متا دیتای ثبت‌های اداری	۶۰
جدول ۱۴-۳	شاخص‌های ارزیابی کیفیت خروجی‌ها برای داده‌های اداری در کشور سوئد	۶۲
جدول ۱۵-۳	شاخص‌های ارزیابی کیفیت خروجی‌ها و ورودی‌ها - شاخص مرتبط بودن	۶۴
جدول ۱۶-۳	شاخص‌های ارزیابی کیفیت خروجی‌ها و ورودی‌ها - شاخص دقت	۶۷
جدول ۱۷-۳	شاخص‌های ارزیابی کیفیت خروجی‌ها و ورودی‌ها - شاخص دقت	۶۹
جدول ۱۸-۳	شاخص‌های ارزیابی کیفیت داده‌های ورودی و فرایند تولید	۷۰
جدول ۱۹-۳	شاخص‌های ارزیابی کیفیت داده‌های ورودی - در دسترس بودن	۷۴
جدول ۲۰-۳	رتبه بندی ثبت بر اساس مدل ارزیابی کیفیت پیشنهادی	۷۶
جدول ۱-۴	چک‌لیست ارزیابی داده‌های ثبتی یا اداری مورد استفاده در سرشماری ثبتی مبنا و ابعاد کیفیت متناظر با هر یک از شاخص‌ها	۹۲
جدول ۲-۴	امتیاز هر یک از ابعاد کیفیت برای شاخص‌های ۱ تا ۲۸	۹۳
جدول ۳-۴	نتیجه ارزیابی داده‌های اداری یا ثبتی مورد استفاده در سرشماری ثبتی مبنا پیش از تبادل ریزداده‌ها و براساس بازبندی‌های صورت گرفته از منبع داده	۹۴

- جدول ۴-۴- امتیاز هر یک از ابعاد کیفیت برای شاخص‌های ۲۹ تا ۴۰ ۹۴
- جدول ۴-۵- نتیجه ارزیابی داده‌های اداری یا ثبتی مورد استفاده در سرشماری ثبتی مینا پس از دریافت ریزداده‌ها و انجام بررسی‌های عمیق ۹۴
- جدول ۴-۶- امتیاز هر یک از ابعاد کیفیت برای تمامی شاخص‌ها ۹۵
- جدول ۴-۷- نتیجه کلی ارزیابی داده‌های اداری یا ثبتی مورد استفاده در سرشماری ثبتی مینا ۹۵
- جدول ۵-۱- فهرست اقلام جغرافیایی و مکانی و جامعه آماری، روش تهیه و سازمان متولی آنها ۹۸
- جدول ۵-۲- فهرست اقلام جمعیتی، جامعه آماری، روش تهیه و سازمان متولی آنها ۹۹
- جدول ۵-۳- معیار ارزیابی کیفیت ۱۰۲
- جدول ۵-۴- نتیجه ارزیابی داده‌های اداری/ثبتی مورد استفاده در سرشماری ثبتی مینا پیش از تبادل ریزداده‌ها و براساس بازدیدهای صورت گرفته از منبع داده ۱۰۸
- جدول ۵-۵- چک‌لیست ارزیابی داده‌های ثبتی/اداری مورد استفاده در سرشماری ثبتی مینا و ابعاد کیفیت متناظر با هر یک از شاخص‌ها ۱۰۹
- جدول ۵-۶- امتیاز هر یک از ابعاد کیفیت برای شاخص‌های ۱ تا ۲۸ ۱۱۰
- جدول ۵-۷- امتیاز هر یک از ابعاد کیفیت برای شاخص‌های ۲۹ تا ۴۰ ۱۱۱
- جدول ۵-۸- امتیاز هر یک از ابعاد کیفیت برای تمامی شاخص‌ها ۱۱۱

فهرست شکل‌ها

- شکل ۱-۲- فرایند سرشماری در و ارزیابی کیفیت ۱۵
- شکل ۲-۲- سیستم سرشماری IT ۱۶
- شکل ۳-۲- چارچوب مشترک انگلستان برای سرشماری های ثبتي مبنا يا تركيبی ۱۹
- شکل ۴-۲- بررسی اجمالی سیستم یکپارچه سرشماری و طرح‌های اجتماعی ۲۶
- شکل ۵-۲- مدل سرشماری آلمان در سال ۲۰۱۱ ۲۹
- شکل ۱-۳- مراحل مختلف در فرایند تولید آمارهای ثبتي ۳۷
- شکل ۲-۳- مدل ثبتي اتریش ۵۰
- شکل ۳-۳- چارچوب کیفیت اتریش ۵۱
- شکل ۴-۳- روش‌های مختلف استفاده از داده‌های اداری در تولید محصولات آماری ۶۱

کلیات

۱-۱- شرح مسئله

سرشماری جمعیت و نتایج حاصل از آن نقش محوری در سیستم آمارهای ملی و رسمی دارد. دستیابی به برآوردهای قابل اعتماد از جمعیت، ترکیب خانوارها، ویژگی‌های آنها و مشخصات محل سکونت آنها هدف انجام سرشماری‌ها است. در حال حاضر سرشماری‌های سنتی هر ده سال و یا هر پنج سال یکبار با صرف وقت و هزینه‌های زیادی انجام می‌شود که سازمان‌های آماری را با چالش نیاز به آمارهای به‌روز و به‌هنگام و در نتیجه گذر از سرشماری سنتی به سرشماری ثبتی مبنا روبرو کرده است. هر چند حرکت از سرشماری سنتی به سرشماری ثبتی مبنا برای سازمان‌های آماری مطلوب است و کاهش هزینه و بار پاسخگویی را به دنبال دارد اما با چالش‌ها و پرسش‌های جدی نیز روبرو است که یکی از مهم‌ترین آنها «ارزیابی کیفیت» داده‌های ثبتی به منظور استفاده در سرشماری ثبتی مبنا است. اندازه‌گیری درستی و صحت داده‌ها باید با روش‌ها و تحت یک مدیریت ارزیابی کیفیت منحصر به خود انجام شود و لزوماً نمی‌توان از معیارهای ارزیابی کیفیت آمارگیری‌های سنتی بهره‌گرفت و تنها در صورتی که داده‌های اداری و آمارهای ثبتی دارای کیفیت کافی بوده و داده‌های در دسترس با مفاهیم مورد نیاز، مطابقت داشته باشند می‌توان از آنها به جای آمارگیری مستقیم استفاده کرد.

طبق تقسیم‌بندی که در گزارش «ارزیابی کیفیت داده‌های ثبتی» توسط اداره آمار اروپا آمده است، ارزیابی کیفیت داده‌های ثبتی در سه مرحله انجام می‌شود:

- مرحله داده‌های ورودی
- مرحله پردازش داده‌ها
- مرحله ساخت فایل سرشماری ثبتی مبنا

این گزارش، به دنبال طراحی «چارچوب ارزیابی کیفیت» برای داده‌های مورد استفاده در مرحله اول یعنی ارزیابی داده‌های ورودی از منابع اداری/ ثبتی جهت استفاده در سرشماری ثبتی مبنا است. چارچوب ارزیابی کیفیت مورد هدف این طرح پژوهشی، باید به نحوی طراحی شود که ابعاد مختلف کیفیت داده‌های ثبتی مورد توجه قرار گیرد.

علاوه بر این در راستای شکل‌گیری اهدافی که برای نظام جامع ثبت‌های آماری ایران^۱ تعیین شده است، انجام تحلیل کیفیت داده‌های ثبتی موجود و فعالیت ارزیابی کیفیت و بهبود ثبت جمعیت مطرح می‌باشد که این موضوع نیز گویای اهمیت ارزیابی کیفیت در استقرار نظام ثبتی کشور می‌باشد و در احکام مصوب شورای عالی آمار نیز آورده شده است که «دستگاه‌های اجرایی موظفانند بهبود کیفیت یا ایجاد ثبت‌های اداری مورد نیاز نظام جامع ثبت‌های آماری ایران را در اولویت کاری خود قرار دهند و مسوولیت ارزیابی کیفیت این ثبت‌های اداری بر عهده مرکز آمار ایران است.»

۱-۲- ضرورت تحقیق

بر اساس مباحث روز و تجربیات نوین بین‌المللی، مانند بسیاری از کشورهای دیگر برنامه‌گذار از نظام سنتی به نوین ثبتی-مبنا در دستور کار نظام آماری کشور قرار گرفته است هرچند در این برنامه‌گذار اصلاح و بهبود قانون آمار کشور، توسعه زیرساخت‌های فناوری اطلاعات، بهبود روش‌های آمارگیری، افزایش ظرفیت نظام آماری و حتی توسعه سواد آماری در کشور مورد توجه قرار گرفته است ولی توسعه و تولید آمار به روش ثبتی از اهمیت کلیدی برخوردار است. بر همین اساس در چارچوب برنامه‌گذار نظام آماری از سنتی به ثبتی مبنا، استقرار نظام جامع ثبت‌های آماری کشور در حال پیگیری است. اما نکته قابل توجه در این گذار این است که بین آمارهای حاصل از آمارگیری‌ها و آمارهای حاصل از منابع اداری تفاوت قابل توجهی وجود دارد و به عبارتی مفهوم ارزیابی کیفیت و به عبارت دیگر مفهوم ذات کیفیت متفاوت خواهد بود. لذا لازم است با نگاه به روش‌شناسی تهیه و تولید آمارهای ثبتی مبنا چارچوبی مطابق با سرشماری ثبتی مبنا طراحی شود. کیفیت داده‌های ورودی به قدری اهمیت دارد که می‌تواند کیفیت کل فرایند سرشماری ثبتی مبنا را تحت شعاع قرار دهد و به عبارت دیگر، ارزیابی کیفیت از ضروریات اجرای سرشماری ثبتی مبنا است که با ارزیابی کیفیت سرشماری‌های سنتی و آمارگیری‌های نمونه‌ای متفاوت است. زمانی که منابع ترکیب می‌شوند، ارزیابی سازگاری وظیفه‌ی جدید منحصر به آمارگیری‌های ثبتی است. ارزیابی کیفیت کاری نظام‌مند جهت تشخیص خطاهای مشخص و محتمل است. بنابراین، ویرایش، تضمین کیفیت و طراحی سرشماری ثبتی مبنا برای این کار مباحثی در ارتباط مستقیم با یکدیگر هستند که در این تحقیق به آن پرداخته می‌شود. علاوه بر آن فراهم بودن زیرساخت‌های مرتبط در حوزه IT و تهیه بسترسیستمی استاندارد، متناسب با بانک‌های اطلاعاتی دریافتی و همچنین فراهم نمودن زیرساخت‌های قانونی به منظور دریافت ریزداده‌های مورد نیاز از سازمان‌ها و استفاده از داده‌های اداری امری ضروری است. دریافت بانک اطلاعاتی حاوی ریز داده با ابلاغ بخشنامه شماره ۵۴۱۵۴/۱۶۸۶۱۴ مورخ ۹۶/۱۲/۲۷ معاون اول محترم رییس جمهور به کلیه وزارتخانه‌ها، سازمان‌ها و مؤسسات دولتی و غیر دولتی و نیز ابلاغ بخشنامه شماره ۱۷۴۶۷۶۴ مورخ ۹۶/۱۲/۱۹ رییس محترم سازمان برنامه و بودجه و کشور به کلیه معاونت‌های ریاست جمهوری و وزرا و رؤسای سازمان‌ها و دستگاه‌های اجرایی، مبنای قانونی استفاده از ریزداده‌ها قرار گرفت.

¹ IRANSTARS

۱-۳- سابقه‌ی علمی مسئله

در سال‌های اخیر سازمان‌های آماری کشورهای مختلف گام‌های مهمی در جهت اجرای سرشماری ثبتی مینا برداشته‌اند. پیش‌گامی بسیاری از کشورها منجر به تولید مستندات و تجربیات زیادی شده است که به عنوان مثال به مستندات توصیه‌های سازمان ملل، اداره آمار اروپا، تجربیات کشورهای نوردیک و استرالیا و ... می‌توان اشاره کرد.

در مرکز آمار ایران نیز مجموعه‌ای با عنوان «چارچوب کیفیت آمارهای مرکز آمار ایران» در سال ۱۳۹۵ تهیه و تدوین شد. این مجموعه که به عنوان اولین چارچوب یکپارچه کیفیت برای مدیریت کیفیت آمارهای مرکز به شمار می‌رود، با در نظر گرفتن ظرفیت‌ها و شرایط نظام آماری کشور و نیز با توجه به استانداردهای بین‌المللی و بر اساس چارچوب پیشنهادی کمیسیون آماری سازمان ملل تهیه شده است. در این گزارش به جنبه‌های مختلف ارزیابی داده‌های ثبتی پرداخته شده که می‌توان از این مجموعه در این طرح پژوهشی بهره گرفت.

کمیسیون آمار سازمان ملل به عنوان بالاترین مرجع بین‌المللی آمار رسمی، در پی چهل و سومین نشست خود در سال ۲۰۱۲ به سازمان‌های ملی آمار کشورهای عضو پیشنهاد کرد تا برای افزایش کیفیت آمارهای خود، طی یک بازه‌ی پنج ساله تهیه چارچوب ملی کیفیت آمارهای خود را مطابق استانداردها و الگوهای معرفی شده در آن نشست در دستور کار قرار داده و بر آن اساس یک نظام تضمین و ارزیابی کیفیت سازمان‌یافته را مستقر نمایند. در حال حاضر اغلب کشورهای توسعه یافته در زمینه آمار دارای چارچوب ملی کیفیت آمارها هستند که خوشبختانه مرکز آمار ایران نیز در راستای حرکت خود به سوی یک مرکز آمار مدرن و پیشرو در آسیا و برای تضمین کیفیت محصولات آماری خود، طراحی، استقرار و اجرای نظام مدیریت کیفیت مرکز آمار ایران را بر اساس این چارچوب در دستور کار قرار داده و امیدوار است ضمن اجرای کامل آن در درون خود، در تعامل فعال با دیگر دستگاه‌های اجرایی تولیدکننده آمار کشور، نسبت به استقرار و اجرای آن در مجموعه نظام آماری کشور اقدام نماید.

۱-۴- اهداف تحقیق

در سرشماری‌های سنتی، منابع بسیار زیادی تدوین و شاخص‌های متعددی برای ارزیابی معرفی شده است اما یکی از چالش‌های حرکت به سوی سرشماری‌های ثبتی چگونگی ایجاد یک چارچوب برای ارزیابی کیفیت داده‌های ثبتی یا اداری است.

هدف اصلی این تحقیق دستیابی به یک چارچوب مدون به منظور ارزیابی کیفیت داده‌های ثبتی از لحظه ثبت در سازمان مربوطه تا بررسی نهایی جهت استفاده در سایر مراحل سرشماری ثبتی مینا است.

اهداف فرعی تحقیق نیز پاسخ به سؤالات زیر است:

- شاخص‌های ارزیابی چگونه تعریف شوند؟
- شاخص‌های ارزیابی در چه زمانی تهیه شوند؟
- پس از محاسبه شاخص‌های ارزیابی نحوه برخورد با داده‌های دارای اشکال یا با کیفیت کمتر از حد مشخص چه باشد؟

۱-۵- مراحل انجام تحقیق

مراحل انجام این گزارش، وابسته به شناخت دقیق سازمان‌های ارائه دهنده داده‌های اداری و فراهم بودن زیرساخت‌های قانونی دسترسی به داده‌ها و زیرساخت‌های حوزه IT، نحوه دریافت فایل‌های اداری، نحوه ارزیابی اولیه، شناسایی متغیرهای شناسایی و محاسبه شاخص‌های موردنیاز جهت ارزیابی کیفیت هر کدام از داده‌ها است. بنابراین می‌توان مراحل انجام این طرح را به صورت زیر بیان کرد:

- تعریف مفاهیم و روش‌شناسی ارزیابی کیفیت
- مطالعه تجربیات کشورها و توصیه‌های بین‌المللی
- شناسایی ابزارهای متداول کیفیت داده‌های ثبتی
- شناسایی ثبت‌های اداری برای ایجاد سرشماری ثبتی مبنا
- مروری بر سرشماری ثبتی مبنا (تعریف و مراحل ایجاد آن)
- تدوین چهارچوب ارزیابی کیفیت سرشماری ثبتی مبنا
- تهیه فرم‌های ارزیابی اولیه، ارزیابی کامل و مشخص کردن زمان و متولی تکمیل هر کدام
- مشخص کردن نحوه برخورد با خطاها و شاخص‌هایی که کیفیت پایین را گزارش می‌دهند.
- مزایا و معایب و بررسی کارایی چهارچوب تهیه‌شده

۰۷. روش

در این طرح، به بررسی روش‌های ارزیابی کیفیت، استخراج شاخص‌های مورد نیاز و تدوین چهارچوب ارزیابی کیفیت داده‌های سرشماری ثبتی مبنا پرداخته می‌شود. بدین منظور طبق گام‌هایی که در قسمت «مراحل انجام تحقیق» معرفی شد، هر کدام از بندها مطالعه و شاخص‌هایی متناسب با اهداف همان بند معرفی می‌شود. در انتها کارایی هر کدام از شاخص‌ها بررسی و در نهایت جهت قرار گرفتن در چهارچوب ارزیابی کیفیت و حد مجاز برای آن شاخص معرفی می‌شود. در تهیه و بررسی کارایی شاخص‌های ارزیابی کیفیت شاخص‌های متداول در ارزیابی کیفیت داده‌های ثبتی یا اداری جهت استفاده در سرشماری ثبتی مبنا به عنوان نمونه معرفی می‌شود. بانک اطلاعاتی یکی از دستگاه‌ها که داده‌های آن مربوط به سال ۱۳۹۵ است در بخش نهایی، به عنوان یک منبع ثبتی مورد ارزیابی قرار خواهد گرفت.

مطالعات تطبیقی، توصیه‌ها و استانداردهای بین‌المللی موجود در خصوص ارزیابی کیفیت داده‌های ثبتی

۱-۲- مقدمه

در این فصل تجارب کشورهای عضو کمیسیون اقتصادی اروپا مورد بررسی قرار می‌گیرد. هر یک از این کشورها در مسیر انتقال از روش‌های سنتی به شیوه‌های نوین جمع‌آوری اطلاعات، تجارب مفیدی دارند. برای افزایش سطح دانش، در مورد چگونگی عبور از این گذار و نیز نوع مشکلات عملی سر راه کشورها، بعضی از تجارب کشورهای عضو شبکه آماری اروپا بیشتر توضیح داده می‌شود. این نمونه‌ها براساس تجارب کشورهایی است که اخیراً از سرشماری سنتی تغییر جهت داده یا قصد انجام آن را در آینده نزدیک دارند لازم به ذکر است که مطالعات اشاره شده در این پژوهش، منعکس‌کننده وضعیتی است که در زمان آماده‌سازی دستورالعمل مذکور در اواسط سال ۲۰۱۷ در کشورها اتفاق افتاده است. تغییرات در متدلوژی و فرایندها ممکن است از آن زمان صورت گرفته باشد. با این وجود، کشورهای دیگر می‌توانند از این تجربیات سود ببرند.

۲-۲- مطالعات تطبیقی کشورها

۱-۲-۲- اتریش

اولین سرشماری به شیوه نوین در اتریش در سال ۱۸۶۹ انجام شد. پس از آن سرشماری‌ها در سال‌های ۱۸۸۰، ۱۸۹۰، ۱۹۰۰ و ۱۹۱۰ برگزار شد. بعد از جنگ جهانی اول در سال‌های ۱۹۲۳ و ۱۹۳۴ انجام شد و در سال ۱۹۳۹ آلمان سرشماری را برگزار کرد. از سال ۱۹۵۱ نیز سرشماری به صورت منظم و ده ساله برگزار شده است. در سرشماری‌های مذکور، مسئولیت برنامه‌ریزی، طراحی پرسشنامه‌ها، آموزش رده‌های سرشماری، استخراج و انتشار داده‌ها بر عهده سازمان آمار اتریش و اجرا بر عهده شهرداری‌ها بوده است. در سال ۱۸۹۰ برای تدوین

جداول از دستگاه‌های مجهز به کارت‌های پانچ استفاده کردند. در سال ۱۹۷۱ از جانپی و پرسشنامه‌هایی که توسط دستگاه قابل خواندن بود، استفاده شد. از دستگاه‌های اسکنر و نرم‌افزارهای تشخیص کاراکتر و همچنین کدگذاری خودکار در سرشماری ۲۰۰۱ استفاده شد. در این سال از ثبت‌های جمعیت محلی در پایگاه داده‌ی مرکزی استفاده شد.

آخرین سرشماری سنتی در سال ۲۰۰۱ انجام شد. در سال ۲۰۰۶ اولین آزمایش سرشماری ثبتی مبنا انجام شد. از سال ۲۰۰۹ آمارهای نیروی کار به صورت ثبتی مبنا جمع‌آوری می‌شود. اولین سرشماری ثبتی مبنا کامل در تاریخ ۲۰۱۱/۱۰/۳۱ انجام شد. در این سرشماری از ۷ ثبت پایه برای داده‌ها و ۸ ثبت برای مقایسه استفاده شد. در این سرشماری اطلاعات مربوط به نفوس (افرادی که مقیم اتریش هستند)، مسکن، بنگاه‌ها و اشتغال جمع‌آوری شد. لازم به ذکر است که جمع‌آوری اطلاعات مربوط به سؤالات: «زبانی که بیشتر صحبت می‌شود»، «دین»، «نحوه رفتن به مدرسه یا محل کار» و «مدت زمانی که برای حمل و نقل صرف می‌شود» با استفاده از ثبت‌ها امکان‌پذیر نبود و از طرح‌های نمونه‌گیری حاصل شد.

به این ترتیب کشور اتریش پیشرو در میان کشورهای جهان در تجربه سرشماری ثبتی مبنا است. این کشور نخستین سرشماری مبنا ثبتی خود را در سال ۲۰۱۱ انجام داد. بنابراین این کشور در بررسی ثبت‌های اداری و ارزیابی کیفیت داده‌ها تجربه غنی و ارزشمندی دارد. با توجه به منابع و داده‌های موجود در این زمینه فصل سوم تحقیق حاضر به طور کامل تجربه کشور اتریش را از نظر ارزیابی و کیفیت داده‌ها مورد بحث و بررسی قرار خواهد داد.

۲-۲-۲- ایرلند

اداره مرکزی آمار ایرلند^۱ پیشرفت قابل توجهی در ایجاد منابع اطلاعات اداری برای اهداف آماری انجام داده است (مک فلی و دان، ۲۰۱۴). این کار با توسعه یک زیرساخت داده ملی^۲ در ایرلند آغاز شد. به طور خلاصه، زیرساخت داده ملی یک چارچوب مفهومی است که ترویج استفاده کارآمدتری از اطلاعات در بخش دولتی ایرلند است. این ترویج و حمایت از استفاده از شناسه‌های رسمی برای افراد، کسب و کار و اموال در تمام سیستم‌های رسمی (به دنبال مدل Nordic) است. یک سیستم ثبت مرکزی جمعیت با اطلاعات در مورد محل زندگی افراد، پیش بینی نشده است (همان).

در ایرلند کد شناسایی شخصی یکتا وجود دارد که به طور کلی در همه سیستم‌های اداری مبتنی بر اطلاعات فردی (مدارس، رفاه، اشتغال و غیره) استفاده می‌شود. شماره شناسایی یکتا در هنگام تولد یا زمان ورود هر فرد به داخل این کشور به افراد تخصیص داده می‌شود. این شماره شناسایی در یک فایل اصلی یا ثبت اداری ذخیره می‌شود. این ثبت توسط اداره تأمین اجتماعی^۳ (وزارت کشور نظارت بر مدیریت رفاه اجتماعی را بر عهده دارد) به روز می‌شود. با این حال، نیازی نیست که همه سازمان‌های بخش دولتی برای به روز رسانی اداره تأمین اجتماعی ارائه دهند. به تازگی، یک پروژه برای خلاصه کردن فعالیت سالانه هر فرد در سیستم‌های مدیریت کلیدی عمومی با یک نشانگر بله یا خیر در ثبت فعالیت شخصی^۴ لحاظ شده است. ثبت فعالیت شخصی استفاده از یک کلید شناسه حفاظت

^۱ . Central Statistics Office

^۲ . National data Infrastructure

^۳ . Department of Social Protection

^۴ . Person Activity Register, PAR

شده^۵ برای کاهش خطر حفظ حریم خصوصی در حالی که قابلیت اتصال در طول زمان و در سراسر منابع داده را دارد، می‌باشد. منابع داده کلیدی اداری عبارتند از: تولد، سود کودکان^۶، تحصیلات (ابتدایی، ابتدایی، متوسطه، بالاتر و بیشتر)، اشتغال، بیکاری، حقوق بازنشستگی شغلی و رفاه اجتماعی (از جمله حقوق بازنشستگی). PAR سه هدف به شرح زیر دارد:

۱- شناخت تحلیل طولی جمعیت نسلی در سیستم‌های مختلف اداری؛

۲- شناخت ساختار جمعیت در طول زمان؛

۳- ارائه یک کلید اصلی خلاصه با توجه به منابع اطلاعاتی مختلف اداری برای بررسی امکان پروژه‌های مختلف بالقوه.

ثابت شده که تحلیل طولی جمعیت‌های خاص، از اهمیت قابل توجهی برخوردار است. مفهوم «نیازهای دولت متصل به داده‌های پیوسته»^۷ در کل بخش دولتی ایرلند ترویج پیدا کرده است. به معنای دقیق تعریف این اصطلاح، PAR یک ثبت نیست بلکه فقط یک مجموعه داده آماری از جمعیت^۸ است (همان).

تعدادی از شماره‌های شناسایی رسمی برای استفاده در کسب و کار، سیستم‌های مالیاتی، استخدام و سیستم ثبت شرکت وجود دارد. این شماره‌ها به منظور استفاده از منابع داده برای اهداف آماری در فهرست مشاغل اداره مرکزی آمار ایرلند مرتبط و در دسترس هستند. سازمان امور مالیاتی نیز با این شماره‌ها ارتباط دارد. در حال حاضر، برنامه‌ریزی شده تا استفاده از شماره‌های شناسایی رسمی برای مشاغل در سراسر بخش عمومی گسترش یابد (همان).

ارتباط بین فایل آماری کارفرما و مستخدم بر مبنای بازپرداخت مالیات کارکنان و کارفرما ایجاد شده است. این فایل تمام استخدام‌ها در ایالت مختلف را پوشش می‌دهد و ارتباطات پایه‌های تجاری و شخصی زیرساخت داده ملی را تسهیل می‌کند (همان).

در ایرلند اخیراً، سیستم کد پستی جدیدی راه اندازی شده است. کد پستی یکتا هر نامه‌ای را در ایالت شناسایی می‌کند. در حال حاضر استفاده از کد پستی اجباری نیست. احتمالاً چند سال طول خواهد کشید تا سیستم کد پستی به عنوان مرجعی منحصر به فرد و قابل دسترس در تمام سیستم‌های به کار رود. در ایرلند، بیش از ۳۰ درصد از آدرس‌ها منحصر به فرد نیستند و تنها با استفاده از نام فرد و اطلاع محلی پستی، نامه تحویل داده می‌شود. سامانه کدپستی و ثبت اصلی آن، چارچوب نمونه‌گیری و چارچوب آدرس اصلی برای بررسی‌های اجتماعی و عملیات سرشماری سال ۲۰۱۱ و ۲۰۱۶ است (همان).

شکاف‌های آشکاری در منابع داده برای انجام سرشماری با استفاده از داده‌های اداری مربوط به پوشش جمعیت و جایی که جمعیت آن ساکن است وجود دارد. علاوه بر تلاش مداوم برای دستیابی به داده‌های با کیفیت بهتر، این شکاف‌ها با توسعه فرایندهای تحول -مرحله دوم چارچوب- مورد رسیدگی قرار می‌گیرند، جایی که داده‌های اداری با استفاده از تکنیک‌هایی مانند مدل‌سازی و پیوند به یک ثبت آماری تبدیل می‌شوند (همان: ۳۸).

۲-۲-۱- پیشینه سرشماری

⁵. Protected Identifier Key

⁶. Children's benefit

⁷. Joined-up government needs joined-up data

⁸. Statistical Population Dataset

ایرلند به طور معمول دارای جا به جایی‌های جمعیتی و جریان‌های مهاجرتی زیاد است که هر ۵ سال یکبار نیاز به سرشماری دارد (۴/۶ میلیون نفر در سال ۲۰۱۱ در ۱/۶ میلیون مسکن زندگی می‌کنند). ایرلند سرشماری سنتی را تقریباً با تعداد ۴۰۰۰ آمارگیر انجام می‌دهد. به هر خانوار یک فرم سرشماری ارائه می‌شود تا توسط سرپرست خانوار تکمیل شود. برای تکمیل فرم حضور افراد در شب سرشماری ضرورت دارد. سپس فرم‌ها توسط آمارگیران جمع‌آوری می‌شوند تا توسط سازمان مرکزی اداره مرکزی آمار ایرلند پردازش شود. تجربه سال ۲۰۱۶ نشان داد که ارتباط با خانواده‌ها سخت‌تر شده است (همان: ۳۸).

۲-۲-۲-۲- اهداف سرشماری آینده

هنوز مشخص نیست که چگونه ممکن است سرشماری در ایرلند براساس منابع داده‌های اداری و طرح‌های موجود انجام شود اما اگر ایرلند بتواند چنین سرشماری را انجام دهد، ممکن است در مراحل زیر انجام دهد:

مرحله ۱: ایجاد یک فایل آدرس اصلی به روز شده که همچنین ویژگی‌های ساختمان را از سرشماری قبلی و داده‌های اداری برای کمک به بخش مسکن سرشماری جمع‌آوری کند. ممکن است امکان شناسایی وضعیت سکونت (خالی یا اشغال شده) مسکن‌ها براساس سوابق شرکت‌های ساختمانی نیز وجود داشته باشد (همان).

مرحله ۲: شناسایی همه افراد در ارتباط با سیستم‌های بخش دولتی برای سال سرشماری و خلاصه کردن آن در یک ثبت با یک شناسه شخصی. این ثبت فعالیت شخص (PAR) است که قبل از آن ذکر شده است. PAR حاوی لیستی از افرادی است که در یک سال خاص فعال هستند (همانطور که در رویکرد «نشانه‌های زندگی»^۹). منابع کلیدی داده‌های اداری در ایرلند عبارتند از: تولد؛ پرداخت سود کودکان؛ پایگاه داده ثبت نام مدارس ابتدایی؛ منابع تحصیلات عالی و تحصیلی؛ ثبت نام و جوایز؛ منابع استخدامی (از جمله خود اشتغالی) از مقامات مالیاتی؛ افراد ثبت شده در اجاره ملک؛ افرادی که از مزایای پزشکی استفاده می‌کنند (یا ثبت نام در یک سال داده شده)؛ رفاه اجتماعی و حقوق بازنشستگی شغلی / دولتی (همان: ۳۹).

مرحله ۳: استفاده از روش‌های آماری برای اصلاح کم پوششی و اضافه کردن یک ضریب اصلاح برای هر رکورد. روش‌های آماری ممکن است به بررسی نمونه‌های موجود یا یکی دیگر از منابع مستقل متکی باشند. در حال حاضر یکی از روش‌ها به عنوان یک منبع مستقل ثانویه برای ناشناس سازی در ثبت فعالیت با استفاده از روش (Capture-Recapture) در حال بررسی است (همان: ۳۹).

مرحله ۴: تخصیص هر فرد با آدرس یا کد پستی از فایل آدرس اصلی. از یک الگوریتم درخت تصمیم‌گیری مناسب استفاده می‌شود که قادر به ترکیب موقعیت‌هایی است که بیش از یک آدرس با شخص در منابع داده‌های اداری شناسایی می‌شود. روابط خانواری مشخص شده در منابع داده‌های اداری (مانند سود کودکان) نیز ممکن است با این الگوریتم تغذیه شود (همان).

مرحله ۵: از روابط خانواری با استفاده از روابطی که در منابع داده‌های اداری و افراد شناسایی شده در همان خانوار شناخته می‌شود (همان).

مرحله ۶: برآورد و صفات برای هر شخص در PAR با استفاده از طرح‌های موجود، منابع داده‌های اداری و روش‌شناسی مناسب است (همان).

مرحله ۷: تمام خروجی‌های سرشماری اکنون از PAR به ترتیب در مراحل فوق تکمیل شده است.

^۹ . Signs of life