

رئیس جلسه: آقای دکتر حسینزاده

دبیر جلسه: سعید عباسی

شماره جلسه: ۳۲

تاریخ جلسه: ۱۴۰۰/۰۹/۲۲

ساعت شروع: ۱۱:۰۰

ساعت پایان: ۱۳:۰۰

دستور جلسه:

- گزارش دستورالعمل پاکسازی داده‌ها توسط کمیته طراحی و روش شناسی
- پیگیری مصوبات جلسه قبل و سایر موارد

جلسه با تلاوت قرآن آغاز گردید. آقای عباسی ضمن خوش آمدگویی به اعضای شورای هماهنگی و تلفیق دستور جلسه را قرائت کردند. سپس جناب آقای دکتر حسینزاده رئیس شورای هماهنگی و تلفیق ضمن تشکر از حضور اعضای شورا، اظهار داشتند: از آنجا که نشست ستاد سرشماری ثبتی مبنا قرار است برگزار شود، نیاز است کمیته‌ها مباحث خود را برای جلسه ستاد مطرح کنند تا با دریافت مصوبات کامل برنامه‌ریزی دقیقی برای انجام سرشماری ارائه شود.

- دستور اول: ارائه گزارش دستورالعمل پاکسازی داده‌ها توسط کمیته طراحی و روش شناسی

در ابتدای جلسه آقای رنجی گزارشی از دستورالعمل پاکسازی داده‌های ثبتی شامل فرآیند پاکسازی داده‌ها، بررسی، تشخیص و اقدامات اولیه، تصمیم‌گیری در مورد مقادیر گمشده، نادرست، دورافتاده و رکوردهای تکراری و مستندسازی ارائه دادند ایشان اظهار داشتند پاکسازی داده‌ها، فرایند بهبود کیفیت یک مجموعه داده بوسیله تصحیح رکوردهای نادرست است. این اصطلاح مشخصاً به شناسایی و ادیت/جانپ/حذف رکوردهای ناقص، نادرست، دارای فرمت نامناسب، اضافه و نامربوط اشاره دارد. در تحلیل و کاربرد داده‌ها، وجود داده‌های باکیفیت تر مطرح است به خصوص وقتی داده‌ها از منابع مختلفی می‌آیند و استاندارد مشخص و یکسانی در این منابع استفاده نشده است، نیاز به چنین فرایندی اهمیت ویژه‌ای پیدا می‌کند. فرایند پاکسازی داده‌ها شامل بررسی، تشخیص و اقدامات اولیه تصمیم‌گیری در مورد مقادیر گمشده، نادرست، دورافتاده و رکوردهای تکراری، مستندسازی برای پاک سازی داده‌ها می‌باشد. اعتبارسنجی شماره ملی برای مجموعه داده، از آنجا که ابزار اصلی در شناسایی رکوردهای تکراری یا اتصال رکوردهای یکسان در مجموعه داده‌های متفاوت متغیر شماره ملی است، پیش از اتخاذ هر تصمیمی در خصوص رکوردهای تکراری اعتبارسنجی شماره ملی ضروری است و مجموعه داده ثبت احوال ملاک ارزیابی صحت شماره ملی در نظر گرفته می‌شود. پس از اعتبارسنجی شماره ملی، رکوردهای مشکل‌دار برای تعیین تکلیف به دستگاه اجرایی مربوطه ارسال می‌شود و در صورت حل نشدن مشکل باید در مرکز تعیین تکلیف شود و اگر بعد از ارجاع به دستگاه اجرایی مربوطه مشکل حل نشد، باید ابتدا رکوردهایی که قبل از این با فایل ثبت احوال لینک داده شده‌اند را کنار گذاشت و تعدادی رکورد باقی مانده که در مجموعه داده ثبت احوال پیدا نمی‌شوند، برای استفاده در مرحله جانپ حفظ می‌شوند. برای بررسی مجموعه داده اتباع، در صورتی که دو یا چند رکورد بر اساس شماره اتباع، یکسان باشند رکوردی که دارای اطلاعات کاملتری است حفظ و بقیه رکوردها حذف می‌شود.

سایر اعضا این موارد را مطرح کردند:

آقای محزون اظهار داشتند: تکلیف دستگاه است که داده‌های پاک را در اختیار مرکز آمار قرار دهد و بعضی موارد مربوط به داده دستگاه‌های اجرایی است که قادر به تعیین تکلیف آن نبوده و لازم است داده پاک به مرکز آمار برسد و در دستگاه‌ها چک شود و اگر مشکلی بود در فرآیند رفت و برگشت، داده پاک به دست مرکز آمار برسد. در آموزش و پرورش شماره اتباع وجود ندارد و مجموع داده‌های وزارت علوم رکوردها به دلیل چندگانه بودن سطح تحصیلی تکرار می‌شود.

آقای پیمان اظهار داشتند: باید تفاوت پاکسازی و ادیت را مورد توجه قرار داد در پاکسازی به دنبال اعتبارسنجی داده‌های دریافتی هستیم و باید ساز و کار نحوه تعامل با دستگاه‌ها را توجه کنیم و برای داده‌های دستگاه مختلف لازم است ساز و کار اینکه کدام داده پاک است را تعیین کنیم. همچنین موضوع مستند سازی برای استفاده بعدی مهم است. وجود نقشه برای انجام فعالیت‌ها نیاز است. وضعیت و ظرفیت موجود به ستاد ارائه و برای بالابردن ظرفیت دستگاه‌ها آنچه که نیاز است مطرح شود. داده‌های اصلی و کمکی را بر اساس فرآیند مشخص و تحت کنترل و کار کمیته‌های مربوط که قرار است بر روی آن کار انجام دهند مشخص کنیم.

آقای دکتر عزیزی اظهار داشتند: آن چه که مهم است بررسی داده‌های پایگاه هاست. در خصوص ایجاد پایگاه‌های موجود دستورالعمل استاندارد سازی برای دستگاه‌ها تهیه شود و زمان بهنگام سازی داده‌های دستگاه‌ها مشخص نیست و مرجع زمان ارسال داده‌ها اهمیت زیادی دارد.

خانم دکتر بنی هاشمی اظهار داشتند: پیشنهاد می‌شود جلساتی با کارگروه کیفیت، کمیته طراحی و روش شناسی داشته باشیم. کمیته فاوا هم دستورالعمل‌ها را برای فایل دستگاه‌ها داشته‌اند. لازم است دستورالعمل یکسانی برای کمیته‌هایی که با فایل‌ها کار می‌کنند وجود داشته باشد و این گزارش تبدیل به فلوجارتی برای اجرا شود. این گزارش اشاره به کمیته ندارد. همچنین آدرس و مکانی که افراد ساکن هستند باید مورد توجه جدی قرار گیرد و کیفیت داده مکانی افراد برای سال ۱۴۰۵ بسیار مهم است. برای دریافت داده‌های دانشگاه آزاد شرط محرمانگی، امنیت و عدم کار اقتصادی گذاشته شده است.

آقای نصیری اظهار داشتند: هر سه کمیته فایل پاکسازی داده‌ها را بررسی کنند و همپوشانی‌ها را رفع نمایند تا بازخورد مناسبی به دستگاه‌ها داده شود. در موضوع پاکسازی داده باید کدپستی مورد توجه قرار گیرد. در جلسه‌ای که با ثبت احوال برگزار می‌شود باید درخواست یکسان سازی پایگاه‌ها را کرد که نیاز به همکاری دستگاه‌ها و پیگیری دارد.

آقای محزون بیان کردند: سه فاز مورد نیاز است: اول ورود داده و دریافت داده که از نظر استانداردهای کیفی دریافت داده باید مورد توجه قرار گیرد دوم بررسی کمیته تامین داده از لحاظ دیتا مدل و استانداردهای در نظر گرفته شده ببیند مطابقت دارد یا خیر و در مرحله سوم از نظر موضوعی و نسبت‌ها و سازگاری جمعیت شناختی بررسی شود.

آقای فاضلی اظهار داشتند: گردش کار داده‌های ارائه شده و ترکیب کار در کمیته‌ها و کارگروه‌های مربوطه مشخص شده و موضوع تهیه شناسنامه در دریافت و بررسی فایل‌ها انجام شده است. ناسازگاری در فایل داده‌های دریافتی وجود دارد که لازم است شیوه نامه‌ای برای این موضوع دیده شود. در کیفیت اول باید کل شاخص‌های کیفیت بررسی شود بعد انطباق دادن داده‌های پایگاه بر اساس دستورالعمل تهیه شده است و در آخر انطباق دادن پایگاه‌های اداری است که در استاندارد سازی داشتن ترتیب و توالی بسیار مهم است و نکته دیگر اینکه که فایلی که تشخیص داده شد برای انجام کار مناسب نیست آیا لازم است در چرخه پاکسازی قرار بگیرد؟

آقای کرامتی اظهار داشتند: فرآیند تهیه و ایجاد ثبت آماری دیده شده بود آیا فعالیت‌ها بر اساس این فرآیند انجام شده است.

آقای رنجی اظهار داشتند: در حال انجام کار بر اساس برنامه زمانبندی تایید شده هستیم و ارائه این دستورالعمل گویای این است که پاکسازی بر اساس مکانیزم تعریف شده انجام می‌شود. این وظیفه مرکز آمار نیست که افراد را به مکان منتسب کند و بهنگام سازی در حال حاضر بر عهده ثبت احوال است و موضوع کد پستی همیشه مطرح است که لازم است افراد به مکان منتسب شوند. نکته این است که در فایل‌ها پوشش کامل صورت نگرفته است و بسیاری از اطلاعات دریافتی در خصوص انتساب کد ملی به کد مکان است. در دستورالعمل پاکسازی هر کدام از فایل‌ها به طور جداگانه بررسی می‌شود و در خصوص پوشش داده‌های ثبت احوال باکم پوششی مواجه شدیم. موضوع مرجع زمانی داده‌ها خیلی مهم است و باید توجه شود.

آقای پیمان اظهار داشتند: پیشنهاد می‌شود وضعیت موجود و ظرفیت فعلی را به ستاد ببریم و آنچه که می‌تواند به بالا بردن ظرفیت‌ها کمک کند را مطرح نماییم. داده‌های اصلی و کمکی را مشخص کنیم و بر اساس فرایند تعریف شده و تحت کنترل، کمیته‌ها و کارگروه‌ها فعالیت کنند.

خانم قادری اظهار داشتند: لازم است قبل از کار با داده‌ها این موضوع مورد توجه قرار گیرد که چه موقع شروع به پاکسازی داده شود نکته این است برای کار با داده‌ها برای انطباق داده‌ها لازم است پاکسازی انجام شود و موضوع انجام پاکسازی توسط دستگاه‌ها و مرکز باید مورد توجه قرار بگیرد.

خانم دکتر بنی‌هاشمی اظهار داشتند: رسالت مرکز انجام سرشماری و کمک به دستگاه‌ها برای تولید داده‌های بهتر است. از داده‌های دریافتی چند حالت پیش آمده اول اینکه آنچه که دریافت کردیم با آنچه که مورد نیاز است مطابقت ندارد، حالت دوم فایل درست است اما به دلیل عدم شناخت ما از فایل به مشکل برخوردیم، سوم خروجی گرفته شده درست است اما برای استفاده کامل نیست و حالت آخر اینکه لازم است روی داده‌ها توسط دستگاه‌ها کار انجام شود و این ممکن است طولانی شود و نیاز به برنامه‌ریزی برای انجام دارد. لازم است دریافت فایل بر اساس استاندارد انجام شود و شاخص‌های استاندارد و کیفیت باید مورد توجه قرار گیرد.

خانم میرزایی اظهار داشتند: پیشنهاد می‌شود دستورالعمل و فلوچارتی که تهیه می‌شود در یکی از دستگاه‌ها آزمایش شود. کمیته فاوا با دستگاه‌ها و کمیته‌ها در تعامل برای تهیه دستورالعمل‌ها است.

آقای نصیری اظهار داشتند: داده‌های دریافت شده از دستگاه‌ها با آنچه که توافق شده بود متفاوت است و دلیل آن زیرساخت دستگاه است که ارائه بازخوردها خود به خود سبب ارتقا زیرساخت دستگاه نیز خواهد شد.

خانم پوررضا اظهار داشتند: نتایج آزمایش میدانی و همچنین اقدامات انجام شده روی دیتا مطرح است تا الزامات ثبتی مبنای سال ۱۴۰۵ مشخص و در ستاد تعیین شود. اما آنچه که لازم است برنامه سالانه تا ۱۴۰۵ و موارد کلان و مشکلات و اقدامات مورد انتظار است. پیشنهاد می‌شود در جلسه ستاد اول مصوبات جلسه قبلی و اقدامات کمیته‌ها و آزمایش سرشماری همراه با تصویر موجود و چالش‌های بدست آمده مطرح شود و در ادامه روش‌های تطبیق داده‌ها با داده‌های ثبتی ارائه گردد و فلوچارت انجام فعالیت‌ها با تاکید بر تکالیف دستگاه‌ها ارائه گردد.

در پایان جلسه دکتر حسین زاده ضمن تشکر از همکاری که دستورالعمل پاکسازی داده را تهیه کردند اظهار داشتند: جلسه امروز نشان از فعالیت همکاران بر روی دیتای دستگاه‌ها است و این یعنی ما از کلیات عبور کردیم که بسیار ارزشمند است. در کارهای انجام شده در دیگر کشورها روش‌هایی وجود دارد که کار مرکز هم خارج از آن شیوه‌ها نیست طبیعی است که نمی‌شود ثبت اداری را مستقل از ثبت آماری دید، ثبت‌های اداری و ثبت‌های آماری در تلاقی هم هستند و تعامل با دستگاه‌ها برای دریافت داده باید به نتیجه برسیم. مسائل مربوط به اتصال کد پستی و کد ملی یک مسئله حقوقی است که موارد مستقل از هم نیستند و کمیته‌ها نیز باید با هم فرآیند را انجام دهند تا نقاط مشترک را پیدا کنند و حتی ممکن است در آینده نیاز باشد کمیته‌ها دستخوش تغییر هم بشوند. لازم است فلوچارتی تهیه شود و مشخص کند یک ثبت اداری چگونه به ثبت آماری تبدیل می‌شود و تکالیف کمیته‌ها نیز مشخص شود و

این فلوجارت با توجه به تجربیات و الزامات باید پویا باشد. همچنین لازم است استاندارد مورد نیاز برای تولید داده دستگاهها تهیه شود که اقداماتی در این راستا انجام شده است و این استاندارد به دستگاهها ابلاغ شود تا رعایت کنند. ایشان در خصوص اهمیت جلسه ستاد و مصوبات آن پیشنهاد ترسیم وضع موجود ارائه خلاصه و راهکارهای پیشنهادی را مناسب دانستند که لازم است در قالب فلوجارت نیز ارائه شود.

مصوبات:

بازه زمانی	متولی	مصوبه
-	کمیته تامین داده با همکاری کمیته طراحی و روش شناسی و کمیته فاوا	مقرر شد فلوجارتی بر اساس برنامه زمانبندی کلان مصوب شورا به عنوان نقشه راه انجام کار و تبیین چگونگی تبدیل ثبت اداری به ثبت تهیه شود تا در جلسه ستاد ارائه شود. این فلوجارت با توجه به تجربیات و الزامات لازم است پویا باشد.

دستور و زمان جلسه آینده متعاقبا اعلام می شود.

شرکت کنندگان:

اعضاء	آقای دکتر حسین زاده	آقای میرفلاح نصیری	آقای دکتر وزیر	آقای دکتر فرامرزی	آقای دکتر شباک	خانم میرزایی	خانم دکتر بنی هاشمی	آقای محزون
زمان	۲:۰۰	۲:۰۰	۲:۰۰	-	۱:۰۰	۲:۰۰	۲:۰۰	۲:۰۰
اعضاء	آقای دکتر اصغری	خانم بنه	آقای رحیمی	آقای پیمان	آقای عباسی	آقای نصحی	آقای دکتر عزیزی	آقای رنجی
زمان	-	۲:۰۰	۲:۰۰	۲:۰۰	۲:۰۰	۲:۰۰	۲:۰۰	۲:۰۰
اعضا	آقای فاضلی	خانم پور رضا	خانم فقیه	آقای فرحزادی	آقای کرامتی	خانم قادری	آقای جعفری	خانم شاکری
زمان	۲:۰۰	۲:۰۰	۲:۰۰	۲:۰۰	۲:۰۰	۲:۰۰	۲:۰۰	-