

استراتژی ادارات ملی آمار در مواجهه با دنیای دیجیتال، حرکت به سمت استراتژی داده هوشمند^۱

۱۵ ام جلسه کمیته آمار و سیاست‌های آماری، سازمان توسعه و همکاری اقتصادی^۲، ۲۰ و ۲۱ ژوئن ۲۰۱۸، ژنو

ترجمه: سحر صاحبی

با گسترش اینترنت و شبکه‌های اجتماعی، انفجاری از داده‌های جدید پدید آمده است. این داده‌های جدید حوزه‌های متفاوتی در تجزیه و تحلیل‌های آماری فراهم آورده و در عین حال چالش‌هایی را نیز ایجاد کرده است و زمینه‌های مهمی را برای بهبود تجزیه و تحلیل‌های آماری ارائه کرده که می‌تواند اطلاعات بهتری را برای در اختیار سیاست‌گذاران و شهروندان قرار دهد. ادارات ملی آمار باید با این دنیای دیجیتال همراه شوند. از فرصت‌ها استفاده کنند و در مواجهه با این دنیای جدید داده، استراتژی‌های جدیدی اتخاذ کنند.

محیط جدید، تقاضاهای جدید و چالش‌های آن، پیچیده و پویا هستند.

این یادداشت، اجزای یک مدل مفهومی را بر اساس رویه‌های دنبال شده توسط سازمان توسعه و همکاری اقتصادی در خصوص توسعه "چارچوب داده هوشمند" ارائه می‌دهد و تلاش می‌شود به سوالات زیر پاسخ دهند:

- ادارات ملی آمار^۳ چگونه فرصت‌ها و چالش‌هایی را که در ارتباط با محیط دیجیتال جدید دارند، مشخص می‌کنند - آیا آن‌ها استراتژی‌های جدید دارند و اگر دارند اجزا اصلی آن‌ها کدام است؟

- چگونه ادارات ملی آمار، استراتژی‌ها را پیاده‌سازی کرده و تغییر فرهنگی لازم را در مدل کسب و کار خود ایجاد می‌کنند؟

- چگونه با "اکوسیستم‌های آماری"^۴ در سطح ملی و بین‌المللی مواجه می‌شوند؟ و چه مناطقی برای همکاری در این "اکوسیستم‌ها" وجود دارد؟

- آیا در مورد چارچوب داده هوشمند توصیف شده در این سند بازخوردی وجود دارد؟

۲-۱- تقاضای جدید^۵

در عصر دیجیتال و با فراهم شدن دسترسی آسان مردم به اطلاعات به موقع در امور زندگی، انتظارات نسبت به آمارها نیز تغییر کرده است. حداقل چهار ویژگی، این تقاضای جدید را نشان می‌دهد:

۱- تعیین حداقل زمان برای عرضه داده به بازار، متقاضی‌های بسیاری دارد. در جوامع امروز پذیرش تاخیر در ارائه آمارهای رسمی وجود ندارد و سیاست‌گذاران و شهروندان می‌خواهند ارائه داده در زمان واقعی و با روند جامعه هماهنگ شود و از قبول آمارهایی - که مانند اخترشناسان، پدیده‌هایی را نشان می‌دهند- که سال‌ها پیش اتفاق افتاده‌اند، اجتناب می‌کنند. مردم انتظار تجزیه و تحلیل سریع از سیاست‌ها و مشاهده تأثیر آنها بر زندگی‌شان را دارند. شتاب ارائه اطلاعات در جهان دیجیتال به آمارها نیز منتقل شده است.

۲- دقت در جزئیات: علاوه بر به موقع بودن، این انتظار وجود دارد که داده‌ها دقیق و با جزئیات باشند. برای مثال داده‌ها در سطوح محلی ارائه شود («وضعیت جامعه من چگونه است؟») و تفاوت گروه‌های اجتماعی-اقتصادی مختلف نشان داده شود و پوشش گسترده‌ای از موضوعات

^۱ Which Strategies for NSOs in the Digital Era? Towards 'Smart Data' Strategies

^۲ OECD

^۳ NSO

^۴ statistical ecosystems

^۵ The new demand

اجتماعی، اقتصادی و محیطی زیستی وجود داشته باشد. بسیاری از فعالیت‌های خصوصی مبنی بر تولید شاخص‌ها برای استفاده محلی و روزمره نیز به این جنبه از تقاضای جدید اشاره دارند.

۳- تلاش در راستای بهبود کیفیت، نظر به افزایش روز افزون عرضه داده، این انتظار از طرف کاربران وجود دارد که ادارات ملی آمار به داده‌های بزرگ حساس باشند و به داده‌های خامی که از طریق شبکه‌های اجتماعی و رسانه‌ها و منابع جدید ایجاد می‌شوند، اعتبار داده شود. اعتبار بخشی به داده‌ها به این معنی است که به داده‌ها از طریق ایجاد ساختار، فراداده، روش شناسی، مستندسازی و تجزیه و تحلیل، اعتبار آماری داده شود. ایجاد استانداردهای اخلاقی (شفافیت و احترام به حریم خصوصی داده‌ها)، ایجاد قابلیت خواندن داده توسط ماشین به ایجاد «اعتماد» (که یک دارایی گرانبها و کمیاب است) منجر می‌شود.

۴- با در نظر گرفتن این امر که طیف وسیعی از داده‌های جدید، خدمات و تولیدات آماری وجود دارد، انتظارها در سطوح مختلف کارشناسان (دانشجویان، سازمان‌های مردم‌نهاد^۶ و مشاغل) افزایش پیدا کرده است.

۳-۱- چالش‌ها

۵- تشکیلات، ساختار و رویکرد ادارات ملی آمار، در زمینه تولید آمار ممکن است با تقاضاهای جدید منطبق نباشد. در یک ساختار سنتی فرایند جمع‌آوری داده عمودی است، داده از یک کارگاه یا سازمان جمع‌آوری می‌شود و به مرحله انتشار می‌رسد. این شیوه عمودی گردآوری داده با منحصربفرد بودن آن مشخص می‌شود (مانند سرشماری) و ارزش آن به کیفیت قابل اعتماد داده است ولی این شیوه تنها به یکی از تقاضاهای جدید که ادارات ملی آمار با آن مواجه هستند پاسخ می‌دهد اما ادارات ملی آمار با این چالش مواجه هستند، تا جریان داده‌ای خود را با بهره‌برداری از چندین منبع داده‌ای جدید تغییر دهند. اتخاذ تکنیک‌های جدید علوم داده، دستیابی به پلتفرم‌های جدید، ساخت الگوریتم‌های در دسترس کاربران، "زنجیره عمودی تامین داده" را به یک "زنجیره یکپارچه تامین داده" تغییر می‌دهد و زمینه را برای آمادگی مواجه با منابع و روش‌های جدید فراهم می‌کند.

۶- ادارات ملی آمار به منظور غلبه بر شیوه‌های قدیمی تولید آمار و پاسخ به تقاضاهای جدید تحت فشار هستند. این جریان ممکن است با مدل فعلی کسب و کار امکان پذیر نباشد.

- شکاف در منابع داده: در حالی که ادارات ملی آمار، به جمع‌آوری داده‌های اصلی خود (مانند سرشماری، نیروی کار و ...) می‌پردازند، ایجاد ظرفیت برای ترکیب داده‌ها از منابع دیگر، ایجاد پایگاه‌های داده بر اساس داده‌های جدید، ارزش و اهمیت بسیاری پیدا کرده است. ادارات ملی آمار، تمرکز بیشتری بر روی داده‌های بازی که تولید می‌کنند پیدا کرده‌اند؛ ادارات ملی آمار نیاز دارند به کاربران داده‌هایی که توسط شخص یا سازمان‌های سوم تولید می‌شود، تبدیل شوند. اگر این اتفاق رخ دهد، دسترسی به داده‌های جدید یک موضوع استراتژیک می‌شود که نیاز به فرمول‌بندی از طریق یک استراتژی جهانی تأمین داده دارد. این موضوع ممکن است یک فعالیت جدید برای بسیاری از ادارات ملی آمار باشد و نیاز به یک تغییر اساسی در فرهنگ ادارات ملی آمار داشته باشد - از تولید کننده محتوای داده به تولید کننده و کاربر داده - . چنین تغییری شامل ایجاد شراکت‌های جدید، به ویژه با بخش خصوصی می‌شود.
- شکاف پلتفرم داده: بهره‌برداری از این فرصت‌ها نیازمند سرمایه‌گذاری در پلتفرم‌های جدید برای ایجاد دریاچه داده^۷ و ارائه خدمات علوم داده است. پلتفرم‌ها از راه دور (در فضای ابری) میزبانی خواهند کرد، که نیازمند تطبیق مدل‌های مالی با فناوری اطلاعات (از هزینه‌های سرمایه‌ای به مدل‌های هزینه‌های عملیاتی) و اجرای سیاست‌های حفظ حریم خصوصی است.

^۶ NGO

^۷ Data lakes

تکنیک‌های هوش مصنوعی نوظهور کار ما را تغییر خواهند داد؛ بسیاری از فعالیت‌های موجود از بین خواهند رفت و شغل‌ها تغییر شکل خواهند یافت. بلاک‌چین‌ها شیوه‌های اعتماد سازی به داده را مجدداً تعریف می‌کنند و توسعه‌های دیگری به وجود خواهد آمد که هنوز پیش بینی آن‌ها دشوار است. تمام این موارد روند از بین رفتن (زنجیره‌های ارزش) داده‌های موجود تسریع کرده و منجر به رویکرد استراتژیک به منظور یکپارچه‌سازی پلتفرم‌های داده خواهد شد.

- شکاف در مهارت‌های داده‌ای: یکی از چالش‌های موجود کمبود کارشناسان علوم داده است. اپراتورهای آماری، باید بیشتر به کارشناس علوم داده تبدیل شوند و وظایفی مانند منبع‌یابی داده، مدل‌سازی داده، کاوش داده و تجزیه و تحلیل داده را بر عهده گیرند. رفع این شکاف نیازمند وجود استراتژی مهارت‌های داده و تجدید نظر در ساختارهای موجود است.

-۷ با توجه به سرمایه‌گذاری‌های مورد نیاز برای پوشش سه شکاف ذکر شده در بالا، ادارات ملی آماری باید به تغییرات ساختاری خود بیندیشند و مدل کسب و کار متناسب را پیاده کنند. اعمال چنین تغییراتی شامل ادارات ملی آمار و سازمانهای ملی و بین‌المللی نیز می‌شود.

۴-۱- چارچوب و استراتژی

۸- فضای جدید، تقاضاهای جدید و چالش‌های مرتبط با آن پیچیده، پویا و مرتبط هستند. گذار از این وضعیت نیازمند اعمال استراتژی‌ها و چارچوب‌های استراتژیک مشخص است تا اجزای مختلف را به یکدیگر مرتبط کند، انتقال منابع را توالی بندی کند و حرکت به سمت جلو را سرعت بخشد. در اینجا تعدادی سوال پیش می‌آید:

- ادارات ملی آمار چگونه فرصت‌ها و چالش‌هایی را که در محیط دیجیتال جدید روبه‌رو می‌شوند، توصیف می‌کنند؟ آیا آن‌ها استراتژی‌هایی دارند و اگر دارند، مؤلفه‌های اصلی آن‌ها چیست؟

- ادارات ملی آمار چگونه استراتژی‌ها را اجرا می‌کنند و تغییرات لازم در مدل کسب و کار خود را پیاده سازی می‌کنند؟

- ادارات ملی آمار چگونه تحولات «اکوسیستم‌های آماری» را در سطح ملی و بین‌المللی می‌بینند که به اجرای استراتژی‌های خود کمک کند؟

۹- در بخش بعدی به رویکرد سازمان همکاری و توسعه اقتصادی برای پاسخ به این سوالات پرداخته شده است:

چارچوب داده هوشمند

۱۰- در چند سال اخیر، تحلیلگران و آمارشناسان سازمان همکاری و توسعه اقتصادی به طور فزاینده از منابع داده غیر معمول مانند داده‌های مکانی، اطلاعات رسانه‌های اجتماعی یا داده‌های جمع آوری شده از وبسایت‌ها در کار خود استفاده کرده‌اند. حدود ۱۰۰ پروژه مورد تجزیه و تحلیل قرار گرفتند. در همان زمان، تقاضاهای جدیدی مطرح شد که توسعه‌های فنی و ضرورت تجدید نظر در شیوه‌های جمع‌آوری، پردازش و انتشار سنتی داده‌ها را ایجاد کرد. سازمان همکاری و توسعه اقتصادی با ارائه مفهوم «چارچوب داده هوشمند»^۸، روشی را برای ساختاردهی استراتژیک در این زمینه مطرح کرده است.

۱۱- چرا "داده‌های هوشمند" نسبت به "مه داده"^۹ ارجحیت پیدا کرده است؟ در پروژه‌هایی که مورد تجزیه و تحلیل قرار گرفته، ارزش لزوماً از این ناشی نمی‌شود که داده‌ها «بزرگ» باشند؛ بلکه ارزش، به زوایای خلاقانه داده‌ها مرتبط می‌شود که منابع داده‌ای موجود (کوچک و بزرگ) را به شیوه‌های جدید ترکیب می‌کند، همچنین از منابع غیر معمول بهره می‌برد و روش‌ها و الگوریتم‌های جدیدی را پیش می‌برد. ترکیب رویکردهای آماری سنتی با تکنیک‌های جدید علوم داده به پروژه‌های آماری ارزش می‌دهد. در این رویکرد «داده هوشمند» را می‌توان به عنوان یک پیشنهاد ارزشمند در نظر گرفت که آمارشناسان و داده‌شناسان را در کنار هم قرار می‌دهد.

^۸ Smart Data Framework

^۹ Big data

۱۲- «چارچوب داده هوشمند» حوزه پیوند خورده‌ای است که سازمان‌ها باید در آن سرمایه‌گذاری کنند تا به نتایج مورد نظر دست پیدا کنند. ترکیبی از زنجیره ارزش و اکوسیستم‌هایی است که باید ایجاد شوند، تا رویکردهایی را که بیشتر به یک پروژه یا تیم محدود شده‌اند را یکپارچه‌تر در تمام اجزای سازمان به اجرا درآید. با این شیوه سرعت پذیرش روش‌های نوظهور افزایش خواهد یافت و فرصت‌هایی برای رشد و گسترش همه جانبه داده درون و خارج سازمان ایجاد خواهد شد.

۱- ایجاد شواهد جدید برای سیاست‌گذاری: سیاست‌هایی که اتخاذ می‌شود باید نوآورانه باشد. در بیشتر مواقع رویکردهایی که سازمان‌ها در قبال مه داده دارند، منفعت طلبانه است («این داده‌ها را دارم، چه کاری می‌توانم با آنها انجام دهم؟») یا رویکردهایی را انتخاب می‌کنند که ابعاد محدودی دارد («چگونه مسائل حریم خصوصی را حل کنم؟») این موارد مهم و حیاتی هستند ولی نباید بازدارند باشند. دیدگاه سازمان همکاری و توسعه اقتصادی این است که رویکرد باید مبتنی بر ارزش باشد: سیاست ما برای حل مشکلات چیست؟ کدام مشکلات را نمی‌توانیم با منابع داده‌ای موجود و روش‌های فعلی حل کنیم؟ کدام محصولات جدید می‌توانند به حل مسائل کمک کنند؟

۲- منابع (جدید) داده: دسترسی به داده مانعی اساسی است، به ویژه زمانی که داده‌های جدید در اختیار بخش خصوصی است و استفاده از آنها تحت تاثیر حریم خصوصی قرار دارد. دسترسی به داده می‌تواند به صورت مستقیم یا غیرمستقیم («داده به صورت کد باشد») باشد. ادارات ملی آمار تولید کننده داده هستند ولی معمولاً در دستیابی به منابع داده مشکل دارند. در آینده، استفاده از منابع داده خارجی احتمالاً مهم‌تر از داده‌هایی خواهد بود که خود سازمان‌های آماری تولید می‌کنند.

۳- توسعه پلتفرم داده هوشمند برای آمارشناسان، داده‌شناسان و تحلیل‌گران شامل کار با داده‌ها به شیوه‌ای کارآمد و خلاقانه‌تر، دسترسی آسان تر به داده، کاوش در آن‌ها و به اشتراک‌گذاری آمار است. این پلتفرم هسته مرکزی داده را تشکیل می‌دهد که داده‌های آماری را مدیریت کرده و بستری برای جمع‌آوری انواع داده است. این پلتفرم شبکه‌ای از دیگر پلتفرم‌ها است که به اشتراک‌گذاری الگوریتم‌ها و داده‌ها، و تسهیل در تحقیقات و نوآوری‌ها داده می‌پردازد.

۴- ترویج اکوسیستم‌های چند رشته‌ای (شامل حوزه‌های مختلف سیاست، دیجیتال و داده)، موضوعات مشترکی را در منابع داده (مانند محیط زیست، کشاورزی، شهرسازی که بر روی یک مجموعه داده‌های مکانی کار می‌کنند ایجاد می‌کنند) به اشتراک می‌گذارند. اکوسیستم‌ها می‌توانند در سطوح مختلف (درون سازمان‌ها و موسسات و در سطوح ملی و بین‌المللی) توسعه یابند. اکوسیستم‌ها می‌توانند با شهروندان نیز ارتباط برقرار کنند، تا از طریق کانال‌های دیجیتال داده‌های خود را به اشتراک بگذارند.

۱۳- مشارکت در این زمینه به ساخت داده‌های هوشمند در زمینه حقوق و اخلاق، کیفیت و روش‌شناسی، فناوری، حاکمیت و توانایی مواجهه با چالش‌های مرتبط با کمک می‌کند. چالش‌های حقوقی و اخلاقی اهمیت ویژه‌ای دارند، همانطور که در یک دهه گذشته حفظ حریم خصوصی مرتبط با داده‌های اداری و دیگر داده‌ها از چالش‌های مهمی بوده اند که با آن مواجه بودیم.

مدل‌های کسب و کار برای منابع داده چیست؟

۱۴- از سالیان گذشته ادارات ملی آمار، داده را از طریق طرح‌های آمارگیری یا سرشماری (کارگاه و خانوار) یا اخذ داده از سازمان‌ها و موسسات (به عنوان مثال، ادارت آمار ملی داده‌ها را از وزارت‌خانه‌ها، آژانس‌ها و دولت‌های محلی جمع‌آوری می‌کنند؛ یا بانک مرکزی داده‌ها را از مؤسسات مالی اخذ می‌کنند) جمع‌آوری کرده‌اند. همواره تلاش ادارات ملی آمار بر بهبود روش‌های گزارش‌دهی و دستیابی به منابع ثبتي به منظور دستیابی به داده و کاهش موانع در دستیابی به داده‌ها (به ویژه در بخش کسب و کار) است.

۱۵- "طوفان داده‌ای"^{۱۰} منجر به ظهور منابع جدید داده شده است، (الف) به ویژه با گسترش دنیای دیجیتال که خدمات خود را برای صدها میلیون مشتری ارائه می‌دهند و عمومی شدن «اینترنت اشیاء»؛ و (ب) نظر به توسعه‌های خوب صورت گرفته در حوزه

^{۱۰} Data deluge

تولید داده، تغییرات شگرفی در حوزه‌های فناوری و کیفیت و کمیت داده‌های تولیدی و تفسیر پذیری داده با بکارگیری تکنیک‌های جدید علوم داده (مانند تصاویر ماهواره‌ای با رزولوشن بالا و تکنیک‌های تشخیص الگو) صورت گرفته است.

۱۶- دسترسی به داده، بسیار اهمیت دارد، اما با افزایش منابع، دسترسی به داده پیچیدگی بیشتری پیدا می‌کند.

۱۷- ادارات ملی آمار باید توانایی و مهارت‌های لازم، برای رویارویی با پیچیدگی‌های این نوع داده، پویایی آن و مدل‌های تجاری متعدد آن را بدست آورند.

- تنظیمات^{۱۱}: ایده ای که مطرح می‌شود این است که (مانند رگولاتوری تلفن همراه) داده‌ها از طرف تولید کنندگان اصلی آن در اختیار عموم قرار گیرد یا دسترسی به داده فراهم شود و داده به جای اینکه از کارگاه‌ها یا خانوارها اخذ شود از منبع اصلی آن گرفته شود مانند اخذ داده از اوبر، آمازون و ... به عنوان مثال برای جمع آوری داده‌های هزینه درآمد خانوار به جای مراجعه مستقیم به خانوار از اطلاعات حساب‌های بانکی یا کارت‌های اعتباری آن‌ها استفاده شود. بکارگیری چنین نظام‌های هدفمندی، ادارات ملی آمار را در مواجهه با طوفان داده، قدرتمندتر میکند.
- داده باز^{۱۲}: داده‌های خصوصی و عمومی به منظور شفافیت و افزایش در نوآوری به رایگان به اشتراک گذاشته می‌شوند. فعالیت‌های مربوط به داده باز بیشتر به سازمان‌های عمومی مانند شهرداری‌ها برمی‌گردد که داده تنها زیرمجموعه آن‌ها است. ادارات ملی آمار می‌توانند تجربیات خود را در افزایش کیفیت آمار به سایر سازمان‌ها ارائه دهند تا بتوان داده‌های بهتری برای سیاست گذاری ارائه داد.
- داده‌های تجاری: ایجاد درآمد از طریق فروش داده در حال گسترش است و ساختار بازارهای مرتبط با فروش داده دستخوش تغییراتی است. تا سال ۲۰۲۰ تغییرات اساسی در ساختار بازار داده رخ خواهد داد و بازیگران اصلی با مقیاس بزرگ و توانایی سرمایه‌گذاری پلتفرم‌های اصلی، بازیگران کوچکتری که قادر به توازن نگه داشتن نیستند را خریداری خواهند کرد.^۱ (گارتنر، ۲۰۱۸). گروه‌هایی که در زمینه تبادل (فروش) داده فعالیت می‌کنند باید با اداره آمار ارتباط داشته باشند. این امر فرصتی را برای سازمان‌ها فراهم می‌کند تا در سطوح ملی و بین المللی با یکدیگر همکاری کنند و فرصت‌های موثری را برای مذاکری به ویژه در حوزه‌های رقابتی ایجاد کرده فرصت‌های همکاری را ایجاد کنند.
- مشارکت داده^{۱۳}: مشارکت بخش خصوصی و دولتی فرصت‌های خوبی را مانند به صرفه بودن، به موقع بودن در ارائه شاخص‌های جدید ارائه می‌کند. با این حال چالش‌هایی مانند مشکلات فنی، محرمانگی، نبود انگیزه وجود دارد. پروژه‌های مختلفی تعریف شده است که می‌توان آن‌ها را به ۴ گروه دسته بندی کرد. ۱- تولید آمار توسط ارائه کننده داده ۲- ارائه داده‌های شخصی یا خصوصی به کاربران نهایی ۳- ارائه داده‌های شخصی به شخص ثالث برای پردازش و تحلیل ۴- برون سپاری وظایف ادارات ملی آمار
- جمع سپاری^{۱۴}: به اشتراک گذاری داده می‌تواند منجر به بهبود شرایط زندگی شود. جمع سپاری می‌تواند شهروندان، گروه‌های کاری و تجاری را هدف قرار دهد. با گسترش دنیای دیجیتال و داده‌های وسیعی که از طریق گوشی‌های تلفن همراه بدست می‌آید، می‌توان به داده‌های وسیعی از جامعه هدف دست پیدا کرد و این امر فرصتی را برای سازمان‌های آماری فراهم کرده که هزینه‌های ثابت جمع آوری داده را کاهش دهند. نکته کلیدی، فراهم آوردن انگیزه برای مشارکت در مراحل جمع سپاری است به عنوان مثال با ترکیب آن با جامعه هدفی که همیشه تمایل دارند آنلاین باشند (به عنوان مثال در ارائه خدمات دولتی). چالش مهم حفظ حریم شخصی و انتخاب درست نمونه‌ها به

^{۱۱} Regulation

^{۱۲} Open data

^{۱۳} Data partnership

^{۱۴} Crowdsourcing

- عنوان نماینده جامعه است. یک رویکرد سیستمی درست در فرایند جمع سپاری، می‌تواند راه حلی قدرتمند برای دسترسی به داده باشد و داده‌های شهروندان را به داده‌هایی که معرف کل جامعه است، تبدیل کند.
- ۱۸- مواردی که در بالا به عنوان مدل‌های کسب و کار ارائه شده، کامل نیست. ادارات ملی آمار برای دستیابی به نتایج و تحلیل‌های مورد نیاز باید فعالیت‌های تحقیقاتی انجام دهند تا بتوانند به هماهنگی‌های لازم در سطوح ملی و بین‌المللی دست یابند. این تلاش‌ها شامل موارد زیر می‌شود:
- ۱- توجه و تمرکز بر "سیاست داده"^{۱۵} به عنوان استراتژی ملی داده باز
 - ۲- ایجاد استانداردها در حوزه پروتکل‌های توزیع شده و جمع‌آوری اطلاعات از طریق سرمایه‌گذاری مشترک در برنامه‌های ملی و بین‌المللی
 - ۳- برقراری مکانیزم مذاکره با برخی از گردانندگان اصلی (بزرگ) عرصه دیجیتال و ارائه‌دهندگان داده تجاری، به منظور ایجاد رویکرد تامین داده مشترک و شراکت جهانی

پلتفرم نسل بعدی: آمارهای فعال با هوش مصنوعی

۱۹- دهه‌ها است که تسلط بر پلتفرم‌ها و تکنولوژی از دغدغه‌های اصلی ادارات ملی آمار است. چارچوب تضمین کیفیت در فناوری اطلاعات خودنمایی می‌کند. ایجاد فضای ذخیره‌سازی مناسب و انجام محاسبات دقیق برای آمارشناسان و متخصصان داده اهمیت بسیار دارد. تلاش برای ایجاد استانداردهایی مانند SDMX^{۱۶} در نهایت تبدیل به نرم افزارهای داده باز می‌شود. این امر در مورد مه داده بیشتر مصداق دارد. تغییر رویکرد از آمار به ترکیب آمار و داده، به تغییر در معماری پلتفرم‌های چرخه داده منجر خواهد شد.

۲۰- اصولی که در طراحی پلتفرم‌های نسل بعدی راهبردی هستند عبارتند از:

- آمارها به سمت آمارهای قابل اجرا توسط هوش مصنوعی تکامل پیدا می‌کنند: ادارات آمار، سرمایه‌گذاری قابل توجهی در استانداردهای مشترک (DDI^{۱۸}, SDMX, GSBPM^{۱۷}) برای بهینه‌سازی مدل‌ها و فرآیندهای داده، تبادل داده و فراداده و مکانیزم‌های گزارش‌دهی از طریق رابط‌های برنامه‌نویسی باز کرده‌اند. این تلاش‌ها باید به صورت مستمر تا بهره‌مندی کامل از نتایج و دستیابی به سیستم‌های آماری ملی یکپارچه و مکانیزم‌های گزارش‌دهی بین‌المللی ادامه پیدا کند. تلاش‌های بیشتری نیز در توسعه کیفیت داده از طریق طراحی در سیستم‌ها (VTL) پیش‌بینی می‌شود. آمارها (و شاید، آمارهای قابل اجرا توسط بلاکچین) می‌توانند این مسیر را پیگیری کنند. در صورتی که این امر واقعیت پیدا کند، آمارهای تولید شده توسط هوش مصنوعی می‌توانند رقیبی برای ادارات آمار باشند تا به بهبود بهره‌وری مورد نیاز برای تبدیل شدن به سازمان‌های علم داده برسند.
- علم داده آمادگی پذیرش طیف گسترده‌ای از تکنیک‌ها با درجات مختلف تخصص را دارد: طیف انواع داده و تکنیک‌های توسعه بسیار گسترده است (تحلیل داده‌های میکرو، تحلیل فضایی، مدل‌سازی اجتماعی-اقتصادی، استخراج متن، تحلیل داده‌های تصویری، شبیه‌سازی سیاست...). پروژه‌های منبع باز (Jupyterhub, Spark, Python, R, ArcGIS...) ابزارهای علم داده رایگان، فرصت بزرگی برای سازمان‌های آماری فراهم می‌کنند.

۲۱- نسل بعدی پلتفرم‌های با قابلیت هوش مصنوعی بسیار پیچیده خواهند بود. این نیازمند سرمایه‌گذاری‌های قابل توجه و پایدار در ایجاد زیرساخت‌های فناوری اطلاعات و به‌روزرسانی مهارت‌های کارکنان و مقابله با مسائل قانونی و اخلاقی است. دستیابی به

^{۱۵} Data for policy

^{۱۶} Statistical data and metadata exchange

^{۱۷} Generic Statistical Business Process Model

^{۱۸} Data Documentation Initiative

این چالش‌ها، برای کارکنان ادارات آمار دشوار خواهد بود و مدل‌های کسب و کار جدید باید برای بهره‌برداری در اکوسیستم‌های ملی و بین‌المللی داده ابداع شوند.

۲-۳ ساختارهای اکوسیستمی: فرمت‌های جدید همکاری

۲۲- با توجه به حجم بالای تغییرات پیش رو، برای پیشرفت در پیدا کردن منابع، پلتفرم‌ها و تفاوت‌های مهارتی، باید اکوسیستم‌ها در سطوح مختلف ساخته، تقویت شوند و ارتباط برقرار شود.

- داخل سازمان: ادارات ملی آمار باید ساختارهای خود را بازنگری کنند. یکی از راه‌حل‌ها می‌تواند توسعه دورن سازمانی باشد، به طوری که دانش و روش‌ها در ساختارهای ایجاد شده تبادل شوند. این ادارات باید توسط متخصصین بیرون سازمان مانند دانشگاه، شرکت‌های تخصصی و ... غنی شوند. مقیاس سرمایه‌گذاری در منابع و پلتفرم‌های جدید داده نیازمند درجه بیشتری از تخصص است. همانطور که قبلاً گفته شد، مدل منبع باز در این حوزه حاکم است و امکان اتصال به نرم افزار و الگوریتم‌های منبع باز قدرتمندی وجود دارد که اهمیت بسیاری دارند.
- **در حوزه نظام ملی آمار:** ادارات ملی آمار، در موقعیتی طبیعی برای هدایت یک استراتژی داده هوشمند ملی قرار دارند. کمبود مهارت‌ها، ضرورت توسعه ارائه دهندگان داده و سرمایه‌گذاری‌های مورد نیاز در پلتفرم‌ها و مفهوم نقش قوی برای ادارات ملی آمار را نمایان می‌کنند.
- **در حوزه بین‌المللی:** مجامع بین‌المللی می‌توانند فضایی را ارائه دهند تا: (الف) به لطف به اشتراک‌گذاری تجربه و شیوه‌های خوب، آگاهی را درباره اهمیت توسعه استراتژی‌های داده هوشمند افزایش دهند؛ (ب) چارچوب‌ها و استانداردهای مشترک را توسعه دهند، به عنوان مثال در حوزه پروتکل‌های توزیع شده یا جمع‌سپاری؛ (ج) هماهنگی در گفتگو با بازیکنان دیجیتال بزرگ و ارائه‌دهندگان داده را انجام دهند؛ (د) هزینه‌ها را با سرمایه‌گذاری مشترک کاهش دهند.

۲۳- توسعه اکوسیستم‌ها - درون و خارج از هر سازمان آماری - یک جزء حیاتی از یک استراتژی داده هوشمند است، که برای غلبه بر نواقص منابع، مهارت‌ها و پلتفرم‌ها و به طور بنیادی‌تر، برای تقسیم هزینه‌ها که هیچ سازمانی نمی‌تواند به تنهایی تحمل کند کاربرد دارد.

۲۴- به طور خلاصه، دنیای آمار و داده به سرعت در حال تغییر است. تقاضاهای جدیدی برای به چالش انداختن برخی کارکردهای ثابت ادارات ملی آمار به وجود آمده است، اما همچنین فرصت‌های زیادی نیز ارائه شده است. ادارات ملی آمار، شروع به فعالیت در پذیرش تغییرات کرده‌اند و به بررسی فرصت‌های منابع داده جدید، شراکت‌های جدید و مدل‌های کسب و کار جدید پرداخته‌اند. جنبه‌های تغییر چندگانه و پیچیده هستند و برای مقابله با آنها به پیاده‌سازی استراتژی‌ها و چارچوب‌هایی نیاز است. به دلیل شباهت مسائل در سراسر کشورها، همکاری‌های بین‌المللی نیز در این حوزه مفید خواهد بود.

منابع:

Which Strategies for NSOs in the Digital Era? Towards 'Smart Data' Strategies ۱۵th meeting of the Committee on Statistics and Statistical Policy (CSSP) ۲۰-۲۱ June ۲۰۱۸ Palais des Nations, Geneva, Organization for Economic Co-operation and Development