

# دو یار دیرینه جدا نشدنی!

مهندسید دولت مدلی - معاونت شبکه و خدمات رایانه ای

پاییز ۱۴۰۰

مراکز آماری و فناوری اطلاعات و ارتباطات از دیرباز در کنار همدیگر بوده‌اند. از زمانی که برای محاسبات از ماشین های بزرگ (mainframe) استفاده می شد تا کنون که منابع داده‌ای جدیدی را در اختیار مراکز آماری قرار داده است. اکنون نیز مراکز آمار در شرایطی قرار گرفته‌اند که می‌خواهند با کمترین هزینه بیشترین خروجی را داشته باشند. از سوی دیگر توقع می رود که بتوانند داده‌های جمع‌آوری شده را در کمترین زمان ممکن به مخاطبین خود ارائه نمایند. اما مواردیکه باعث شده فناوری اطلاعات و ارتباطات در فرایند مراکز آماری تاثیرگذار باشند را می توان به صورت زیر نام برد:

۱. همکاری بین دستگاه‌ها: با توجه به اینکه نحوه کارکردن دستگاه‌های آماری متفاوت شده و از داده‌های سایر تولیدکنندگان استفاده کنند باید ارتباطی بین دستگاه‌ها شکل بگیرد. این ارتباط شامل: برقراری یک ارتباط مخابراتی است که امکان تبادل اطلاعات را خواهد داشت. پس از آن که این ارتباط در نهایت دقت و سرعت انجام شود، نحوه یکپارچه سازی داده و اجرای آن باید به نحو مطلوبی و با الگوریتم‌هایی در فناوری اطلاعات و ارتباطات اجرایی خواهد شد.

۲. استفاده مجدد از نرم افزارها: مراکز آماری می توانند از نرم افزارهای تولید شده در سایر مراکز آماری دنیا استفاده کرده و آگاهانه آنها را به کارگیرند تا از این طریق در زمان و هزینه صرفه جویی شده و علاوه بر آن باعث افزایش در کیفیت امکانات نرم افزارها شوند.

۳. افزایش استفاده از اینترنت \_ وب: استفاده از اینترنت برای دانلود فایل های PDF گرفته تا ارتباطهای ماشین \_ ماشین باعث شده است که ارتباطات مراکز آماری تحت تاثیر قرار گرفته و بتوانند داده ها را با یکدیگر به اشتراک بگذارند. در گام بعدی داده هایی که بر روی این بستر منتشر شده منبع داده مهمی برای تولید آمار شده اند.

۴. استفاده از اینترنت اشیا (IOT): از سال ۲۰۰۸ اتصال اشیا به اینترنت آغاز شده است. اتصال دستگاه‌های خانگی مانند: یخچال، تلویزیون، لباس، ساعت و .... باعث شده که منابع داده‌ای جدیدی برای مراکز آماری به وجود بیایند. تخمین زده شده است که در حال حاضر حدود ۵۰ بیلیون دستگاه به اینترنت متصل شده و در حال جمع آوری داده‌ها هستند.

۵. استفاده از تکنولوژی ابری (Clouding): یکی از تکنولوژی هایی که از سال ۱۹۹۶ بر پایه اینترنت شکل گرفته است. این تکنولوژی باعث می شود خدماتی شامل: سرورها، ذخیره سازی‌ها و نرم افزارهای کاربردی،

از طریق شبکه و اینترنت به اشتراک گذاشته شده و در دسترس بقیه قرار گیرد. این کار موجب می‌شود که در استفاده از فضای مورد نیاز و زیرساخت‌های مورد نیاز صرفه‌جویی شود.

۶. تلفن‌های هوشمند و تبلت: این دستگاه‌ها به دلیل قابلیت‌های محاسباتی و ارتباطی در حال حاضر حدود ۶ میلیارد تلفن هوشمند در حال استفاده از این دستگاه‌ها هستند. این دستگاه‌ها می‌توانند علاوه بر این که در جمع‌آوری داده‌ها کمک کنند، از سوی دیگر برای اجرای طرح‌های آماری‌گیری تعاملی نیز مورد استفاده قرار گیرند. استفاده از این دستگاه‌ها موجبات بالارفتن دقت و سرعت انتقال داده‌ها می‌شود.

۷. بصری‌سازی داده‌ها (visualization): در قبل بصری‌سازی به معنی این بود که بتوان در یک محیط ساده انواع نمودارها را پیاده‌سازی کنند. اما اکنون نیاز است که بصری‌سازی به صورت تخصصی‌تر پیاده‌سازی شود. بصری‌سازی می‌تواند به صورت: متنوع، برخط و بر حسب دسترسی‌های مورد نیاز پیاده‌سازی می‌شود. از سوی دیگر می‌توان با قرار دادن دیتاست‌ها (Dataset) نمودارهای متفاوتی به دست آورد. از سوی دیگر ساخت گزارش‌های چند بعدی نیز در حال افزایش است که این موضوع نیز با ابزار موجود در فناوری اطلاعات و ارتباطات ممکن شده است. برای جذاب‌سازی و همچنین رواج آمار در جامعه نیز محصولات انیمیشنی نیز به کمک آمده است.

۸. استفاده از داده‌های حجیم (Big Data): داده‌های حجیم منابع داده‌ای جدیدی هستند که دارای ویژگی‌های حجم بالا، سرعت تغییرات و فرمت‌های متفاوت است. این داده‌ها معمولاً به صورت غیر ساختاریافته هستند. برای این که بتوان از داده‌هایی با این ویژگی استفاده کرد، فناوری اطلاعات و ارتباطات این نوع داده‌ها را جمع‌آوری و با قابلیت انتقال داده و ساختاریافته کرده و از آنها گزارش‌های مورد نظر را تهیه می‌نماید.

۹. استفاده از داده‌های باز (Open Data): داده‌های باز که از سال ۲۰۰۷ شکل گرفته است داده‌هایی است که باید به گونه‌ای ارائه شود تا برای کاربران مجدداً قابل استفاده شود. این داده‌ها باید با فرمت‌های قابل استفاده (نه عکس و PDF) مانند: ساختار متنی (CSV) یا زبان علامتگذاری شده (Markup) ارائه شوند.

۱۰. استفاده از نرم‌افزارهای منبع‌باز (Open Source): نرم‌افزارهایی که منبع‌باز هستند می‌توانند برای کاهش هزینه‌ها در مراکز آماری مورد استفاده قرار گیرد. از نمونه این نرم‌افزارها می‌توان به لینوکس، R، MySQL و ..... نام برد.

۱۱. اطلاع‌رسانی: جداول ساده با استفاده از فلش مموری و لوح فشرده به کاربران آماری قابل ارائه است. اکنون با استفاده از وب‌سایت‌ها، اطلاع‌رسانی‌های مبتنی بر اطلاعات مکان جغرافیایی نیز قابل ارائه شده است. با استفاده از واسط برنامه‌نویسی نرم‌افزارهای کاربردی (API) باعث شده که انتقال اطلاعات بین

نرم افزارهای مختلف نیز انجام شود. از سوی دیگر ظهور شبکه‌های اجتماعی ابزاری برای مراکز آماری شده که باعث تسهیل در فرهنگ سازی آماری می‌شود.

۱۲. علم داده‌ها (Data Science): در سال‌های اخیر که داده‌های بزرگ قابل استفاده شده‌اند، علم داده‌ها نیز شکل گرفته است. علم داده‌ها این هدف را دارد که با استفاده از مباحث آماری، ریاضی و فناوری اطلاعات و ارتباطات، اطلاعات را استخراج کرده و بتواند پیش‌بینی‌هایی برای آینده انجام دهد. برای اینکه بتواند این مهم به سرانجام برساند از روش‌های هوش مصنوعی (AI) استفاده می‌کند. هوش مصنوعی به معنی این است که سیستم بتواند با استفاده از برنامه نویسی‌ها و الگوریتم‌ها همانند انسان تصمیم‌گیری کند.

۱۳. امنیت داده: محرمانگی و حریم شخصی از قوانین اولیه و پایه‌ای مربوط به مراکز آماری است. نگرانی داشتن امنیت داده‌ها از مواردی است که باعث می‌شود تا اعتبار هر دستگاه آماری دچار خدشه شود. از سوی دیگر درخواست‌های کاربران آماری دسترسی به ریز داده‌ها است. همیشه بیم این می‌رود که اطلاعات شخصی افراد در دسترس افراد دیگر قرار گیرد. فرایندها، تجهیزات سخت‌افزاری و نرم‌افزارهای مربوطه به کمک آمده است تا بتواند دسترسی به این داده‌ها را مدیریت کند و تا جایی که می‌تواند ریسک دسترسی نادرست به داده‌ها را مدیریت کند.

حال باید منتظر ماند تا با پیشرفت بیشتر فناوری اطلاعات و ارتباطات چه تأثیرات دیگری را بر روی مراکز آماری خواهیم دید.

منابع:

1. Handbook on Management and Organization of National Statistical Systems -Chapter 14 - Information Technology Management
2. <https://www.bankmycell.com/blog/how-many-phones-are-in-the-world>
3. <https://www.zdnet.com/article/what-is-cloud-computing-everything-you-need-to-know-about-the-cloud/>
4. <https://opendatahandbook.org/guide/en/what-is-open-data/>
5. [https://www.sas.com/en\\_us/insights/big-data/what-is-big-data.html](https://www.sas.com/en_us/insights/big-data/what-is-big-data.html)