



صدا و سیما جمهوری اسلامی ایران

معاونت سیاسی

اداره پژوهش‌های سیاسی

## روزنامه‌نگاری داده محور در خبر تلویزیون

فرآورده‌های خبری و تولیدات پژوهشی در بخش‌های زیر قابل دسترس است:

– وب سایت خبرگزاری صداوسیما (سرویس پژوهش) <http://www.iribnews.ir>

پژوهشگر: زهرا فرخی

۱. نکات مهم .....
۲. مقدمه .....
۲. روزنامه نگاری داده محور چیست؟ .....
۴. انواع روزنامه نگاری داده محور .....
۵. فرآیندهای روزنامه نگاری داده محور .....
۵. مهارت‌های لازم برای روزنامه نگاری داده محور .....
۵. موضوعات مناسب و منابع دریافت داده برای روزنامه نگاری داده محور .....
۷. سوژه یابی در روزنامه نگاری داده محور .....
۸. داده‌های کیفی برای روزنامه نگاری داده محورها کدام است؟ .....
۹. ابزارها و نرم‌افزارهای روزنامه نگاری داده محور .....
۱۰. نتیجه گیری .....
۱۰. منابع .....

□ نکات مهم

□ داده‌نگاری نوعی از روایت است که با در نظر گرفتن واقعیت‌های امروز دنیای رسانه یعنی «کوتاه شدن بازه توجه مخاطب» و «عرضه حجم بالای اطلاعات»، روشی برای «انتقال حجم بالای اطلاعات در زمانی کوتاه» است.

□ «روزنامه‌نگاری داده محور» نوعی روزنامه‌نگاری است که برای انجام آن فرآیندهایی همچون یافتن و گردآوری، پاکسازی و ساماندهی داده‌ها، بیان خبر به یاری یک روایت، تحلیل، تصویرسازی داده‌ها و انتشار داده‌ها طی می‌شود.

□ هدف «روزنامه‌نگاری داده محور» صرفاً ارائه محتوای گرافیکی با جذابیت‌های بصری بالا نیست، بلکه گزارش کردن و روایت کردن رویدادها به بهترین شیوه‌های ممکن است.

□ پرکاربردترین گونه‌بندی «روزنامه‌نگاری داده محور» عبارت است از پژوهشی و عمومی

□ فرآیند «روزنامه‌نگاری داده محور» شامل مراحل گردآوری، پاکسازی، ایجاد بافت، ترکیب و انتقال اطلاعات است.

□ موضوعات سیاسی (انتخابات)، آموزشی، اقتصادی، ورزشی، محیط زیست و حتی امور اجتماعی و بهداشت و سلامت مناسب «روزنامه‌نگاری داده محور» است.

□ منابع دریافت داده «روزنامه‌نگاری داده محور» عبارتند از: آژانس‌های دولتی، وب سایت‌ها، منابع مرتبط با دولت، منابع غیر رسمی و سازمان‌های غیرانتفاعی.

روزنامه‌نگاری متأثر از عوامل مختلف، اعم از تحولات تکنولوژیکی، گسترش اطلاعات، نیاز مخاطب، اقتصاد نشر، تحولات نسلی و ... است و به طور مداوم تغییر می‌کند. این تغییرات به شکل‌گیری انواع روزنامه‌نگاری منجر می‌شود. با ظهور فناوری‌های جدید به ویژه شبکه‌های اجتماعی روند گسترش اطلاعات شتاب فراوانی گرفته و مخاطبان و کاربران رسانه علاوه بر مصرف کننده، خود به تولیدکننده تبدیل شده‌اند و حجم وسیعی از اخبار و اطلاعات را به وجود آورده‌اند که نیاز به گردآوری، تحلیل و تصویرسازی متناسب با اهداف رسانه‌ای مدنظرشان دارند. از طرف دیگر هر روزه طیف وسیعی از اطلاعات در اختیار فعالان و خبرنگاران رسانه‌ها قرار می‌گیرد، این اطلاعات خام است و باید پردازش شود تا فعالان رسانه‌ای بتوانند براساس ارزش‌ها و سیاست‌های خبری خود از آنها استفاده کنند.

افزایش داده‌ها، به خصوص داده‌های بزرگ، بهره‌گیری از آنها را ضروری کرده و ورود داده‌ها به دنیای روزنامه‌نگاری، منجر به شکل‌گیری نوع جدیدی از روزنامه‌نگاری به نام «روزنامه‌نگاری داده محور»<sup>۱</sup> یا داده‌نگاری را فراهم کرده است. داده‌نگاری نوعی از روایت است که با در نظر گرفتن واقعیت‌های امروز دنیای رسانه یعنی «کوتاه شدن بازه توجه مخاطب» و «عرضه حجم بالای اطلاعات»، روشی برای «انتقال حجم بالای اطلاعات در زمانی کوتاه» پیدا کرده است. بنابراین به نسبت افزایش داده‌ها نیاز به کسانی که بتوانند این اطلاعات را به زبانی همه فهم تبدیل کنند، بیشتر می‌شود.

با ورود به عصر اطلاعات تجزیه و تحلیل داده‌ها از نیازهای اساسی رسانه است تا بتواند از اطلاعات جهت جریان‌سازی و گفتمان‌سازی استفاده کند. همچنین تلویزیون امکانات منحصر به فردی برای تصویرسازی و ارائه تصویر به مخاطبان دارد، که می‌تواند برای انتقال اطلاعات به مخاطبان از طریق این نوع خبررسانی موفق عمل کند.

## □ روزنامه‌نگاری داده محور چیست؟

تعریف روزنامه‌نگاری داده محور یا داده‌نگاری مشکل است چرا که تعریف داده مشکل است علاوه بر آن با حضور تکنولوژی‌های جدید تعریف سنتی روزنامه‌نگاری هم تغییر کرده است. هنوز هم بسیاری بر این باورند که روزنامه‌نگار باید تایپ کند یا بنویسد که طبیعتاً روش‌های جدید روزنامه‌نگاری امروز را نادیده می‌گیرد. ارائه حجم زیادی از داده در مدت زمان کم را می‌توانیم حاصل کار یک داده‌نگار بدانیم. به نظر آن چیزی که دنیای امروز را دنیای روزنامه‌نگاری داده کرده سه عامل اساسی زیر است: در دسترس بودن حجم بسیار بالای اطلاعات، تکنولوژی‌هایی که مصور سازی را ممکن کرده و از همه مهمتر کمتر شدن بازه زمانی تمرکز و توجه مخاطب که به طور کلی جغرافیای رسانه‌ای امروز را زیر و رو کرده است.

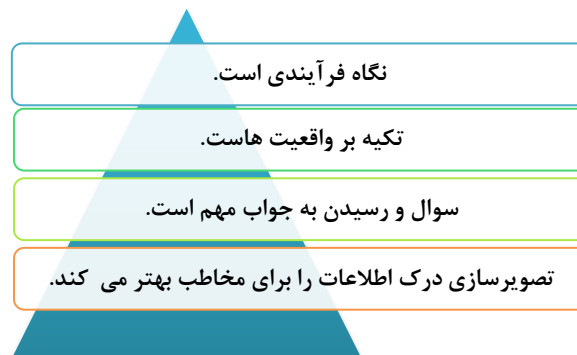
ریموند جوزف در مطلبی با عنوان «روزنامه‌نگاری داده محور»: مجموعه ابزار جدید روزنامه‌نگاری داده محور را این‌گونه تعریف می‌کند: روزنامه‌نگاری داده محور به مجموعه‌ای از ابزار اشاره دارد که به روزنامه‌نگار در کار با حجم زیاد اطلاعات، بررسی عمیق‌تر آنها و ارتباطات بهتر کمک می‌کند. روزنامه‌نگاری داده محور در واقع

<sup>۱</sup>. Data journalism

مفهومی قدیمی با نامی جدید است. گزارشگری به کمک رایانه، که زمانی همان روزنامه‌نگاری داده محور بود شامل استفاده از رایانه برای تحلیل داده‌های زیاد از طریق جست‌وجو و مرتب‌سازی اطلاعات درون پایگاه‌های داده است. (سلیمی، ۱۳۹۶/۶/۲۵، بازیابی از سایت همشهری آنلاین)

در کتاب «راهنمای کاربردی روزنامه‌نگاری داده محور»، روزنامه‌نگاری داده محور این‌طور تعریف شده است: روزنامه‌نگاری داده محور نوعی روزنامه‌نگاری است که برای انجام آن فرآیندهایی همچون یافتن و گردآوری، پاکسازی و ساماندهی داده‌ها، بیان خبر به یاری یک روایت، تحلیل، تصویرسازی داده‌ها و انتشار داده‌ها طی می‌شود. چند نکته در این تعریف هست:

#### محورهای مهم روزنامه‌نگاری داده محور



روزنامه‌نگاری به کمک داده‌ها را روزنامه‌نگاری داده محور می‌نامند. بعضی معتقدند داده مجموعه‌ای از واقعیت‌ها، آمار و ارقامی است که به صورت خام و گسترده جمع‌آوری می‌شود تا به عنوان مرجع مورد استفاده قرار گیرد. پروفیسور نایجل شد بولت<sup>۱</sup> معتقد است: روزنامه‌نگاری داده محور استفاده از داده‌ها و اطلاعات اساسی و عناصری از منابع کلیدی برای شکل دادن به یک خبر است. پاول لوپس سردبیر پروژه‌های خاص گاردین در تعریف دیتا ژورنالیسم می‌گوید: روزنامه‌نگاری داده محور تنها به معنای وجود داده‌ها یا به دست آوردن و انتشار آن نیست بلکه فرآیندی است که برای تولید آن انجام می‌شود و باید سؤالات مناسبی را بپرسید تا پاسخ‌های درست را دریافت کنید.

بیشتر تعریف‌ها نگاه فرآیندی به روزنامه‌نگاری داده محور دارند و اینکه چه فرآیندی را باید طی کرد. در روزنامه‌نگاری داده محور تکیه اصلی بر واقعیت‌هاست. در کل در این شیوه روزنامه‌نگاری؛ بیان خبر به یاری یک روایت، تحلیل، تصویرسازی داده‌ها و انتشار داده‌ها طی می‌شود. چرا که تصویرسازی درک اطلاعات را برای مخاطب بهتر می‌کند. (سلیمی، ۱۳۹۶/۸/۱۳، بازیابی از سایت همشهری آنلاین)

سایمون راجرز (۲۰۱۱) توضیح می‌دهد که هدف این نوع روزنامه‌نگاری صرفاً ارائه محتوای گرافیکی با جذابیت‌های بصری بالا نیست، بلکه گزارش کردن و روایت کردن داستان‌ها به بهترین شیوه‌های ممکن هدف اصلی است.

<sup>۱</sup> - از موسسان سایت data.gov.uk و موسسه داده‌های آزاد

(«گونه شناسی روزنامه‌نگاری جدید»، ص ۶) نتیجه کار یک روزنامه‌نگاری داده محور می‌تواند نموداری ساده، نقشه‌ی تعاملی، چند نمودار و یا چند نقشه، تایم‌لاین یا راهنما، اینفوگرافیک یا موشن‌گرافی باشد که به یمن نرم‌افزارهای جدید این شاخه از روزنامه‌نگاری، به آسانی در دسترس روزنامه‌نگاران قرار گرفته است. روزنامه‌نگاری داده محور، مثل فوتوژورنالیسم یا ویدئوژورنالیسم نیست که با یک نوع خاصی از رسانه در ارتباط باشد. داده‌نگاری با مجموعه‌ای از رسانه‌های آنلاین، دیداری و چاپی و محتواهای مختلف در ارتباط است و وابستگی کمی به متن دارد اما به فهم درست از عدد و رقم نیاز دارد. («مجموعه آموزشی روزنامه‌نگاری داده (دیتاژورنالیسم)»، ۱۳۹۵/۱۱/۱، بازیابی از سایت دیتاویز)

## □ انواع روزنامه‌نگاری داده محور

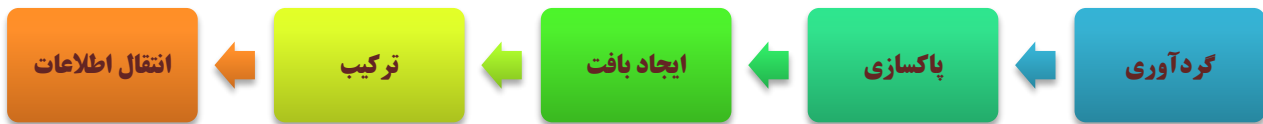
دسته‌بندی‌های متفاوتی متناسب با کشورها و صاحب نظران مختلف از روزنامه‌نگاری داده محور وجود دارد که پرکاربردترین آن گونه‌بندی روزنامه‌نگاری داده محور به پژوهشی و عمومی است.



روزنامه‌نگاری داده محور **پژوهشی**: نگران زمان نیست، در اینجا زمان زیادی برای کار لازم است؛ حتی ممکن است چند سال طول بکشد، مهارت‌های داده‌ای مثل کدنویسی باید در حد پیشرفته باشد، به اطلاعات غیر رسمی، محرمانه و داده‌های افشا شده توجه زیادی می‌شود، کنترل مجدد داده‌ها و اکتفا نکردن به یک منبع در مورد صحت داده‌ها مورد تاکید قرار می‌گیرد.

روزنامه‌نگاری داده محور **عمومی**: برخلاف روزنامه‌نگاری داده محور پژوهشی در این نوع، زمان اهمیت دارد و ممکن است در چند ساعت یا چند روز کار آماده شود، با مهارت‌های داده پایه‌ای مثل excel هم می‌توان کار را پیش برد. برای جمع‌آوری اطلاعات نگاه به مجموعه‌های بزرگ عمومی با امکان دسترسی آسان به داده‌هاست. در رابطه با منابع و کنترل مجدد داده‌ها نسبت به روزنامه‌نگاری داده محور پژوهشی سخت‌گیری کمتری وجود دارد. **نکته:** علاوه بر این دو مورد روزنامه‌نگاری داده محور به نام «روزنامه‌نگاری داده محور همزمان» شکل گرفته که در حال پررنگ شدن است. در این نوع؛ الگوریتم‌ها و ربات‌ها به کمک تیم‌های روزنامه‌نگاری داده محور می‌آیند و به تسریع فرآیندها کمک می‌کنند. («دیتاژورنالیسم، مقدمات و آموزش، ابزار جمع‌آوری و تحلیل داده»، صص ۵ و ۶)

## □ فرآیندهای روزنامه‌نگاری داده محور



گردآوری: می‌تواند به دو صورت باشد؛ داده‌ها را داریم اما پرس وجو می‌کنیم تا داده‌ها را کامل کنیم و یا اینکه داده و پرسشی وجود ندارد و تلاش می‌کنیم داده جمع‌آوری شود.

پاکسازی: حذف خطاهای انسانی و تبدیل داده‌ها به ساختاری که باهم هماهنگ هستند.

ایجاد بافت: رفتن در قالب و بستری مناسب؛ مثلاً آمار جرائم را داریم ولی به تنهایی معنا ندارد و باید در کنار تعداد جمعیت بیاید تا داده به صورت خاص معنی پیدا کند.

ترکیب: استفاده از دو پایگاه داده است.

برقراری ارتباط یا انتقال اطلاعات: مرحله آخر که معمولاً همراه با تصویرسازی داده است. از مرحله جمع‌آوری داده تا نهایه شدن کار به ترتیب بر ارزش اطلاعات افزوده می‌شود. («دیتاژورنالیسم، مقدمات و آموزش، ابزار جمع‌آوری و تحلیل داده»، ص ۶)

## □ مهارت‌های لازم برای روزنامه‌نگاری داده محور

گزارشگری مبتنی بر رایانه: گزارشگر باید ذهنی پرسشگر داشته باشد، دنبال داده‌ها برود، برای دستیابی به آنها تلاش کند و حداقلی از آمار و ریاضیات بداند.

توسعه دهنده‌های اپلیکیشن‌های خبری: وظیفه عمومی کردن داده‌ها را دارند.

متخصص‌های تصویرسازی داده: وظیفه بصری کردن اطلاعات را برعهده دارند. در حال حاضر تقاضا برای این تخصص‌ها زیاد است ولی تعداد افرادی که این تخصص‌ها را در دنیا داشته باشند کم است. پیشنهاد می‌شود رسانه ملی در بخش‌های مختلف تلاش کند چنین نیروهایی را تربیت کند؛ مثلاً برای توسعه اپلیکیشن خبری از متخصصان حوزه آی‌تی کمک بگیرد.

نکته: اهمیت رنگ‌ها در روزنامه‌نگاری داده محور: متخصصان این حوزه اصرار دارند هر رنگی برای روزنامه‌نگاری داده محور مناسب نیست و از اصطلاحی به نام رنگ‌های رنگین‌کمانی استفاده می‌کنند که برای نرم افزارها تعریف شده است. همچنین می‌گویند از رنگ‌هایی که تاثیر کمتری دارد و مخاطب به آنها عادت کرده است نباید استفاده شود. متخصصان پیشنهاد می‌کنند پالت‌های رنگی جدید ساخته شود. مثلاً گاردین برای خودش جدول رنگی تعریف کرده است و تیم روزنامه‌نگاری داده محور و افراد فعال در قسمت بصری سازی از این پالت رنگی استفاده می‌کند. («دیتاژورنالیسم، مقدمات و آموزش، ابزار جمع‌آوری و تحلیل داده»، صص ۶ و ۷)

## □ موضوعات مناسب و منابع دریافت داده برای روزنامه‌نگاری داده محور

تصور عموم این است که هر آنچه مبتنی بر آمار و داده باشد می‌تواند مناسب باشد؛ اما در پژوهشی مشخص شده که آنچه بیشتر مورد توجه است عبارتند از موضوعات: سیاسی (انتخابات)، آموزش (گزارش‌هایی درباره میانگین نمره امتحانات دانش‌آموزان در نقاط مختلف کشور)، اقتصاد (وجود اعداد زیاد و بازار بورس) و همچنین ورزش

(بخش زیادی از آن درباره بهترین و بدترین امتیازات و جایگاه تیم‌ها در جدول لیگ است)، محیط زیست و حتی امور اجتماعی (داده‌های دولت درباره مسائل خاص) و بهداشت و سلامت. موضوعات زیادی وجود دارند و باید به خاطر داشته باشیم که هنگام استفاده از دیتاژورنال‌سیسم، مهارت‌هایی را از دنیای علوم اجتماعی و روش‌های آماری قرض می‌گیریم و این برای گستره عظیمی از موضوعات صدق می‌کند. بنابراین، می‌توان از روزنامه‌نگاری داده محور به شکل‌های مختلف استفاده کرد.

منابع دریافت داده روزنامه‌نگاری داده محور هم عبارتند از: آژانس‌های دولتی، وب سایت‌ها، منابع مرتبط با دولت، منابع غیر رسمی و سازمان‌های غیرانتفاعی. روزنامه‌نگاری داده محور بدون داده غیرممکن و مثل نقاشی بدون رنگ است. اما داده (دیتا) را از کجا بیاوریم؟ پیش از هر چیز باید به چند نکته ضروری توجه کنیم. اول اینکه بخش عمده کار یک داده‌نگار به لحاظ زمانی جمع‌آوری، مرتب کردن و تمیز کردن داده‌هاست و بخش جالب داستان که مصورسازی آن داده‌هاست بخشی است که به کمک نرم‌افزارهای امروز، زمان زیادی نمی‌برد. هرچند که پیدا کردن بهترین روش برای ارائه اطلاعات می‌تواند زمان‌بر باشد.

نکته مهم این است که در بسیاری از موارد اساساً اطلاعاتی وجود ندارد که از مراکز دولتی کسب کرد حتی در کشورهای صنعتی که در حوزه آمار پیشرو هستند در خیلی از موارد بسیاری از نهادها مثل پلیس یا سایر نهادهای دولتی در جمع‌آوری اطلاعات دچار مشکل هستند. ممکن است این اطلاعات جمع نشده یا اطلاعات مفید یا دقیق و به روز نباشد، یا داده‌ها به نوعی گردآوری شده‌اند که نمی‌شود از آنها استفاده کرد چون تصفیه آنها ممکن نیست یا اینکه اطلاعاتی که از نهادهای مختلف جمع شده با هم هماهنگ نیستند و به همین دلیل قابل اتکا و استناد نیستند.

نکته آخر اینکه اساساً داشتن اطلاعات منظم و مرتب که بدون آماده‌سازی و تصفیه بشود از آنها استفاده کرد غیرممکن است. با در نظر گرفتن این نکات منابع را معرفی می‌کنیم:

از ایرادهای مهم سایت‌های منبع این است که معمولاً اطلاعات را به صورت pdf در اختیار کاربر قرار می‌دهند. فراموش نکنید به عنوان یک داده‌نگار شما همیشه از پی دی اف فرار می‌کنید همیشه به دنبال فرمت‌های CSV یا اکسل هستید؛ گاهی در گزارشی که در حال نوشتن آن هستید داده وجود دارد؛ مثلاً وقتی که از بحران کم آبی در استان‌ها صحبت می‌کنید می‌توانید از نقشه ایران برای نشان دادن استان‌های کم آب استفاده کنید. ابزارهایی که می‌تواند به شما کمک کند برای اینکه داده‌های خوبی را پیدا کنید عبارتند از: ابزار جستجوی پیشرفته گوگل که به شما امکان جستجو براساس فرمت را می‌دهد. همین‌طور موتور جستجوی بینگ<sup>۱</sup> متعلق به شرکت مایکروسافت و ... («مجموعه آموزشی روزنامه‌نگاری داده (دیتاژورنال‌سیسم)»، ۱۳۹۵/۱۱/۱، بازیابی از سایت دیتاویز) مزایای روزنامه‌نگاری داده محور عبارتند از: کمک به اینکه یک خبرنگار یا رسانه بتواند ادعای خودش را ثابت کند. امکان پرداختن به موضوعات بزرگ را فراهم می‌کند. می‌تواند مسائل اجتماعی مبهم را با استناد به داده‌ها روشن کند. اطلاعات ارزشمند را به گفتمان اجتماعی تبدیل کند. موانعی را که سر راه درک درست مخاطب وجود دارد کم کند. داده‌ها معمولاً زیبا نیستند و نقش روزنامه‌نگاری داده محور علاوه بر مزیت‌های دیگری که دارد؛ تبدیل داده‌های زشت به اطلاعات زیباست.

<sup>۱</sup>. Bing



چالش‌های روزنامه نگاری داده محور: بهره‌گیری نادرست و جهت‌دار از داده‌ها، دانش کم در استفاده از داده‌ها، نحوه دستیابی به داده‌ها (در کشورهای مختلف متفاوت است)، فرمت داده‌ها، هزینه‌ها که بستگی به سطح پروژه دارد و در آخر خطای انسانی از اهم چالش‌های روزنامه نگاری داده محور محسوب می‌شوند.

### آثار روزنامه نگاری داده محور عبارتند از:



(سلیمی، ۱۳۹۶/۸/۱۳، بازیابی از سایت همشهری آنلاین)

### □ سوژه‌یابی در روزنامه نگاری داده محور

از ویژگی‌های مهم دنیای امروز ما ارائه حجم بالای اطلاعات یا داده است. این داده‌ها طبیعتاً موضوعات مختلفی را شامل می‌شوند. کار داده‌نگار دیدن داستان‌ها و سوژه‌هایی است که در این داده‌ها نهفته است. همان‌طور که روزنامه‌نگار برای پیدا کردن یک موضوع خوب سیاسی یا اجتماعی باید شامه تیزی داشته باشد و محیط اطرافش را به خوبی رصد کند، داده‌نگار هم باید همین حس را وقتی با حجم بالایی از داده روبرو می‌شود؛ داشته باشد. البته هر داده‌نگاری بعد از مدتی کار با اطلاعات و داده‌ها یاد می‌گیرد که چطور و چگونه به دنبال سوژه‌ها باشد. اما به صورت ساده و کلی چند پیشنهاد ارائه می‌شود.

باید داده‌ها را مرتب کنید، سر و ته و اول و آخر داده‌ها را بررسی کنید. پیش از هر چیز چهار گوشه آن را چک کنید، با چه حجمی از داده روبرو هستید؟ داده شما از چه جنسی است؟ عدد، تاریخ، تگ جغرافیایی یا مثلاً کلمه است. آیا این داده‌ها با هم همخوان هستند؟ ستون‌ها یا ردیف‌های مختلف فایل‌های اکسل چه تناسبی با هم دارند؟ حواستان باشد که مصورسازی با روایت سوژه فرق می‌کند. مصورسازی بخشی از روایت شماست. تصویرسازی لزوماً داستان و سوژه‌ای را در خود ندارد این وظیفه شماست که اول سوژه را کشف کنید و بعد از تصویرسازی به عنوان ابزاری برای روایت آن استفاده کنید. برای دیدن داستان به سراغ بیشترین‌ها یا کمترین‌ها بروید. به سراغ ترندها یا روندها بروید. آیا افزایش یا کاهش در کار است؟ داده‌هایی که میانگین فاصله زیادی دارند کمک زیادی می‌کنند که سوژه‌تان را کشف کنید. مثلاً تصور کنید نرخ اعتیاد جوانان در همه استان‌ها بین ۳ تا ۵ درصد است. اگر در استانی این نرخ یک درصد یا ۱۵ درصد باشد، این داستان شماست. اگر داده‌ها در یک بازه زمانی عرضه شده‌اند، به دنبال نرخ افزایش یا کاهش باشید. در پی چیزی باشید که از روند اصلی بیرون باشد. چیزی که بتوانید آن را در فضایی که می‌شناسید معنی کنید. مثلاً نرخ دزدی در مناطق مختلف تهران را دارید. آیا همه این نرخ‌ها به هم نزدیکند؟ آیا تفاوت معنی‌داری بین سطح درآمد این مناطق و میزان دزدی وجود دارد؟ یا مثلاً هر مترمربع زمین

با تعداد دزدی ارتباط دارد؟ تنها ابزار برای این مقایسه‌ها عدد و رقم است. («مجموعه آموزشی روزنامه‌نگاری داده (دیتاژورنالیزم)»، ۱۳۹۵/۱۱/۱، بازیابی از سایت دیتاویز)

## □ داده‌های کیفی برای روزنامه‌نگاری داده محورها کدام است؟

داده‌هایی که ما و مخاطب‌مان را گمراه می‌کنند و باید از آن‌ها فرار کرد و با احتیاط با آن‌ها برخورد کرد داده‌های کیفی هستند. شاید به صورت کلی بتوانیم بگوییم که جمع‌شدن قابلیت استفاده از عدد و کلمه کار ساده‌ای نیست. داده‌نگارها هم باید روایت‌گرهای خوبی باشند هم از پس فهم عدد و رقم بر بیایند.



به عنوان مثال در این خبر واضح است که روزنامه‌نگار به عدد دقت نکرده است چون هیچ کدام از این اعداد به هم جور نیستند. جمع ۱۲۰ هزار و ۹۸۰ هزار، یک میلیون و ۳۰۰ هزار نمی‌شود و سه برابر ۱۲۰ هزار، ۹۸۰ هزار نیست. اما اگر ۱۲۰ هزار را به ۳۲۰ هزار تغییر بدهیم آن وقت اعداد معنی‌دار می‌شود. این اشتباهی است که نمی‌شود از یک داده‌نگار انتظار داشت، چون این اشتباه وقتی به صورت یک کاف و یا یک نمودار روی صفحه رسانه نمایان شود دیگر یک اشتباه ساده‌تایی نیست.

اما داده‌های بد داده‌های غیرقابل اطمینان هستند به طور عمومی می‌توانیم بگوییم که داده‌ها کیفی هستند. داده‌های کیفی یا نامرتب را می‌توان با استفاده از ابزارهای مختلف تمیز یا مرتب کرد اما باید تشخیص دهیم از چه ابزاری و به چه صورتی استفاده کنیم و این در گرو فهم درست ما از مجموعه داده‌هاست. در حالت عادی، داده‌های خام هرگز به صورت تمیز و مرتب و ایده‌آل وجود ندارند و این وظیفه ماست که آنها را تمیز کنیم، اما پیش از آن باید مطمئن باشیم که چنین کاری ارزش دارد یعنی داده‌های ما قابل استناد، اعتماد و اطمینان هستند، دوم اینکه تمیز کردن داده‌ها امکان‌پذیر است.

ابزارهای زیادی برای تمیز کردن داده وجود دارد. معمولاً دیتابیس‌ها مملو از اطلاعات تکراری، شبیه به هم و قسمت‌های خالی هستند. تصور کنید در یک مجموعه اطلاعات از سه عنوان نزدیک به هم استفاده شده، مثلاً «بیمارستان مدرس» یا «بیمارستان شهید مدرس» یا «بیمارستان شهید مدرس (ره)». شما می‌خواهید در مجموعه اطلاعاتتان این سه عنوان داده یکی در نظر گرفته شود و در عین حال با عنوان «بیمارستان شهید مدرس» اشتباه گرفته نشود، وقتی با هزاران هزار داده در مجموعه بزرگی از داده‌ها روبرو می‌شوید این نکات اهمیت حیاتی پیدا می‌کنند. («مجموعه آموزشی روزنامه‌نگاری داده (دیتاژورنالیزم)»، ۱۳۹۵/۱۱/۱، بازیابی از سایت دیتاویز)

## □ ابزارها و نرم‌افزارهای روزنامه‌نگاری داده محور

تقریباً ۶۱ نرم افزار برای استفاده در روزنامه‌نگاری داده محور وجود دارد که براساس کاربردهایی که دارند دسته‌بندی می‌شوند. نرم‌افزارهای تحلیل داده‌ها، نرم‌افزارهای تصویرسازی داده‌ها، نرم‌افزارهایی که مخصوص کور رنگی هستند و در تنظیم رنگ دقت بالایی دارند و ... هر کدام از اینها کاربرد متفاوتی دارند.

روزنامه‌نگاری داده محور بدون ابزارها و نرم‌افزارهایی که انبوه داده‌های ما را مرتب می‌کند و آن را به صورت تصویری روایت می‌کند، فقط در حد یک ایده است. آشنایی با این نرم‌افزارها قدم اول است که بدانیم چه انتظاراتی می‌توانیم از آنها داشته باشیم و هر کدام کجا ممکن است به درد ما بخورد.

این ابزارها را می‌توانیم در ۴ گروه عمده تقسیم‌بندی کنیم: گروه اول ابزارهایی که کارشان نگهداری و جمع‌آوری داده است؛ معروف‌ترین این ابزارها اکسل در برنامه آفیس ماکروسافت است. تسلط روی این دسته از نرم‌افزارها از اساسی‌ترین مهارت‌های داده‌نگار است. گروه دوم نرم‌افزارهایی که برای استخراج فایل‌ها و داده‌ها از اینترنت و برای ساختن فایل‌های داده از آنها استفاده می‌کنیم. از بهترین این ابزارها [import.io](http://import.io) است. گروه سوم ابزارهایی که برای تصفیه و تمیز کردن داده از آنها استفاده می‌کنیم که بهترین ابزار این حوزه [Open Refine](http://Open Refine) است. گروه چهارم نرم‌افزارهای مصورسازی داستان است [Tableau](http://Tableau) از بهترین نرم‌افزارهای این حوزه است و کار کردن با آن خیلی ساده است و امکانات زیادی دارد. اگر کار کردن با نقشه و ساختن اینترکتیو مپ برایتان جالب است دو نرم‌افزار خوب دیگر عبارتند از: [MAPBOX](http://MAPBOX) و [CARTODB](http://CARTODB). اگر قصدتان ساختن تایم‌لاین است که بتوانید مطالب را روی یک گذر زمانی نشان بدهید سایت رایگان [Timeline](http://Timeline) نایت لب را ببینید. [Datawrapper](http://Datawrapper) از نرم‌افزارهای اولیه این حوزه است که سال ۲۰۱۱ عرضه شده و طرفداران زیادی دارد. اگر دنبال کارهای پیشرفته نیستید از [infogr.am](http://infogr.am) شروع کنید. [Silk.co](http://Silk.co) از جمله سایت‌هایی است که هم امکان مصورسازی و هم امکان ذخیره اطلاعات را فراهم می‌کند. اگر خواستید نمودار سازمانی، نقشه ذهنی، فلوجارت و کارهایی از این قبیل رسم کنید سایت‌هایی مثل [creately.com](http://creately.com) مناسب این کار است.

**اوپن‌ریفان<sup>۱</sup>، ابزاری برای تصفیه داده‌های خام:** دو نکته را نباید فراموش کرد اول اینکه داده‌ها اصولاً کثیف هستند یعنی داده‌ها تکراری، خالی هستند و شباهت زیادی به هم دارند. داده‌های خام زیاد قابل استفاده نیست مگر اینکه از قبل آن‌را تصفیه کنیم. دوم استفاده از اطلاعات و داده‌های تصفیه نشده کار غیرمسئولانه‌ای است چرا که داده تصفیه نشده می‌تواند شما و مخاطب را گمراه کند درست مثل این است که با چند نفر به عنوان منبع گزارش‌تان مصاحبه کنید و این افراد اطلاعات غیردقیق و نادرست در اختیار شما قرار بدهند که به طور طبیعی نتیجه‌گیری شما غلط خواهد بود و به طبع آن مخاطب‌تان هم گمراه خواهد شد. («مجموعه آموزشی روزنامه‌نگاری داده (دیتاژورنالیسم)»، ۱۳۹۵/۱۱/۱، بازیابی از سایت دیتاویز)

<sup>۱</sup> - اوپن‌ریفان یکی از پروژه‌های گوگل بود که در سال ۲۰۱۰ معرفی شد اما بعد از ۲ سال گوگل آن را رها کرد چون خیلی برنامه به درد بخوری بود به صورت مستقل به کارش ادامه داد و الان یک برنامه متن باز یا اوپن سورس است. اسم آن هم از گوگل‌ریفان به اوپن‌ریفان تغییر یافت.

## □ نتیجه گیری

فضای رقابتی میان رسانه‌ها در کنار رسانه‌های اجتماعی، تحولات نسلی به وجود آمده، اقتصاد نشر، تحولات تکنولوژیکی، امکان دسترسی آزاد به داده‌ها، افزایش روز افزون داده‌ها، ورود به عصر اینترنت اشیا و نیاز رسانه‌ها به جلب رضایت مخاطبان موجب شده تا روزنامه‌نگاری داده‌محور مورد توجه قرار گیرد و چند سالی بیشتر نیست که پایش به رسانه‌های عمومی باز شده است. در این میان تلویزیون با توجه به امکانات گسترده تصویری خود در ارائه بسیاری از خبرهای اقتصادی و اجتماعی که با آمار، اعداد و ارقام و اطلاعات گسترده روبروست، می‌تواند از روزنامه نگاری داده محور استفاده کند.

روزنامه نگاری داده محور برای رسانه‌های خبری به‌ویژه رسانه ملی، ابزاری برای گزارشگری بهتر و محقق کردن وظیفه‌های اساسی آن‌ها و نیز ابزاری برای فراهم کردن اطلاعات مرتبط برای مخاطبان و کمک به درک اطلاعات توسط مخاطبان؛ به کمک روش‌های معنادار است. بخش‌های خبری رسانه ملی می‌تواند با استفاده از داده‌هایی که مراکز همچون بانک مرکزی، سازمان آمار و ... منتشر می‌کنند، اخبار نهفته در داده‌ها را کشف و حقایق بیشتری برای تقویت روایات ارائه کند. مخاطبان در این شیوه نه تنها می‌توانند از اخبار لذت ببرند، بلکه می‌توانند با تصویرسازی‌های داده‌ای و مجموعه داده‌های خام تعامل داشته باشند. همچنین از آنجا که مجموعه داده‌های مورد استفاده در روزنامه نگاری داده محور معمولاً به صورت آنلاین منتشر می‌شوند و برای عموم قابل دسترسی هستند، به گسترش توان اطلاع رسانی، شفافیت و صحت گزارش‌های خبری کمک می‌کنند. بنابر مزایای روزنامه نگاری داده محور و فضای رقابتی میان رسانه ملی و رسانه‌های آنلاین و شبکه‌های اجتماعی، ضروری است که معاونت سیاسی رسانه ملی برای تشکیل روزنامه نگاری داده محور اهتمام ورزد.

## □ منابع

- (۱) سلیمی، مریم، ۱۳۹۶/۶/۲۵، بازیابی از سایت همشهری آنلاین، <https://www.hamshahrionline.ir/news/۳۸۲۲۳۰/>
- (۲) سلیمی، مریم، «دیتا ژورنالیسم؛ آینده روزنامه‌نگاری»، ۱۳۹۶/۸/۱۳، بازیابی از سایت همشهری آنلاین <http://www.hamshahrtraining.ir/news-۳۹۶۳.aspx>
- (۳) «گونه شناسی روزنامه‌نگاری جدید»، مرکز آموزش صدا و سیما
- (۴) «مجموعه آموزشی روزنامه‌نگاری داده (دیتاژورنالیسم)»، ۱۳۹۵/۱۱/۱، بازیابی از سایت دیتاویز <http://datavis.ir/۱۳۹۵/۱۱/data-journalism-courses-ijnet> (۵)
- (۶) کتابچه راهنمای «دیتاژورنالیسم، مقدمات و آموزش، ابزار جمع‌آوری و تحلیل داده»، شبکه بین‌المللی روزنامه‌نگاران (IJNET)