

ارتباط محتوایی

ارتباط داده‌های سایت را به دو گونه می‌توان تقسیم کرد:

۱. داده‌هایی که از ابتدا با هم مرتبط می‌باشند
۲. داده‌هایی که ارتباط اولیه ندارند

هنگامی که فایل تصویری یک سخنرانی به دفتر می‌رسد، به فایل صوتی تبدیل می‌شود. فایل صوتی به صورت متنی پیاده شده و تبدیل به متن بیانات می‌گردد. عکس‌های گرفته شده از این مراسم هم که آمد، چهار نوع داده داریم که همگی با هم مرتبط هستند. البته ارتباطی خام که هنوز به صورت موضوعی تقطیع نشده‌اند.

اگر متن این سخنرانی تقطیع شده و برچسب‌گذاری شود، عملاً فیلم و صوت نیز تقطیع و برچسب‌دار شده است. به شرطی که هنگام پیاده نوار زمان‌بندی‌ها ثبت شود (Sync) و یا پس از تقطیع، قطعات به فایل صوتی منتقل شود (Resync). اگر فایل صوتی از فیلم استخراج شده باشد، تقطیع مزبور به خودی‌خود در فیلم نیز به دست خواهد آمد.

تا این‌جا ما با یک‌بار برچسب‌گذاری (Tag)، سه نوع داده را تقطیع کرده‌ایم. اما تصاویر به تمامی سخنرانی به صورت یکپارچه متصل هستند و معنایی ندارد تقطیع شوند. طبیعتاً اخباری که درباره این سخنرانی به دست می‌رسند نیز مانند تصاویر، به صورت یک‌جا به سخنرانی متصل هستند.

پس در این نوع از داده‌های ورودی ما دو نوع خروجی خواهیم داشت:

۱. داده‌های کلی

- a. متن کامل سخنرانی
- b. صوت کامل سخنرانی
- c. فیلم کامل سخنرانی
- d. اخبار مرتبط با سخنرانی
- e. تصاویر مرتبط با سخنرانی

۲. داده‌های جزئی

- a. قطعات متن سخنرانی
- b. قطعات صوتی سخنرانی
- c. قطعات فیلم سخنرانی

نوع دیگری از ورودی‌ها را هم داریم که از ابتدا خبر یا نوشته می‌باشند. پیام‌ها و انتصابات، تریک‌ها و تسلیت‌ها و تمامی اخباری که از دیدارها و ملاقات‌ها و سایر فعالیت‌ها به دست می‌رسد. این داده‌ها قابلیت تقطیع ندارند. پس در داده‌های کلی قرار می‌گیرند.

۱. ادامه داده‌های کلی

- f. اخبار دیدارها، ملاقات و سایر فعالیت‌ها
- g. پیام‌ها، انتصابات، تریکات و تسلیت‌ها

البته پاره‌ای از پیام‌ها در بعضی مناسبت‌های خاص ممکن است قابلیت تقطیع داشته باشند که به داده‌های جزئی تبدیل خواهند شد، با حفظ وضعیت کلی خود.

داده‌ها و اطلاعاتی نیز در خارج از دفتر نشر تولید می‌شوند. این اطلاعات همگی کلی محسوب می‌شوند و اهمیت‌شان در حدی نیست که نیاز به تقطیع داشته باشند.

به دلیل دو نوع نگاه به اطلاعات؛ کلی و جزئی، با دو نوع فرآیند برچسب‌زنی مواجه خواهیم بود. اطلاعات جزئی هم‌زمان با تقطیع برچسب خواهند خورد و اطلاعات کلی به دو صورت در دستور کار قرار خواهند گرفت؛

آن دسته از اطلاعات کلی که تقطیع شده‌اند نیازی به برچسب خوردن نخواهند داشت، زیرا پس از اتمام برچسب‌زنی قطعات، یک فهرست غیرتکراری از برچسب‌ها به آن داده کلی به صورت خودکار الصاق خواهد شد.

اما پاره‌ای از اطلاعات کلی، مانند: نوشته‌ها، اخبار، تصاویر و تمامی داده‌هایی که از «دیگران» دریافت شده است، باید برچسب‌زنی شوند.

بانک برچسب‌ها

هر برچسبی که به داده‌ای ملصق می‌شود، باید در فهرستی قرار گیرد که با تایپ یک حرف از ابتدا یا وسط، فراخوانی شده و به نظر کارشناس برسد، تا کارشناس بتواند آن را انتخاب نماید و نیازی به ایجاد برچسب جدید یا تکراری نداشته باشد.

بانک مترادفات

مترادفات هر واژه یا اصطلاحی که در برچسب‌ها ذکر شده در بانک مترادفات ذخیره خواهد شد. این بانک هم در جستجوهای یاری می‌رساند و هم هنگام درج برچسب می‌تواند کمک نماید تا موارد تکراری به راحتی شناسایی شده و با اصطلاح اصلی جایگزین شوند. اصطلاح اصلی آن برچسبی است که از میان چند مترادف برگزیده می‌شود تا به عنوان برچسب اولیه نمایش داده شده و به کار گرفته شود. واژه‌ای خواهد بود که بیشتر کاربرد دارد و رایج‌تر است.

گزینش برچسب‌های ویژه

آنچه امروز به عنوان واژه‌های اصلی «جستار» برگزیده می‌شود، در حقیقت پاره‌ای از واژه‌هایی است که به عنوان برچسب مورد استفاده قرار گرفته‌اند، ولی به دلیل اهمیت خاص، گزینش شده‌اند. در یک روند دائمی باید چنین برچسب‌هایی شناسایی شوند و ارتباطشان با سایر برچسب‌ها مشخص شود. به این ترتیب که تمامی برچسب‌هایی که ارتباط محتوایی با این برچسب‌های منتخب دارند به آن‌ها متصل شوند.

ایجاد پرونده

این گزینش و ارتباط برچسب‌ها مقدمه‌ای است برای تولید «پرونده» برای داده‌ها و اطلاعات. پرونده مهم‌ترین بخش محتوایی سایت خواهد بود که تمامی انواع داده را حول یک موضوع تجمیع می‌کند. پرونده در حقیقت توسعه‌یافته همان «جستار» می‌باشد.

البته پرونده‌های ویژه بر جای خواهند ماند؛ پرونده‌هایی که موضوع بسیار خاص و ویژه دارند و در حقیقت گزینشی از پرونده‌های سایت خواهند بود که به دلیل اهمیت فراوان به صورت پرونده ویژه معرفی خواهند شد.

برای تسهیل ایجاد برچسب به غیر از روش مرسوم، دو روش دیگر پیشنهاد می‌گردد. در روش نخست، می‌توان از متن بیانات به صورت رایانه‌ای و خودکار برچسب‌های اولیه را استخراج کرد، به روش ذیل:

۱. تمامی واژه‌ها تفکیک شده و در یک بانک اطلاعات هر کدام به صورت یک رکورد قرار می‌گیرند
۲. واژه‌های تکراری حذف شده و یک فهرست غیرتکراری در اختیار کارشناس قرار داده می‌شود
۳. غربال‌گر تمامی افعال، حروف و صفت‌ها و قیده‌های کم‌ارزش را حذف می‌کند
۴. واژه‌های باقی‌مانده مترادفیابی شده و موارد هم‌معنا به یکدیگر متصل می‌شوند
۵. از گروه‌های مترادف، واژه اصلی انتخاب می‌شود

اکنون حجم زیادی برچسب به دست خواهد آمد که می‌تواند پایه اصلی عملیات برچسب‌زنی قرار بگیرد. خصوصاً با توجه به این نکته که تمامی این واژگان از ادبیات رهبری استخراج شده‌اند.

روش دیگر مشابه همان است که در «آرشیو پیشرفته» مورد استفاده قرار گرفته است؛ ایجاد رده‌هایی از اطلاعات پایه. بر اساس موضوعات نظام، اقشار، مناطق جغرافیایی و شهرها و... می‌توان رده‌هایی از اطلاعات پایه را در نظر گرفت و تمامی زیرمجموعه‌های آن را احصاء نمود و به عنوان بخشی از برچسب‌ها مورد استفاده قرار داد.

یکپارچه‌سازی اطلاعات

اگر یک نظام واحد در برچسب‌زنی اطلاعات به کار گرفته شود، به این معنی که تمامی برچسب‌ها در یک استخر ریخته شده و از همان اخذ شوند، ارتباط تمامی داده‌ها تضمین خواهد شد. خصوصاً با توجه به این‌که بعضی از اطلاعات از ابتدا با هم مرتبط بوده‌اند و در نظر گرفتن این ارتباط اولیه، عملیات برچسب‌زنی را ساده‌تر نموده است.

این یکپارچگی از یک‌سو پرونده‌های موضعی قدرتمند و جامعی ایجاد می‌کند که در حقیقت هر برچسب ما را به یک پرونده می‌رساند. از دیگر سو اطلاعات کلی دیگران و خارج از دفتر را به اطلاعات داخلی متصل می‌نماید. این ارتباط می‌تواند بازتاب فرمایشات رهبری در رسانه‌ها و جامعه را قابل رهگیری نماید. این داده‌ها را می‌توان در قالب بولتن‌های داخلی به استحضار مسئولان دفتر رساند.

سرویس‌های موجود

فضای فعلی سایت علی‌رغم حجم زیاد داده‌ها و نظم بسیار خوبی که دارند، اندکی سبب سردرگمی می‌شود. یک کاربر معمولی در ابتدای ورود نمی‌داند باید در بیانات جستجو کند یا سراغ آرشیو پیشرفته برود و یا در پرونده‌ها دنبال مطلب مورد نظر خود بگردد و شاید هم آن‌چه می‌خواهد در بخش دیگران بهتر تبیین شده باشد.

اگر وحدت تمامی اطلاعات حول «پرونده» در نظر گرفته شود، در حقیقت «جستار» توسعه یافته و محور تمام سایت می‌شود. تمامی داده‌ها و اطلاعات حول پرونده قابل ملاحظه است.

یکی از مهم‌ترین استفاده کاربران نیز خبرخوان یا RSS است. اگر چه در حال حاضر در بخش‌های مختلف سایت این گزینه قرار دارد، ولی جای یک صفحه خالی است که بتوان گزینه‌های مختلف را تیک زد و یک RSS دلخواه دست یافت که تنها مثلاً مطالب مربوط به «اقتصاد» و شهر «قم» را نمایش دهد. بسیاری از سایت‌های خبری امروزه از این نوع RSS‌های قابل شخصی‌سازی استفاده می‌نمایند.