

(۱) شما به عنوان یک متخصص فناوری اطلاعات اگر با حجم عظیمی از اطلاعات روبرو شوید چه راهکارهایی جهت سازماندهی اطلاعات و دسترسی بهتر به اطلاعات پیشنهاد می کنید؟ به نظرتون نقش نمایه در این سازماندهی اطلاعات چیست؟

خلاصه پیشنهاد من : به عنوان کسی که در حوزه فناوری مشغول به کاره می باشد، قطعاً راه کارهای نرم افزاری و مدیریت اطلاعات از طریق تکنولوژی های بروز را پیشنهاد میدم ک باید اول کار فهمید که این اطلاعات چه نوع داده های هستند که بتوان برای آن تصمیم گیری کرد چون هر سیستم روش مورنظر خودشو پیشنهاد میشه مثلاً سیستم های خبره باید با هوش مصنوعی و الگوریتم های مخصوص خودش هندل بشه ولی من ابتدایی ترین سیستم ها رو در نظر میگیرم و پیشنهاد های خودمو میدم که جزء ابتدایی اون سیستم مدیریت پایگاه داده (DBMS) که با توجه به نیاز و بسترهای مورد استفاده برای آن میتوان از مدیریت پایگاه های داد های مختلفی از جمله `Sql Server , Oracle, Mysql` که این نوع پایگاه داده های از نوع رابطه ای میباشند و نوع دیگر آن که به آن ها `Nosql` میگویند مانند `redis , mongodb` که این نوع پایگاه داده ها معمولاً به خاطر ساختار کلید و مقدار (`key & Value`) سرعت بالایی در بازیابی داده ها در داده های بزرگ دارن ، که انتخاب هر کدام بسته به سیاست های شرکت و چندین پارامتر دیگر مورد انتخاب قرار میگیرد. بعد از انتخاب سیستم مدیریت پایگاه داده باید سیستم نرم افزاری که بتوان داده ها خام را به اطلاعات (`data mining`) دسته بندی شده و قابل فهم برای کاربران قرار داد که این اطلاعات را میتوان با توجه به نمودار ها و گزارش ها نمایش داد. نکته ای که پیش میاد اینکه مهم بودن نحوه ذخیره سازی داده ها در سیستم مدیریت پایگاه داده که بحث تحلیل پروژه و دیاگرام های پروژه و دیتایس می باشد که بتوان به بهترین راه ممکن داده ها دسته بندی و ذخیره کرد که بتوان برای بازیابی آن دچار چالش نشده و با سرعت بالا و کارایی های بالاتر (منظور بیشترین گزارش ها و مفید بودن آن) بتوان این کارها را انجام داد. و حالا بحث زبان های برنامه نویسی پیش میاد که هر کدام ویژگی های خودشونو دارن که با توجه به نیاز انتخاب میشوند و برای دسترسی بهتر (`UI & UX`) و راحتی کار کاربران باید نرم افزار هایی توسعه داد که بیشترین تقاضا را دارند (معمولاً در این زمان چون پلتفرم تلفن همراه رواج پیدا کرده گوشی های همراه مدنظر ماست) که بازم بستگی به مدل اطلاعات سازمان و نوع کاربران سازمان و سیاست های سازمان انتخاب میشود.

نظر من نقش نمایه در این سازماندهی اطلاعات: نمایه یا همون ایندکسینگ باعث میشه که ما خیلی راحت تر و دسته بندی تر به اطلاعات ذخیره شده دست پیدا کنیم که این نمایه کردن اطلاعات با توجه به الگوریتم های خاص خود و راه کار های زیاد که با توجه با موجودیت داده های راه کارهای مختلفی برای اون وجود دارد اگر بخوام خلاصه بگم نمایه سازی باعث میشه ما به اطلاعاتمون ارزش بدیم یعنی با راه کاری های نمایه سازی ما اطلاعات را به داده های شاخص و مرتب تبدیل کردیم که این نوع اطلاعات نمایه شده باعث سرعت بخشیدن به جستجوهای کاربر، منظم تر شدن اطلاعات، بخش بندی اطلاعات که با این کار علاوه بر بهتر شدن سرعت بازیابی اطلاعات و جستجوهای منصر به فرد و خاص با توجه به دسته بندی ها و... باعث میشود داده ها بتوانند با دیگر اطلاعات دیگر

ارتباط داشته باشند که تشکیل یه درخت را بدهند که این نوع داده ها در هوش مصنوعی و تکنیک های شبکه های عصبی مورد استفاده قرار میگیرد.

۲) به نظر شما فرایند سازماندهی، جستجو و بازیابی اطلاعات در موتورهای جستجو بهتر است به چه صورتی باشد تا بتوان نیاز کاربران را بهتر برآورده نمود.

در ابتدا کاربر درخواست خود را از طریق User interface وارد میکند که این مسئله که ui و ux خوب خیلی مفیده که هرکاری با هر میزان تخصصی بتونه با این سیستم کار بکنه.

اطلاعات سازمان ها باید به صورت سازماندهی شده و ایدکس گذاری شده در پایگاه های داده ذخیره بشه که بتوان با توسعه موتور سرچ از داخل ایندکس هایی که ثبت شده بتوان بهترین نتیجه را همراه بشه، نتایج حاصل در بخش Ranking بر اساس میزان مرتبط بودنشان با درخواست کاربر طبقه بندی میشوند و نتایج به کاربر نمایش داده میشود، موتورهای جستجوگر باید از نظرات کاربر درمورد نتایج نمایش داده شده برای بهبود نتایج بعدی استفاده میکنند. بطور مثال شما در گوگل وقتی روی یکی از نتایج جستجو کلیک میکنید، گوگل امتیاز ویژه ای را برای آن مورد در نظر میگیرد. زیرا شما از بین ۱۰ نتیجه ی نمایش داده شده در آن صفحه فقط روی آن مورد خاص کلیک کردید. همچنین جدیداً گوگل سرویس دیگری را در بخش نتایج خود فراهم کرده است که به کاربران امکان میدهد تا نتایجی که از نظر آنها با موضوعی که جستجو کرده اند بیشتر مرتبط است، به بالای سایر نتایج منتقل کنند و با اینکار گوگل میفهمد که شما در مقابل جستجوی آن عبارت خاص، این نتیجه را از بین سایر نتایج ها بیشتر پسندیده اید. همچنین میتوانید نتایج نامناسب را از داخل نتایج نمایش داده شده حذف کنید، مسلماً گوگل از این Feedback ها برای بهبود نتایج جستجویش استفاده میکند.

در بخش جستجوی تصاویر گوگل وقتی شما یک عبارت را جستجو میکنید و نتایجی را به شما نمایش میدهد، شما فقط بر روی چند عکس خاص کلیک میکنید. گوگل از این طریق میفهمد که عکس موردنظر با جستجوی شما رابطه ی نزدیک تری دارد تا عکس هایی که رویشان کلیک نکردید.

پس مطمئن موتورهای جستجوگری بزرگی مثل گوگل باید روش های ایندکس گذاری خاصی داشته باشند تا بتوانند این عمل را به بهترین نحوه انجام داده و نتایج خروجی به کاربر را نزدیک ترین جستجو به درخواست کاربر میباشد، و یه نکته مهم باید موتورهای جستجو گر دیگر از گوگل یاد بگیرن اینکه باید از نظرات کاربران استفاده کرد و سیستم جستجوگر قابلیت یادگیری یا همون بحث ماشین لرنینگ یا همون یادگیری ماشین(Machine learning) که بحث مهمی توی هوش مصنوعی

و شبکه های عصبی میباشد که باید یادگیری موتور جستجوگر از نتایج به دست آمده و فیدبک از کاربران میتواند به اطلاعات امتیاز داده و یا نحوه ایندکس گذاری ها را تغییر و بهبود ببخشد که در نتایج بعدی بتواند بهترین پیشنهاد ها را بدهد. و اینکه سازمان باید الگو مشخص شده و واحد برای ذخیره و ایندکس گذاری داده ها و اطلاعات خود در سیستم خود لحاظ کند.

۳) به نظر شما کاربران در سازماندهی و نمایه سازی اطلاعات چه نقشی دارند؟

کاربر میتواند با توجه به دانش و حرفه خود نقش های مختلفی داشته باشد از جمله با جستجو کردن در اطلاعات میتوان به موتور جستجوگر به یادگیری آن کمک کرد و یا با کمک کاربران میتوان به تکمیل اطلاعات سازمان برای جامع کردن آن که مثال اون میشه به سرویس گوگل مپ که با کمک کاربران دارع محل های مختلف نقشه را تکمیل میکند که ایرانی هم این موضوع هم هستش که واسه مشهد اسمش نقشه همراه مشهد که این امکانو به کاربر میده که حالا یا کسب کارو خودشو و یا محل های مختلفی که اطراف خودش هستشو با متن یا عکس معرفی کنه و مکان معرفی شده به نامه خودش ثبت نرم افزار بشه و امتیاز دریافت کنه. و یا مثالی بروزی که میشه زد بحث رمز ارزها که یکیش بیت کوین معروف میباشد که میشه گفت رمز ارزها اطلاعات رمز نگاری شده و نمایه شده ای هستند که هرکس برای به دست آوردن آن باید این نمایه رمز نگاری شده رو بشکنه تا بتونه به این ارز دست پیدا کنه که نقش کاربر در این جا به بالا رفتن میزان ارزش رمز ارز و حتی سخت تر کردن ارزها از طریق نمایه کردن و کد کردن سخت تر آن ها برای بدست آوردن آن میتوان نام برد.

۴) به نظر شما آیا می توان اطلاعات را به گونه ای سازماندهی نمود تا کاربران بتوانند معنا و مفهومی که مدنظر دارند را بازیابی کنند؟ اگر اینگونه است سازماندهی اطلاعات باید چگونه باشد؟

به طور قطع نمی توان به این سوال بله گفت چون خواسته های هرکابر و طرز هرکابر با هم فرق دارن و این فرق باعث میشه که در نتایج جستجو هرکابر تاثیر گذار باشه، به همین علت هستش که ما آموزش هایی تحت عنوان روش های جستجو در وب یا گوگل را داریم که بتونیم با طرز بهینه و کارآمد سرچ گوگل متوجه بشیم تا در سرچ کردمون به بهترین نتیجه و هدفمون برسیم. ولی باز هم بعد یادگیری ماشین و اینجور موارد میتواند ضریب خطا های این بخشو به کمترین حده خودش برسونه. برای سازماندهی اطلاعات سازمان بسته به نوع سازمان و بزرگی سازمان و سیاست ها و عناوین دیگر باید در نظر گرفته شود ولی خوب در عصر حاضر فناوری اطلاعات معمولاً از روش سنتی و دستی زیاد استفاده نمیشه (بازم بستگی به سیاست های سازمان دارد) به همین علت روش خودکار یا ماشینی یا پردازش اطلاعات توسط نرم افزار ها مورد استفاده قرار میگیرد. برای همین مورد معمولاً قوانین سئو SEO سایت ها و کسانی که مطالب انتشار میدن در سایت ها حالا با هر عنوان سایتی، سعی میکنن علاوه محتوا های همه جانبه که هرکسی با توجه میزان سواد درک کنه (بازم بسته به نوع محتوا دارد تخصصی باشد یا عمومی) باید

قوانین تگ گذاری یا همان هشتگ گذاری معروف را به بهترین نحوه و شکل انجام دهد که هرکابر با هر ذهنیتی بتواند پشتیبانی کند که سایت های جستجوگر با توجه با خزنده وب و الگوریتم های مخصوص خودشان بتوانند این داده ها رو به بهترین نحوه ذخیره و ایندکس گذاری بکنند تا جوابگوی تمام کاربران با مدل های مختلف جستجو باشد.

(۵) به نظر شما از چه نظام هایی می توان جهت نمایه سازی اطلاعات استفاده نمود؟ آنها را شرح دهید.

دو روش برای این مسئله وجود دارد یکی اینکه داده های و واژه های قابل اندیکس گذاری را از داخل اطلاعات بدست آوریم روش بعدی اینکه واژه های قابل تطبیق با واژه هایی که در روش اول بدست آمده را تعیین کنیم، خوب مطمئن روش دوم نیاز به هزینه بالاتر (از لحاظ نرم افزاری و سخت افزاری و نیروی کار و هزینه زمان و...) دارد ولی روش اول معمولا کمتر است. ولی خوب باید در نمایه سازی داده ها از دو روش استفاده شود یعنی ترکیب دو نظام که همینجور که گفتم حتما باید روش اول که به نام واژه گزینی معروف است را به درستی و کامل انجام بدهیم که بتوانیم روش دوم که واژه گذاری نام دارد را به طور کامل انجام پذیرد.

و برای به اجرا در آوردن این نظام های میتوان از قوانین از جمله پیش همارا و یا پس همارا استفاده کرد. که من خودم به شخصه پس همارا انتخاب میکنم چون میتوان از قوانین عبارت های باقاعده regular-expressions که خیلی راحت میتوان به نتیجه مورنظر دست پیدا کرد و حتی تشکیل درخت و دسته بندی موضوعی و...

(۶) شما به عنوان متخصص فناوری اطلاعات اگر مجموعه ای از تصاویر داشته باشید براساس استانداردها چگونه آنها را سازماندهی و نمایه سازی می کنید.

باید تصاویر را براساس موضوعات دسته بندی و بخش بندی که و برای هر تصویر از قوانین تگ گذاری و کلیدواژه و .. استفاده کرد تا بتوان موضوعات مرتبط با آن به خوبی جستجو بشن و هم بتوان براساس برچسب های موضوعی کلید واژه و دسته بندی و ... به کاربر پیشنهاد هایی مشابه را داد.

ولی اگر تصاویر توسط کاربر به سیستم اطلاعات ما وارد شود علاوه بر گرفتن دسته بندی و تگ های توسط کاربر (که میتواند خطای کاربری دچار شود) باید پردازش تصاویری که با تشخیص بافت و رنگ و... با توجه به الگوریتم های خاص و سیستم خبره تصاویر برای تشخیص موضوع تصویر و مشخصه های دیگر برای ایندکس گذاری این عکس های وارده درست کرد، موضوع دیگر که میتوان بهش اشاره کرد اینکه برای جلوگیری از خطای کاربری و یا موارد که از سمت کاربر اتفاق میوفتد باید به کاربر ها یا حتی صفحات رتبه بندی و یا نمره داد که اگر کاربری

دچار خطا بود نمره پایین تری داشته باشد و الگوریتم های پیشرفته تری نسب به کاربر های نمره بالا برای پردازش تصویر و ایندکس گذاری تصویر انجام داد.