

بهبود تولید پاسخ در انجمن‌های پرسش و پاسخ با استفاده از مدلسازی دانش کاربر-محتوا

سعید فرضی

استادیار دانشکده کامپیوتر دانشگاه صنعتی خواجه نصیرالدین طوسی

چکیده

انجمن‌های پرسش و پاسخ، به دلیل فراهم نمودن بستری برای تبادل دانش و تعامل با متخصصین سرتاسر جهان، به یکی از محبوب‌ترین روش‌های کسب دانش تبدیل شده‌اند. امید به دستیابی سریع به پاسخ‌هایی صحیح و کامل از جانب بهترین متخصصین، کاربران وب را به عضویت در انجمن و ثبت پرسش‌های خود تشویق می‌کند. هر پرسش جدید، بلافاصله پس از ثبت در انجمن، به سوی متخصصین مرتبط که واجد شرایط پاسخگویی هستند هدایت می‌شود. اگرچه متخصصین در اسرع وقت نسبت به پاسخگویی به پرسش‌ها اقدام می‌کنند، با این وجود نمی‌توان از تأخیر ذاتی بین زمان ارسال پرسش تا دستیابی به یک پاسخ صحیح چشم پوشید. در این راستا، هدف اصلی پژوهش جاری، بهبود تولید خودکار پاسخ زبان طبیعی برای پرسش کاربر است تا بدین وسیله، پرسش‌گر به سرعت به یک پاسخ باکیفیت و صحیح دست یابد. در این راستا، سیستمی برای تولید پاسخ مبتنی بر دانش نهفته در محتوا و دانش کاربر متشکل از پنج مؤلفه ارائه شده است. مؤلفه‌ی اول، یک بازنمایی معنایی از دانش نهفته در پرسش ارائه می‌دهد و دومین مؤلفه، به مدلسازی دانش پیش‌زمینه‌ی پرسش‌گر براساس سوابق فعالیت‌های وی در انجمن می‌پردازد. استنباط دانش کاربر می‌تواند به درک مضمون پرسش کمک نموده و فضای جستجو برای پاسخ را محدود نماید. پس از آن، مؤلفه‌ی سوم، مجموعه‌ای از پاسخ‌های تأیید شده برای پرسش‌های متضمن در بایگانی انجمن را که منبعی غنی از دانش و الگوهای انسانی محسوب می‌شوند، استخراج می‌کند. بطور همزمان، در مؤلفه‌ی چهارم، مجموعه‌ای از حقایق مرتبط با پرسش از یک گراف دانش خارجی واکشی می‌شوند. در پایان، با تلفیق این دو مجموعه از پاسخ‌های کاندید، پاسخی به زبان طبیعی به کاربر ارائه می‌گردد.

کلید واژه‌ها: انجمن پرسش و پاسخ، بازیابی پرسش، مدلسازی دانش پیش‌زمینه‌ی کاربر، دانش محتوا، تولید پاسخ

افزایش روزافزون اطلاعات منتشر شده در بستر وب، آن را به مخزنی^۱ گسترده و غنی از اطلاعات مبدل کرده است. از این رو، در اکثر مواقع، محیط وب، اولین منبعی است که مردم برای رفع نیازهای اطلاعاتی^۲ خود به آن مراجعه می کنند [1-4]. اما سؤال اصلی که به ذهن خطور می کند این است که کاربران چگونه می توانند در میان این حجم انبوه از اطلاعات، با صرف کمترین زمان ممکن، به پاسخی مطلوب برای پرسش های خود دست یابند؟ در پاسخ به این سؤال، مروری کوتاه بر تحولات صورت گرفته در حوزه ی بازیابی متن، از دهه ۱۹۹۰ تاکنون، خالی از لطف نیست. برخی از مهمترین رخدادها در این بازه ی زمانی، در شکل **Error! No text of specified style in document.** ۱- نشان داده شده اند. همانطور که ملاحظه می شود، موتورهای جستجو^۳ را می توان پیشگام در عرصه ی بازیابی صفحات وب و اسناد متنی مرتبط با پرسمان^۴ کاربر دانست. رشد سریع وب، محققان را در جهت ساخت موتورهای جستجوی پیشرفته تر و تکامل آنها سوق داد [5]. در حال حاضر گوگل، بینگ و یاهو به ترتیب، محبوب ترین موتورهای جستجو در میان کاربران اینترنت هستند^۵.

^۱ Repository

^۲ Information need

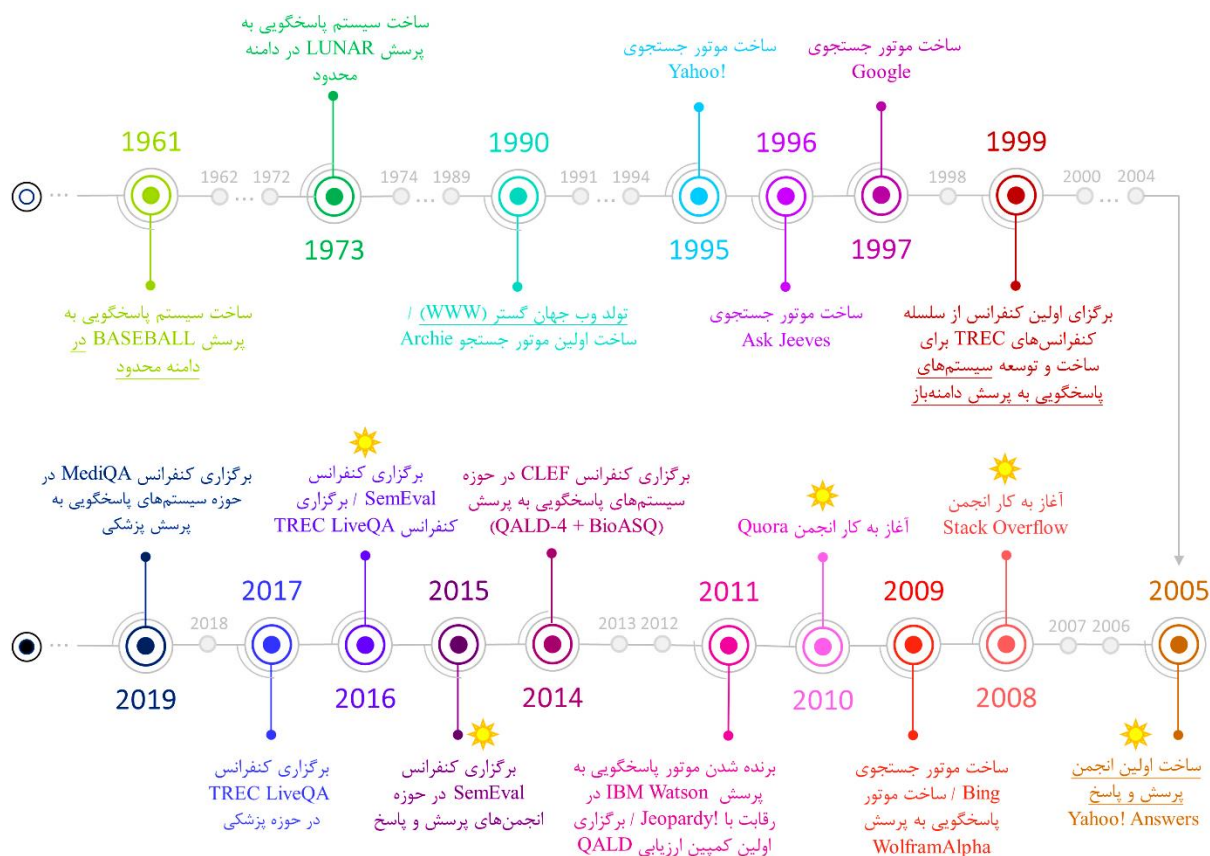
^۳ Search engine

^۴ Query

^۵ طبق آخرین گزارش منتشر شده برای سال ۲۰۲۱ در دو وبسایت

الف) <https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines>

ب) <https://gs.statcounter.com/search-engine-market-share>



شکل ۱- خط زمانی برخی از مهمترین تحولات رخ داده در حوزه بازیابی

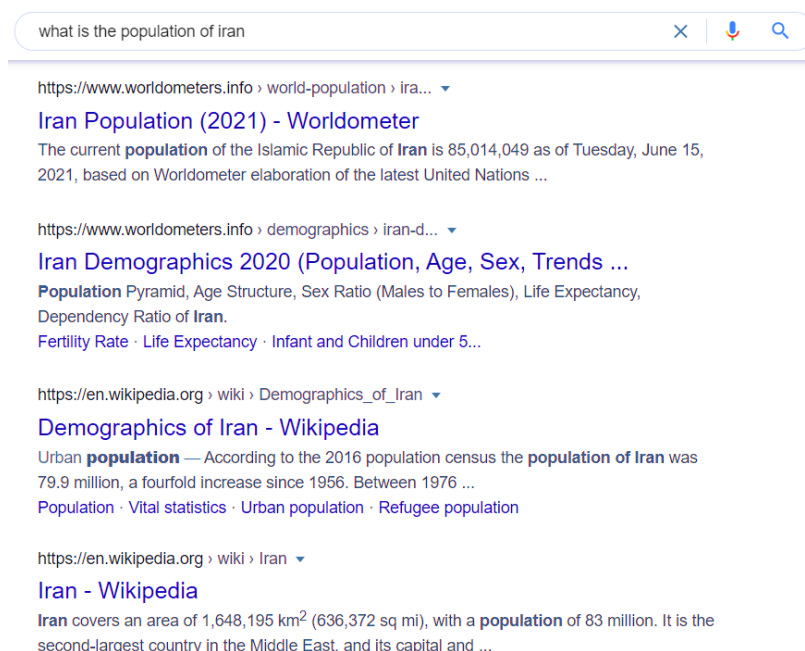
متن و پاسخگویی به پرسش (رخدادهای مربوط به انجمن‌های پرسش و پاسخ، با نماد مشخص شده‌اند).

موتورهای جستجو، نیاز اطلاعاتی کاربر را در قالب پرسمانی متشکل از کلیدواژه‌ها^۶ دریافت می‌کنند [6, 7]. پس از آن، محتوای اسناد را با کلیدواژه‌های پرسمان مطابقت داده و به هر سند، امتیازی مبنی بر میزان مرتبط بودن^۷ آن سند با نیاز اطلاعاتی نسبت می‌دهند. در پایان، اسناد را بر اساس امتیازشان به صورت غیرصعودی مرتب نموده و به کاربر ارائه می‌دهند [8, 9]. به عنوان مثال، در شکل ۲-**Error! No text of specified style in document.** بخشی از خروجی موتور جستجوی گوگل در پاسخ به پرسش "جمعیت ایران چقدر است؟" نمایش داده شده است.

^۶ Keyword

^۷ Relevance score

همانطور که مشاهده می‌شود، کاربر با وارد شدن به اولین پیوند، جمعیت ایران در سال ۲۰۲۱ را می‌یابد و چنانچه در جستجوی جمعیت ایران در سال‌های قبل باشد، باید به مطالعه و بررسی پیوندهای این لیست ادامه دهد.



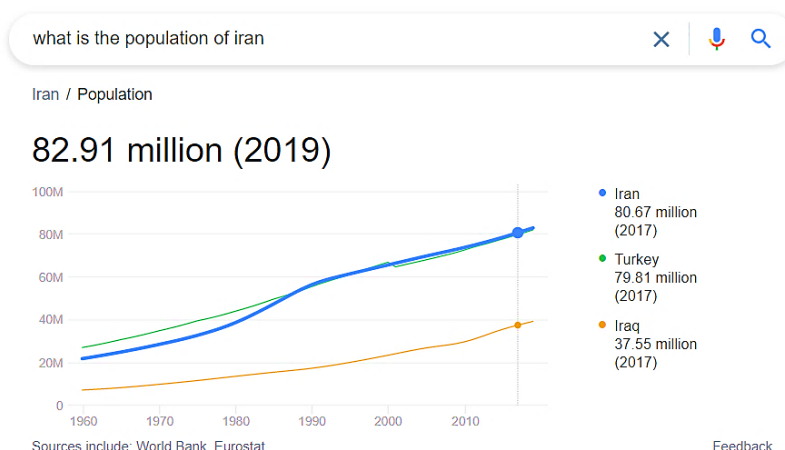
شکل ۲-Error! No text of specified style in document. چهار نتیجه اول ارائه شده توسط موتور جستجوی گوگل، در پاسخ به پرسش "جمعیت ایران چقدر است؟"

اگرچه موتورهای جستجو، دستیابی به اسناد مرتبط با نیاز اطلاعاتی را برای کاربران تسهیل و تسریع نموده‌اند، با این حال کاربران در خوش‌بینانه‌ترین حالت باید چندین صفحه را برای دستیابی به پاسخ نهایی مرور کنند و چنانچه به نتیجه مطلوب دست نیابند، ناگزیر به اصلاح پرسمان و جستجوی مجدد آن در موتور جستجو هستند [10, 11]. در واقع، یکی از مهمترین نقطه ضعف‌های موتورهای جستجو، ناتوانی آنها در استنتاج و جمع‌بندی نتایج است که این امر بر عهده‌ی خود کاربر گذاشته شده است [12]. تلاش برای برطرف نمودن این نقطه ضعف، با حمایت جدی کنفرانس سالیانه‌ی بازیابی متن^۸ (TREC) در سال ۱۹۹۹، مسیر تحقیقاتی تازه‌ای را تحت عنوان ساخت و توسعه‌ی سیستم‌های پاسخگویی به پرسش^۹ (QA) شکل داد [13, 14]. یک سیستم QA موظف است

^۸ Text REtrieval Conference (TREC)

^۹ Question Answering (QA)

به جای غرق کردن کاربر در سیلی از اسناد یا پاراگراف‌های مرتبط، به سرعت او را به یک پاسخ صحیح، کامل و در عین حال خلاصه برساند [2, 8]. به عنوان مثال، در شکل **Error! No text of specified style in document.** پاسخ مستقیم گوگل برای پرسش "جمعیت ایران چقدر است؟" نمایش داده شده است. همانطور که ملاحظه می‌شود، گوگل علاوه بر ارائه یک عدد، نموداری تعاملی نیز به کاربر ارائه داده است. چنین سیستم‌هایی که در پاسخ به پرسش کاربر، یک تصویر، نمودار، صوت یا ویدئو ارائه می‌دهند، سیستم‌های پاسخگویی به پرسش چندرسانه‌ای^{۱۰} نامیده می‌شوند [11, 15, 16]. لازم به تأکید است، پژوهش جاری تنها بر توسعه سیستم‌های متنی متمرکز است.



شکل **Error! No text of specified style in document.** ۳. پاسخ مستقیم ارائه شده توسط موتور جستجوی گوگل به پرسش "جمعیت ایران چقدر است؟"

با تعریف سیستم‌های QA، چالش‌های جدیدی نیز پیش روی محققین قرار گرفت [1]. آنها باید راهکارهایی برای پاسخگویی بی‌درنگ^{۱۱}، درک پرسش^{۱۲}، دسته‌بندی پرسش^{۱۳}، بازیابی قطعه متن^{۱۴}، جمع‌بندی یا خلاصه‌سازی پاسخ^{۱۵}، مدیریت و بهره‌برداری از منابع چندزبانی^{۱۶}، چگونگی شخصی‌سازی^{۱۷} نتایج و ... ارائه می‌دادند. به میان

¹⁰ Multimedia question answering

¹¹ Real-time

¹² Question understanding

¹³ Question classification

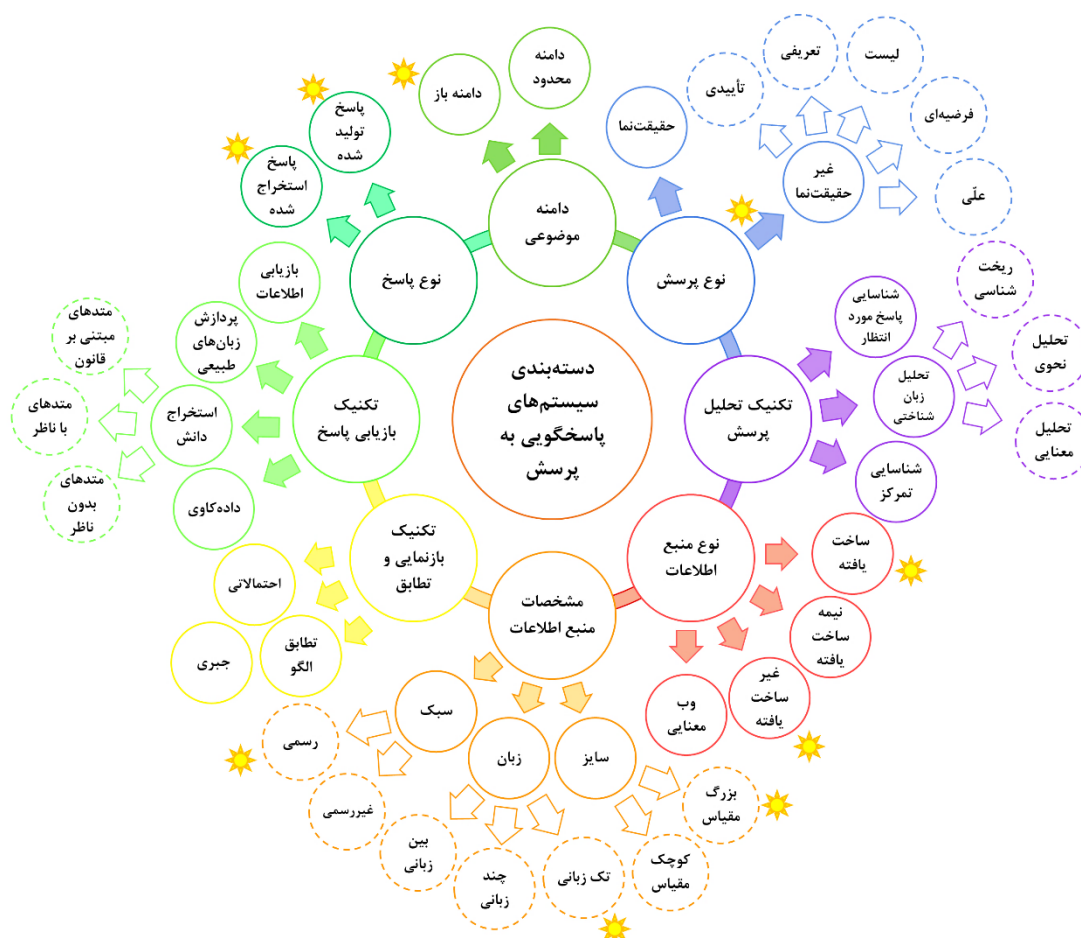
¹⁴ Passage retrieval

¹⁵ Answer summarization

¹⁶ Multilingual

¹⁷ User profiling (Personalization)

آمدن این چالش‌ها باعث شد، تنوعی از پژوهش‌ها انجام شوند که هر کدام با دیدگاهی متفاوت، به یک یا چند مورد از این چالش‌ها پرداخته بودند. در شکل ۴-Error! No text of specified style in document، دسته‌بندی سیستم‌های QA از جنبه‌های مختلف، برگرفته از مطالعات مروری [1] و [3]، نمایش داده شده است. دسته‌بندی‌هایی که با پژوهش جاری، مطابقت قطعی دارند در شکل با نماد ☀ مشخص شده‌اند.



شکل ۴-Error! No text of specified style in document. دسته‌بندی سیستم‌های پاسخگویی به پرسش از جنبه‌های مختلف

هنوز یک دهه از روی کار آمدن سیستم‌های QA نگذشته بود که زمزمه‌هایی مبنی بر کارایی پایین این سیستم‌ها در ارائه‌ی پاسخ مناسب به پرسش‌های غیر حقیقت‌نما^{۱۸} به گوش رسید [17, 18]. در واقع، سیستم‌های QA

¹⁸ Non-factoid

موجود، علیرغم برخورداری از صحت^{۱۹} بالا در پاسخگویی به پرسش‌های حقیقت‌نما^{۲۰}، در بازیابی پاسخ برای پرسش‌هایی که چرایی^{۲۱} یا چگونگی^{۲۲} یک مسئله را خواستار بودند غالباً ناموفق عمل می‌کردند [1, 3, 18]. این عدم موفقیت از دو عامل نشأت می‌گرفت: ناتوانی سیستم‌ها در درک نیازمندی کاربر و فقدان سرنخ‌های اطلاعاتی کافی، واضح و سراسر در منبع اطلاعات برای ارائه پاسخی مستقیم به کاربر [18]. اینجا بود که انجمن‌های پرسش و پاسخ^{۲۳} (CQA)، به عنوان یک ابزار جایگزین برای رفع نیازهای اطلاعاتی پیچیده^{۲۴} کاربران، پا به عرصه‌ی وجود گذاشتند. سیستم‌های CQA به کاربران خود این امکان را می‌دهند که پرسش‌های خود را به صورت برخط^{۲۵} ارسال نموده و پاسخ‌ها، عقاید^{۲۶} و دیدگاه‌های مختلف را از متخصصین^{۲۷} سرتاسر جهان دریافت نمایند [19]. از جمله انجمن‌های شناخته شده در زبان انگلیسی می‌توان به Yahoo! Answers^{۲۸}، Stack Overflow^{۲۹} و Quora^{۳۰} اشاره کرد.

در مقایسه‌ی سیستم‌های CQA با سیستم‌های QA، به چند تفاوت کلیدی می‌توان اشاره کرد:

- (۱) **نوع پرسش:** در سیستم‌های QA، اکثر پرسش‌ها، کوتاه، تک‌جمله‌ای و حقیقت‌نما هستند [17]. در مقابل، پرسش‌های ثبت شده در CQA، اغلب طولانی و چندجمله‌ای هستند [17, 20]. طولانی بودن پرسش‌ها از دیدگاه کاربران انجمن، یک نقطه قوت به شمار می‌آید. چراکه اکثر افراد تمایل دارند نیاز اطلاعاتی خود را با توصیفات دقیق و مبسوط مطرح کنند [21]. شکل **Error! No text of specified style in**
- document-۵** نیز اثبات این مدعاست. در این شکل، محور افقی بیانگر طول پرسش و محور عمودی

¹⁹ Accuracy

²⁰ Factoid

²¹ 'Why' questions

²² 'How' questions

²³ Community Question Answer (CQA)

²⁴ Complex

²⁵ Online

²⁶ Opinions

²⁷ Experts

²⁸ یاهو در ۵ آوریل ۲۰۲۱ اعلام نمود که Yahoo Answers! به کار خود خاتمه داده است. (آدرس سابق انجمن: <https://answers.yahoo.com>)

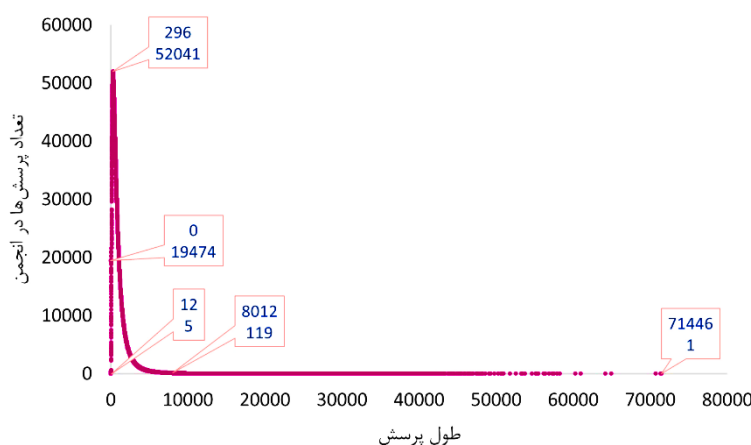
²⁹ <https://stackoverflow.com/> (ملاقات شده در ۱۴۰۰/۰۵/۲۳)

³⁰ <https://www.quora.com/> (ملاقات شده در ۱۴۰۰/۰۵/۲۳)

مبین تعداد پرسش‌های ثبت شده در انجمن Stack Overflow است. میانگین طول پرسش‌ها در این انجمن، برابر با ۱۱۰۹ واژه^{۳۱} و پرتکرارترین طول پرسش، ۲۹۶ است.

(۲) نوع و کیفیت پاسخ: در سیستم‌های QA، پاسخ‌ها به صورت خودکار از اسناد ساخت‌یافته یا غیرساخت‌یافته استخراج شده و معمولاً از کیفیت بالایی برخوردار هستند. در حالیکه در سیستم‌های CQA، پاسخ‌ها توسط انسان تولید شده و بسته به نوع انجمن و سطح تخصص پاسخ‌دهنده‌ها، سطوح کیفی متفاوتی دارند [17].

(۳) تأخیر زمانی: در سیستم‌های QA، پاسخ بلافاصله پس از ارسال پرسش ارائه می‌شود. این درحالیست که در سیستم‌های CQA، هر کاربر، پس از ارسال پرسش باید منتظر بماند تا سایر کاربران به آن پاسخ دهند. مدت زمان انتظار برای دریافت یک پاسخ رضایت‌بخش^{۳۲} توسط کاربر ممکن است از چند ساعت تا چند روز متغیر باشد [22]. این تأخیر، از مهمترین نقطه ضعف‌های CQA محسوب می‌شود.



شکل ۵- Error! No text of specified style in document. توزیع طول پرسش در انجمن Stack Overflow

در سال‌های اخیر، ارائه‌ی راهکار برای کاهش تأخیر زمانی و بهبود کیفیت خدمات در CQA، به یک مسیر تحقیقاتی پرمخاطب مبدل شده است. از جمله‌ی این راهکارها می‌توان به تولید سریع و خودکار پاسخ توسط ماشین و ارائه‌ی

^{۳۱} برای این منظور، به مرورگر داده در Stack Exchange به آدرس <https://data.stackexchange.com/stackoverflow/query/new> مراجعه نموده و دستور `SELECT AVG(CAST (LEN( Body ) AS BIGINT)) FROM Posts` را وارد و اجرا کنید. (ملاقات شده در تاریخ ۱۴۰۰/۰۵/۳۱)

^{۳۲} Satisfactory answer

آن به پرسشگران، به عنوان یک پاسخ کامل، یا پاسخی تکمیلی در کنار پاسخ‌های سایر متخصصین اشاره کرد [23-29]. کیفیت پاسخ تولید شده در این راهکارها، مسئله‌ای است که با توجه به نوپا بودن این حوزه‌ی تحقیقاتی، هنوز با بهترین جایگاه ممکن فاصله دارد. در همین راستا، هدف اصلی پژوهش جاری بر بهبود تولید پاسخ در انجمن‌های پرسش و پاسخ استوار گشته است.

اینک که جایگاه پژوهش در خط زمانی تحولات و پیشرفت بازیابی اطلاعات روشن شد، می‌توان به تشریح سازوکار انجمن‌های پرسش و پاسخ و تبیین مسئله‌ی مورد بررسی در پژوهش با جزئیات بیشتر پرداخت.

فصل ۱- تعریف انجمن‌های پرسش و پاسخ

در شکل **Error! No text of specified style in document.**،^{۳۶} نمایی کلی از فرآیند انجام شده در CQA به تصویر کشیده شده است. همانطور که مشاهده می‌شود، به محض ثبت‌نام در انجمن و قرار گرفتن نام کاربر در فهرست متخصصین، مجوز فعالیت در سه نقش پرسش‌گر^{۳۳}، پاسخ‌دهنده^{۳۴} و رأی دهنده^{۳۵} برای کاربر صادر می‌شود [30, 31]. هر متخصص می‌تواند پرسش خود را به زبان طبیعی در انجمن ارسال کند. این متخصص که پرسش‌گر نامیده می‌شود، قادر است علاوه بر متن پرسش، عنوان^{۳۶} و برچسب‌های^{۳۷} موضوعی را نیز بصورت اختیاری^{۳۸} ثبت نماید. عنوان پرسش، خلاصه‌ای از محتوای پرسش است و برچسب‌ها، کلیدواژه‌هایی هستند که نشان می‌دهند پرسش به چه موضوعاتی مرتبط است. برای پیشگیری از ثبت پرسش تکراری^{۳۹} یا سوءاستفاده‌های^{۴۰} احتمالی، ابتدا محتوای پرسش توسط مدیران انجمن تأیید شده، سپس برای عموم متخصصین، قابل رؤیت می‌گردد [25, 31].

³³ Asker

³⁴ Answerer

³⁵ Voter

³⁶ Title

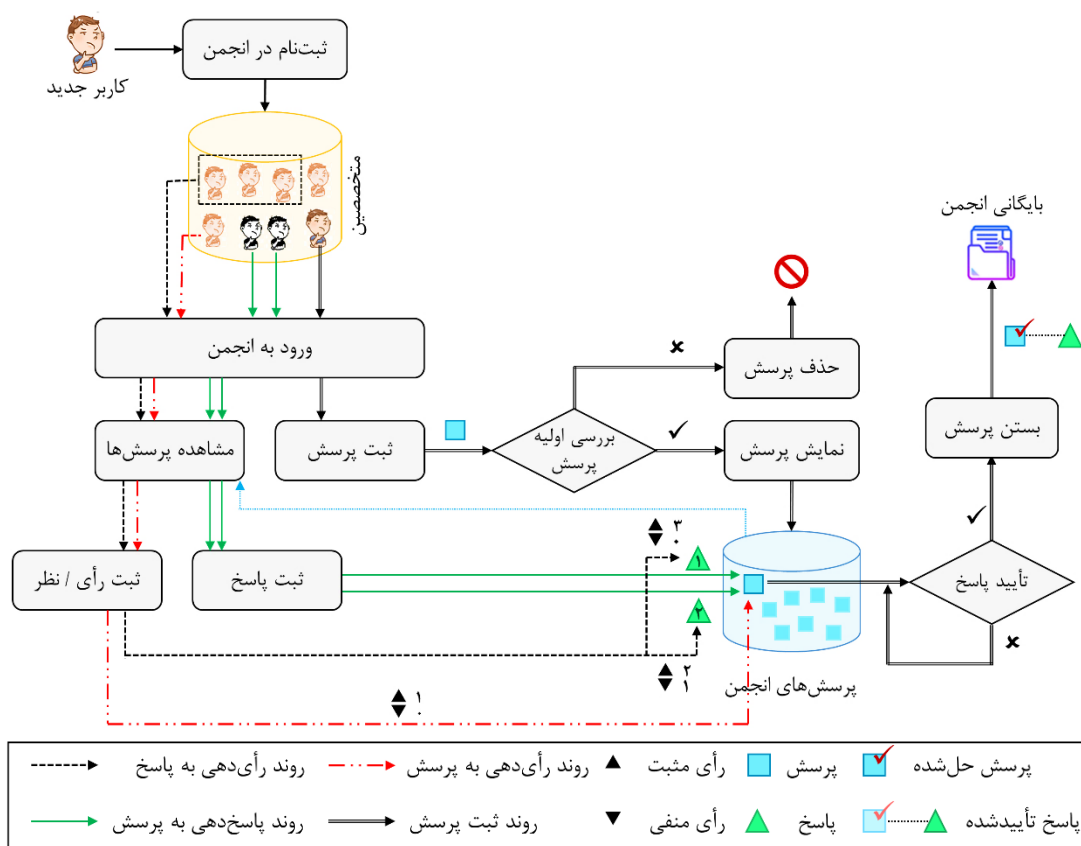
³⁷ Tags

³⁸ Optional

³⁹ Duplicate questions

⁴⁰ Abuse

چنانچه متخصصی این پرسش را مرتبط با حوزه‌ی تخصصی خود بداند، می‌تواند به چهار صورت وارد تعامل شود [19, 32, 33]: (۱) پاسخی برگرفته از دانش، نظرات و تجربیات خود برای پرسش منظور نماید، (۲) با ثبت رأی مثبت^{۴۱} یا منفی^{۴۲} به پاسخ‌های ارائه شده برای پرسش، در قضاوت برای تعیین کیفیت آنها شرکت کند، (۳) با رأی‌دهی به پرسش، کیفیت آن را از منظر واضح، مفید یا به‌روز بودن ارزیابی نماید و (۴) در صورتیکه خواهان توضیحات بیشتر برای پرسش یا پاسخ‌های مبهم^{۴۳} است، به ثبت نظر^{۴۴} مبادرت ورزد.



شکل. Error! No text of specified style in document.-۶. فرآیند ثبت پرسش تا تأیید پاسخ در انجمن پرسش و پاسخ (یک پرسش با ۱ امتیاز مثبت؛ دو پاسخ ثبت شده، اولی با ۲ امتیاز و دومی با ۱ امتیاز)

⁴¹ Upvote

⁴² Downvote

43 Unclear

⁴⁴ Comment

پرسش‌گر، پاسخ‌ها را بصورت دوره‌ای بررسی می‌کند. هر زمان که پاسخی، نیاز اطلاعاتیش را مرتفع سازد، وی می‌تواند آن را به عنوان بهترین پاسخ^{۴۵}، تأیید کند. بلافاصله پس از تأیید، پرسش، با برچسب "حل‌شده"^{۴۶} علامتگذاری و در بایگانی^{۴۷} انجمن ذخیره می‌شود [19, 30, 33, 34]. همچنین ریشه‌ی^{۴۸} پرسش، بسته شده و امکان ثبت پاسخ‌های بیشتر برای آن وجود نخواهد داشت [22]. پاسخ تأیید شده برای پرسش، صرف‌نظر از امتیازی که دارد، بلافاصله پس از پرسش جای گرفته و سایر پاسخ‌ها نیز به ترتیب غیرنزولی امتیازشان نمایش داده می‌شوند. منظور از امتیاز هر پاسخ، تفاضل آراء مثبت و منفی آن پاسخ است [31]. نمونه‌ای از پرسش و پاسخ‌ها در انجمن Stack Overflow در شکل **Error! No text of specified style in document.** نشان داده شده است. در برخی از انجمن‌ها، کاربران نیز براساس نیکویی^{۴۹} پرسش و پاسخ‌های ارسالی‌شان پاداش‌دهی می‌شوند [33]. مجموعه پاداش‌های اعطایی به هر متخصص را اعتبار^{۵۰} می‌نامند. افزایش اعتبار، دامنه‌ی اختیارات بیشتری را برای متخصص رقم می‌زند.

⁴⁵ Best answer

⁴⁶ Resolved

⁴⁷ Archive

⁴⁸ Thread

⁴⁹ Goodness

⁵⁰ Reputation

How to download a pdf file in Android? عنوان پرسش

Asked 10 years, 2 months ago

Active 2 years, 11 months ago

Viewed 87k times

تاریخ ثبت پرسش

آخرین تاریخ فعال بودن پرسش

تعداد دفعات ملاقات پرسش

بدنه پرسش

I want to download a pdf file from an url. For viewing the pdf file I used the code below.

```
File file = new File("/sdcard/example.pdf");
if (file.exists()) {
    Uri path = Uri.fromFile(file);
    Intent intent = new Intent(Intent.ACTION_VIEW);
    intent.setDataAndType(path, "application/pdf");
    intent.setFlags(Intent.FLAG_ACTIVITY_CLEAR_TOP);
    try {
        startActivity(intent);
    }
    catch (ActivityNotFoundException e) {
        Toast.makeText(OpenPdf.this, "No Application Available to View PDF",
            Toast.LENGTH_SHORT).show(); } }
```

It is working but how do I get the pdf file from an url (e.g <http://.../example.pdf>). I want to **download** the pdf file from this url. Please help me. Thanks in advance.

android download برچسب‌های موضوعی

Share Edit Follow

edited Feb 20 '13 at 16:16

Bobrovsky
12.8k ● 19 ● 74 ● 122

asked Jun 16 '11 at 8:00

Ramakrishna
3,978 ● 15 ● 46 ● 71

نشان‌های تخصصی

4 Answers تعداد پاسخ‌های ثبت شده

Active Oldest Votes

Downloading a PDF works the same as downloading any other binary file.

1. Open a `URLConnection`
2. Use the connection's `getInputStream()` method to read the file.
3. Create a `FileOutputStream` and write the inputstream.

Check [this post](#) for example source code.

Share Edit Follow

edited May 23 '17 at 11:33

Community
1 ● 1

answered Jun 16 '11 at 9:34

THelper
14.3k ● 6 ● 60 ● 95

- 2 But in Blundell answer the file will not be downloaded if pdf is not available in mobile. But i want to download pdf file from url and then the pdf will be show in webview. So this is the right answer for me. – Ramakrishna
Jun 16 '11 at 10:54

Any plugin that does that? Sound like something generic – Vlado Pandžić Aug 7 '17 at 21:33

Download a pdf:

```
startActivity(new Intent(Intent.ACTION_VIEW, Uri.parse("www.education.gov.yk.ca/pdf/pdf-te
```

Ah as I've just found out this is device dependent.

scenarios

1. It will download the pdf to your browser/downloaded/ folder
2. You have a google docs account - It will ask you to sign in then view the pdf in the browser
3. You have a pdf reader installed - App dependent may catch it may not

شکل **Error! No text of specified style in document.** ۷- نمونه‌ای از ریشه‌های پرسش و پاسخ در CQA⁵¹. (برای این پرسش

چهار پاسخ ثبت شده است که تنها دو پاسخ آن نمایش داده شده‌اند. پاسخ تأیید شده بلافاصله بعد از پرسش قرار گرفته است.)
برخی از اوقات، پرسش‌گر، هیچ‌یک از پاسخ‌های دریافت شده را تأیید نمی‌کند. این اتفاق، ممکن است به دو علت رخ دهد [22]: (۱) پرسش‌گر با گذشت زمان، علاقه‌ی خود را به دنبال کردن مسئله از دست داده است، (۲)

⁵¹ <https://stackoverflow.com/questions/6368788/how-to-download-a-pdf-file-in-android> (ملاقات شده در تاریخ ۱۴۰۰/۰۵/۲۵)

هیچ یک از پاسخ‌ها دارای کیفیت قابل قبول نبوده‌اند. در برخی از انجمن‌ها، اگر پرسش‌گر هیچ‌گاه پرسش را نبندد، پرسش پس از مدت زمان مشخصی بطور خودکار بسته شده و پاسخی که بیشترین امتیاز یا کیفیت را داراست، به عنوان بهترین پاسخ، تأیید می‌شود [22, 30, 31]. در شکل **Error! No text of specified style in** عنوان بهترین پاسخ، تأیید می‌شود [22, 30, 31]. در شکل **Error! No text of specified style in** عنوان بهترین پاسخ، تأیید می‌شود [22, 30, 31].

document-۸ دو نمونه از پرسش‌های فاقد پاسخ تأیید شده در انجمن Stack Overflow نشان داده شده است.

Download File PDF in Angular from Spring MVC

Asked 7 years, 5 months ago Active 6 years, 1 month ago Viewed 2k times

I have this situation. In AngularJS:

```
var graphHtml = $(".element10")[0];
$http({
  method: 'POST',
  url: '/export/pdf',
  data: "html=" + graphHtml.outerHTML,
  headers: {'Content-Type': 'application/x-www-form-urlencoded'}
}).success(function(data, status, headers, config) {
  var hiddenElement = document.createElement('a');
  hiddenElement.href = 'data:application/pdf;' + encodeURIComponent(data);
  hiddenElement.target = '_blank';
  hiddenElement.download = 'myFile.pdf';
  hiddenElement.click();
}).error(function(data, status, headers, config) { });
```

[...]

Download / Open PDF with Javascript?

Asked 4 years, 7 months ago Active 4 years, 7 months ago Viewed 766 times

I am facing the problem that I want to have two buttons, you can open or download a PDF on a server like:

```
<button id="download">download pdf</button>
<button id="open">open pdf</button>
```

But I don't have any idea how to achieve this. I already found that:

```
<a href="x.pdf" download="x.pdf">PDF Download</a>
```

BUT: HTML5 download attribute is not working in all browsers and it needs to be cross browser save.

javascript pdf download

Share Edit Follow

asked Jan 11 '17 at 19:50



Inkperial
63 ● 7

2 Answers

You should take a look at this:

<https://github.com/eligrey/FileSaver.js/>

JavaScript file which allows for cross-browser download support

Share Edit Follow

[...]

answered Jan 11 '17 at 19:59



Pim
680 ● 6 ● 17

شکل **Error! No text of specified style in document**-۸ بخشی از ریشه‌ی دو پرسش در CQA که توسط شخص پرسش‌گر بسته نشده‌اند (بالا)^{۵۲}: فاقد هرگونه پاسخ یا نظر، پایین^{۵۳}: دارای دو پاسخ که هیچکدام توسط پرسش‌گر تأیید نشده‌اند)

⁵² <https://stackoverflow.com/questions/22306580/download-file-pdf-in-angular-from-spring-mvc> (ملاقات شده در ۱۴۰۰/۰۵/۲۶)

⁵³ <https://stackoverflow.com/questions/41599520/download-open-pdf-with-javascript> (ملاقات شده در ۱۴۰۰/۰۵/۲۶)

فصل ۲- بیان مسئله

بسته نشدن پرسش توسط شخص پرسش‌گر، به خصوص به دلایلی چون عدم دریافت پاسخ یا کیفیت نامناسب پاسخ‌های دریافت شده، همواره زنگ خطری برای انجمن پرسش و پاسخ محسوب می‌شود. این اتفاق، معمولاً بر نارضاضی بودن پرسش‌گر دلالت دارد و این درحالیست که بقا و رشد هر انجمن پرسش و پاسخ، به شدت در گرو رضایت متخصصین آن است [35]. بنابراین، اگر تعداد زیادی از متخصصین از خدمات انجمن نارضاضی باشند، باید منتظر مهاجرت گسترده‌ی آنها به انجمن‌های رقیب یا ابزارهای جایگزین بود [27, 34, 36]. در این راستا، شناخت عوامل مؤثر در کسب رضایت کاربر و ارائه‌ی راهکاری برای کاهش زمان انتظار وی، اهمیت بسیار می‌یابد. یکی از راهکارهای موجود، ارائه‌ی یک پاسخ به کاربر توسط سیستم، بلافاصله پس از ثبت پرسش است [27, 28, 30, 37]. در اینصورت، پرسش‌گر، پیش از آنکه به انتظار دریافت پاسخ از سایر متخصصین انجمن بنشیند، پاسخی را از سیستم دریافت می‌کند. اگر پاسخ صحیح باشد و نیاز اطلاعاتی وی را مرتفع سازد، آن را تأیید می‌کند، در غیر اینصورت می‌تواند منتظر دریافت پاسخ از سایرین بماند.

بر این اساس، مسئله اصلی در پژوهش جاری، بهبود تولید خودکار پاسخ برای پرسش‌ها در یک انجمن پرسش و پاسخ، با کمک دانش کاربر و دانش نهفته در محتوا است. این سیستم باید قادر باشد، به محض ثبت یک پرسش جدید در انجمن، مفهوم پرسش را ضمن توجه به سوابق پرسش‌گر، درک نموده، سپس با تکیه بر منابع اطلاعاتی در دسترس، خواه بایگانی انجمن باشد خواه یک منبع دانش خارجی ساخت‌یافته^{۵۴} یا غیرساخت‌یافته، پاسخی تا حد امکان صحیح، معتبر و مطابق با انتظارات کاربر، به زبان طبیعی به او ارائه نماید. بر این اساس، پرسش‌های زیر باید در پژوهش پاسخ داده شوند:

(۱) چگونه می‌توان مفهوم پرسش و منظور پرسش‌گر را با دقت بالا درک نمود؟

(۲) چگونه می‌توان تخصص و دانش پیش‌زمینه‌ی کاربر را مدل نمود؟

۳) از چه منابعی برای استخراج پاسخ استفاده شود تا هم کیفیت و هم کمیت پاسخگویی تضمین شود؟

۴) چگونه می‌توان از پرسش‌های موجود در بایگانی انجمن که قبلاً پاسخ گرفته‌اند برای پاسخگویی به پرسش

ورودی بهره‌مند شد؟

۵) چگونه می‌توان نتایج حاصل از چند منبع اطلاعاتی را در قالب یک پاسخ زبان طبیعی تلفیق و به پرسش‌گر

ارائه نمود؟

۶) چگونه می‌توان کیفیت پاسخ‌های تولید شده را ارزیابی کرد؟

فصل ۳- ضرورت و اهمیت انجام پژوهش

سرعت و صحت پاسخگویی، دو رکن اساسی در جلب رضایت متخصصین انجمن محسوب می‌شوند؛ چرا که عدم حصول پاسخ صحیح یا دستیابی دیر هنگام به پاسخی با کیفیت، ممکن است پیامدهای سنگینی را برای کاربران به همراه داشته باشد. از آن جمله می‌توان به انجمن‌هایی چون سلامت و پزشکی [38-40] و تجارت الکترونیک [41, 42] که در زندگی روزمره افراد حضور مؤثر دارند، اشاره کرد. در این راستا، درک محتوای پرسش، بهره‌گیری از منابع غنی دانش برای بهبود کیفیت پاسخ تولید شده و همزمان، توجه به دانش پیش‌زمینه و تخصص کاربر برای نزدیک شدن به پاسخ مطلوب وی، امری ضروری است. به این ترتیب، اهمیت پژوهش جاری به شرح زیر است:

۱) محدود نمودن دامنه‌ی جستجو متناسب با دانش پیش‌زمینه و علایق پرسش‌گر

۲) کاهش زمان دستیابی به پاسخ صحیح و رضایت‌بخش برای پرسش

۳) افزایش صحت و جامعیت پاسخ و کاهش ابهام در متن آن با استفاده‌ی همزمان از دانش و الگوهای انسانی

نهفته در پاسخ‌های متخصصین و دانش موجود در منابع دانش خارجی

۴) کمک به جمع‌آوری بهترین پاسخ‌ها از دیدگاه‌ها و جنبه‌های مختلف و تجمیع آنها در قالب یک پاسخ زبان

طبیعی

(۵) پیشنهاد دادن فهرستی از پرسش‌های متضمن^{۵۵} برای مطالعات بیشتر

(۶) ارتقاء کیفیت خدمات انجمن و نائل آمدن به جایگاه‌های بالاتر از نظر اعتماد و رضایت کاربران

(۷) کمک به ارتقاء سطح دانش متخصصین در انجمن، با مطالعه‌ی پاسخ‌های خودکار ارائه شده توسط سیستم

به عنوان یک منبع یادگیری

(۸) ترغیب متخصصین به مشارکت بیشتر در انجمن از طریق ارائه‌ی پاسخ‌های خود به پرسش برای رقابت با

پاسخ تولید شده توسط سیستم

فصل ۴- محدودیت‌های پژوهش

پژوهش جاری، در ارزیابی سیستم پیشنهادی، با سه گروه از محدودیت‌ها به شرح زیر مواجه است:

(۱) دسترسی به دادگان و منابع دانش

در زمینه‌ی دسترسی به دادگان، سه محدودیت عمده وجود دارد. این محدودیت‌ها عبارتند از:

- یافتن دادگان فارسی: تعدد انجمن‌های پرسش و پاسخ در زبان انگلیسی و انتشار دوره‌ای بایگانی

آنها، مسیر تحقیقات در این زبان را بسیار هموار نموده است. با این حال، بسیاری از کاربران

فارسی زبان، ممکن است به دلیل عدم تسلط به زبان انگلیسی، موفق به یافتن پاسخ پرسش‌های

خود در این انجمن‌ها نشوند. این پژوهش قصد دارد، سیستم پیشنهادی را در زبان فارسی نیز

اعمال و نتایج را بررسی نماید. در این مسیر ممکن است، با محدودیت دسترسی به یک دادگان

فارسی جامع و استاندارد، برای آموزش و ارزیابی سیستم و مؤلفه‌های آن مواجه گردد.

- انتخاب گراف دانش^{۵۶}: گراف دانش باید با در نظر داشتن سه فاکتور انتخاب شود. اول، دسترسی

پذیری، دوم، تطابق دامنه‌ی موضوعی پرسش‌های انجمن با موضوعات پوشش داده شده در گراف

⁵⁵ Entailed questions

⁵⁶ Knowledge Graph (KG)

دانش و سوم، مدت زمان لازم برای پاسخگویی^{۵۷} به درخواست‌ها در گراف دانش. متأسفانه دسترسی به کل محتوای گراف دانش Freebase به دلیل برخی محدودیت‌ها، در کشور ایران امکان پذیر نیست^{۵۸}. از این رو، کلیدی ارزیابی‌ها تنها روی گراف دانش DBPedia انجام می‌گردد^{۵۹}.

- پیچیدگی دادگان موجود: داده‌های منتشر شده توسط انجمن Stack Overflow، بسیار بزرگ‌مقیاس و حجیم است^{۶۰}. از یک سو، پردازش این حجم از اطلاعات، زمان زیادی را می‌طلبد. از سوی دیگر، انتخاب تنها بخشی از این دادگان، به دلیل وجود وابستگی بین اجزای مختلف داده، به آسانی میسر نیست. از این رو، برخلاف آزمایشات انجام شده در مقاله‌ی پایه، ناگزیر به انجام آزمایشات روی داده‌های محدودتر هستیم.

۲) محدودیت‌های سخت‌افزاری

- استفاده از مدل‌های یادگیری و گراف دانش می‌تواند منجر به محدودیت‌هایی به شرح زیر گردد:
- آموزش مدل‌های یادگیری عمیق^{۶۱}: پژوهش جاری بر استفاده از روش‌های نوین و مدل‌های یادگیری عمیق متمرکز است. این مدل‌ها حجم زیادی از داده‌ها و پردازش‌ها را می‌طلبند که لازمه‌ی انجام آن، دسترسی طولانی‌مدت به سخت‌افزاری قدرتمند برای پردازش است. فراهم نمودن این امکانات در شرایط آزمایشگاهی، با محدودیت‌هایی روبرو است.
 - استخراج دانش از گراف: جستجوی اطلاعات در میان حجم بسیار زیاد گراف دانش DBPedia، به سیستمی با پیکربندی قوی نیازمند است. این مسئله، لزوم استفاده از نقطه‌ی پایانی^{۶۲} تعریف

⁵⁷ Response time

⁵⁸ <https://developers.google.com/freebase>

⁵⁹ <https://www.dbpedia.org/>

⁶⁰ طبق دادگان منتشر شده در <https://archive.org/download/stackexchange>، حدود ۵۲ گیگابایت حجم دارد.

⁶¹ Deep learning

⁶² Endpoint

شده جهت ارسال پرسمان‌های SPARQL در گراف دانش DBPedia را روشن می‌سازد. این درحالیست که در این نقطه‌ی پایانی نیز، محدودیت‌هایی از نظر نرخ درخواست‌ها برای هر IP در ثانیه، زمان اجرای هر پرسمان و تعداد رکوردهای نتیجه، تعریف شده است که این امر، ارزیابی‌های پژوهش را تحت الشعاع قرار می‌دهد.

۳) چالش‌های زبان‌شناختی

- همانطور که در بخش پیشگفتار گفته شد، پرسش‌های ثبت شده در انجمن‌ها، اغلب طولانی و چندجمله‌ای هستند. این مسئله می‌توان درک محتوای پرسش و تبدیل آن به پرسمان‌های ساخت‌یافته را با چالش‌هایی مواجه کند. از سوی دیگر، عموماً کاربرانی که پرسش‌های طولانی را ارسال می‌کنند، علاقمند به دریافت پاسخ‌های کامل و تشریحی هستند [29]. ارائه‌ی یک چکیده و خلاصه از پاسخ‌های یافت شده، نیازمند یک خلاصه‌ساز قوی است. چنین خلاصه‌سازی با کیفیت قابل قبول، در همه زبان‌ها ممکن است در دسترس نباشد.

فصل ۵- نوآوری‌های پژوهش

بطور کلی، نوآوری‌های ارائه شده در پژوهش جاری به شرح زیر است:

- (۱) پیشنهاد سیستمی^{۶۳} برای پاسخگویی خودکار به پرسش‌های انجمن مبتنی بر مدلسازی دانش پیش‌زمینه‌ی کاربر و دانش موجود در دو منبع غنی اطلاعاتی شامل بایگانی انجمن و گراف دانش عمومی، با این هدف که دو منبع اطلاعاتی، نقطه ضعف‌های یکدیگر را پوشش داده، ارائه‌ی حتمی پاسخ به کاربر را تضمین نموده و کیفیت پاسخ‌های ارائه شده را ارتقاء بخشند

(۲) استفاده از الگوریتم‌های یادگیری عمیق برای مدلسازی دانش پیش‌زمینه‌ی پرسش‌گر براساس سوابق فعالیت‌های وی در انجمن، به منظور درک بهتر خواسته‌ها، علایق و انتظارات پرسش‌گر فراتر از واژگان و ساختار پرسشی که ثبت نموده است

(۳) دخالت دادن دانش پیش‌زمینه‌ی پرسش‌گر در فرآیند شناسایی پرسش‌هایی مشابه با پرسش ورودی، که پیش‌تر در انجمن ثبت شده و پاسخ مناسبی دریافت نموده‌اند

(۴) دخالت دادن دانش پیش‌زمینه‌ی پرسش‌گر در فرآیند ساخت و رتبه‌بندی پرسمان‌های SPARQL جهت استخراج مرتبط‌ترین حقایق از گراف دانش

(۵) تولید پاسخ به زبان طبیعی براساس برترین پاسخ‌های تأیید شده برای پرسش‌های متضمن و حقایق استخراج شده از گراف دانش، ضمن توجه به امتیازات هر کدام

(۶) رفع کاستی‌های دادگان موجود و غنی‌سازی آن‌ها، به منظور هماهنگی بیشتر با وظیفه‌ی مورد مطالعه

(۷) گردآوری و آماده‌سازی دادگان زبان فارسی برای آموزش و تست مؤلفه‌های سیستم

مراجع

- [1] B. Ojokoh and E. Adebisi, "A Review of Question Answering Systems," *Journal of Web Engineering*, vol. 17, no. 8, pp. 717–758, 2019.
- [2] V. Guda and S. K. Sanampudi, "Approches for Question Answering Systems," *International Journal of Engineering science and technology*, vol. 3, no. 2, pp. 990–995, 2011.
- [3] A. Mishra and S. K. Jain, "A survey on question answering systems with classification," *Journal of King Saud University - Computer and Information Sciences*, vol. 28, no. 3, pp. 345–361, Jul. 2016.
- [4] R. Soricut and E. Brill, "Automatic question answering using the web: Beyond the Factoid," *Information Retrieval*, vol. 9, no. 2, pp. 191–206, Mar. 2006.
- [5] A. D. Vanberg, "From Archie to Google - Search engine providers and emergent challenges in relation to EU competition law," *European Journal of Law and Technology*, vol. 3, no. 1, pp. 1–18, 2012.
- [6] T. N. Manjunath, D. Yogish, S. Mahalakshmi, and H. K. Yogish, "Smart question answering system using vectorization approach and statistical scoring method," *Materials Today: Proceedings*, Jul. 2021.
- [7] W. Zheng, H. Cheng, J. X. Yu, L. Zou, and K. Zhao, "Interactive natural language question answering over knowledge graphs," *Information Sciences*, vol. 481, pp. 141–159, May 2019.
- [8] A. Bouziane, D. Bouchiha, N. Doumi, and M. Malki, "Question Answering Systems: Survey and Trends," *Procedia Computer Science*, vol. 73, pp. 366–375, 2015.
- [9] A. M. N. Allam and M. H. Haggag, "The Question Answering Systems: A Survey.," *International Journal of Research and Reviews in Information Science*, vol. 2, no. 3, 2012.
- [10] T. Sultana and S. Badugu, "A Review on Different Question Answering System Approaches," in *Advances in Decision Sciences, Image Processing, Security and Computer Vision*, 2020, pp. 579–586.
- [11] T.-S. Chua, R. Hong, G. Li, and J. Tang, "From text question-answering to multimedia QA on web-scale media resources," in *Proceedings of the First ACM workshop on Large-scale multimedia retrieval and mining - LS-MMRM '09*, 2009, p. 51.
- [12] L. A. Zadeh, "Chapter 9 From search engines to question answering systems — The problems of world knowledge, relevance, deduction and precisiation," in *Capturing Intelligence*, vol. 1, no. C, 2006, pp. 163–210.
- [13] E. M. Voorhees, "The TREC-8 question answering track report," in *Proceedings of Text REtrieval Conference*, vol. 99, pp. 77–82, 1999.
- [14] H. T. Dang, D. Kelly, and J. J. Lin, "Overview of the TREC 2006 question answering track," 2007.

- [15] X. Wang, Z. Li, and J. Tang, "Multimedia news QA: Extraction and visualization integration with multiple-source information," *Image and Vision Computing*, vol. 60, pp. 162–170, Apr. 2017.
- [16] R. Hong, M. Wang, G. Li, L. Nie, Z.-J. Zha, and T.-S. Chua, "Multimedia Question Answering," *IEEE MultiMedia*, vol. 19, no. 4, pp. 72–78, Oct. 2012.
- [17] I. Srba and M. Bielikova, "A Comprehensive Survey and Classification of Approaches for Community Question Answering," *ACM Transactions on the Web*, vol. 10, no. 3, pp. 1–63, Aug. 2016.
- [18] Z. Zheng, Y. Tang, C. Long, F. Bu, and X. Zhu, "Question answering system based on community qa," in *The Workshop Programme*, May 2017, vol. 23, no. May, pp. 23–27.
- [19] B. Patra, "A survey of Community Question Answering," *CoRR*, abs/1705.04009, May 2017.
- [20] M. S. Zahedi, M. Rahgozar, and R. A. Zoroofi, "HCA: Hierarchical Compare Aggregate model for question retrieval in community question answering," *Information Processing & Management*, vol. 57, no. 6, p. 102318, Nov. 2020.
- [21] H. Gao, L. Wu, P. Hu, and F. Xu, "RDF-to-Text Generation with Graph-augmented Structural Neural Encoders," in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, Jul. 2020, pp. 3030–3036.
- [22] Y. Liu, J. Bian, and E. Agichtein, "Predicting information seeker satisfaction in community question answering," in *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '08*, 2008, pp. 483–490.
- [23] M. Li, Y. Li, Q. Peng, J. Wang, and C. Yu, "Evaluating community question-answering websites using interval-valued intuitionistic fuzzy DANP and TODIM methods," *Applied Soft Computing*, vol. 99, p. 106918, Feb. 2021.
- [24] M. Li, Y. Li, Q. Peng, and J. Wang, "A hybrid MCDM model combining DANP with TODIM to evaluate the information quality of community question answering in a two-dimensional linguistic environment," *Expert Systems*, vol. 38, no. 2, Mar. 2021.
- [25] P. K. Roy, "Deep Neural Network to Predict Answer Votes on Community Question Answering Sites," *Neural Processing Letters*, vol. 53, no. 2, pp. 1633–1646, Apr. 2021.
- [26] D. Kundu, R. K. Pal, and D. P. Mandal, "Time-aware hybrid expertise retrieval system in community question answering services," *Applied Intelligence*, Feb. 2021.
- [27] Y. Wu and S. Zhao, "Community answer generation based on knowledge graph," *Information Sciences*, vol. 545, pp. 132–152, Feb. 2021.
- [28] Y. Deng *et al.*, "Joint Learning of Answer Selection and Answer Summary Generation in Community Question Answering," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, pp. 7651–7658, Apr. 2020.
- [29] H. Song, Z. Ren, S. Liang, P. Li, J. Ma, and M. De Rijke, "Summarizing answers in non-factoid community question-answering," *WSDM 2017 - Proceedings of the 10th ACM International Conference on Web Search and Data Mining*, no. 3, pp. 405–414, 2017.
- [30] I. Srba and M. Bielikova, "A comprehensive survey and classification of approaches for Community Question Answering," *ACM Transactions on the Web*, vol. 10, no. 3, pp. 1–63, Aug. 2016.
- [31] M. J. Blooma and J. C. Kurian, "Research issues in Community based Question Answering," *PACIS 2011 - 15th Pacific Asia Conference on Information Systems: Quality Research in Pacific*, pp. 1–12, 2011.

- [32] P. Nakov *et al.*, “SemEval-2017 Task 3: Community Question Answering,” in *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, Dec. 2017, pp. 27–48.
- [33] by Doris Hoogeveen and T. M. Baldwin Karin Verspoor, “Real and Misflagged Duplicate Question Detection in Community Question-Answering,” 2018.
- [34] A. Figueroa, “Automatically generating effective search queries directly from community question-answering questions for finding related questions,” *Expert Systems with Applications*, vol. 77, pp. 11–19, Jul. 2017.
- [35] J. Lou, Y. Fang, K. H. Lim, and J. Z. Peng, “Contributing high quantity and quality knowledge to online Q&A communities,” *Journal of the American Society for Information Science and Technology*, vol. 64, no. 2, pp. 356–371, Feb. 2013.
- [36] C. Shah and J. Pomerantz, “Evaluating and Predicting Answer Quality in Community QA Categories and Subject Descriptors General Terms,” *Sigir*, no. March 2008, 2010.
- [37] S. Hu, L. Zou, J. X. Yu, H. Wang, and D. Zhao, “Answering Natural Language Questions by Subgraph Matching over Knowledge Graphs,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 5, pp. 824–837, May 2018.
- [38] H. Samuel, M.-Y. Kim, S. Prabhakar, M. S. M. Jabbar, and O. Zalane, “Community question retrieval in health forums,” in *2017 IEEE EMBS International Conference on Biomedical & Health Informatics*, 2017, pp. 77–80.
- [39] P. Wongchaisuwat, D. Klabjan, and S. R. Jonnalagadda, “A Semi-Supervised Learning Approach to Enhance Health Care Community-Based Question Answering: A Case Study in Alcoholism,” *JMIR Medical Informatics*, vol. 4, no. 3, p. e24, Aug. 2016.
- [40] A. Ben Abacha, C. Shivade, and D. Demner-Fushman, “Overview of the MEDIQA 2019 Shared Task on Textual Inference, Question Entailment and Question Answering,” in *Proceedings of the 18th BioNLP Workshop and Shared Task*, 2019, pp. 370–379.
- [41] H. Mittal, A. Chakrabarti, B. Bayar, A. A. Sharma, and N. Rasiwasia, “Distantly Supervised Transformers For E-Commerce Product QA,” in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2021, pp. 4008–4017.
- [42] S. Chen, C. Li, F. Ji, W. Zhou, and H. Chen, “Review-Driven Answer Generation for Product-Related Questions in E-Commerce,” in *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, Jan. 2019, pp. 411–419.
- [43] D. Hoogeveen, L. Wang, T. Baldwin, and K. M. Verspoor, “Web Forum Retrieval and Text Analytics: A Survey,” *Foundations and Trends in Information Retrieval*, vol. 12, no. 1, pp. 1–163, 2018.
- [44] H. Rithish, G. Deepak, and A. Santhanavijayan, “Automated Assessment of Question Quality on Online Community Forums,” in *International Conference on Digital Technologies and Applications*, Springer, Cham, 2021, pp. 791–800.
- [45] S. K. Maity, A. Kharb, and A. Mukherjee, “Analyzing the Linguistic Structure of Question Texts to Characterize Answerability in Quora,” *IEEE Transactions on Computational Social Systems*, vol. 5, no. 3, pp. 816–828, Sep. 2018.
- [46] J. Trienes and K. Balog, “Identifying Unclear Questions in Community Question Answering Websites,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11437 LNCS, 2019, pp. 276–289.
- [47] K. Balog, A. Schuth, P. Dekker, and N. Tavakolpoursaleh, “Overview of the TREC 2016

Open Search track,” *Proceedings of the Twenty-Fifth Text REtrieval Conference (TREC 2016)*, pp. 1–9, 2016.