



وزارت ارتباطات فناوری اطلاعات

طرح جویشگر بومی جستجوی وب با استفاده از گفتار

۲۵ اردیبهشت ماه ۹۵



وزارت ارتباطات فناوری اطلاعات

سرفصل مطالب

پردازش گفتار	۱
بکارگیری گفتار برای جستجو	۲
جستجوی گفتاری در طرح جویشگر	۳
RFP	۴
گامهای آتی	۵



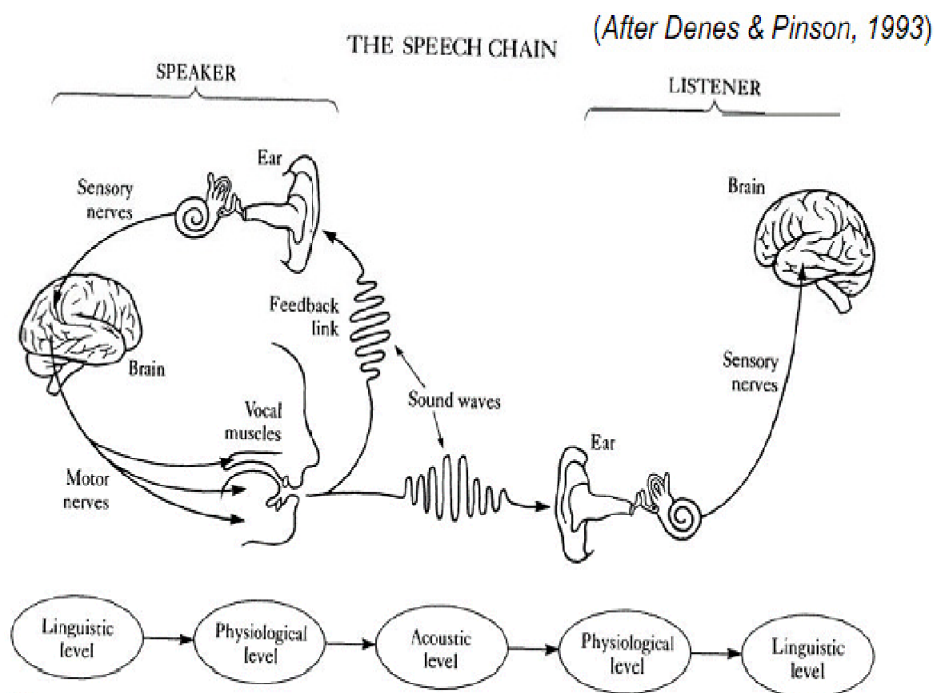
وزارت ارتباطات فناوری اطلاعات

پردازش گفتار

- پردازش سیگنال گفتار
- پردازش سیگنالها در محیط دیجیتال
- پردازش گفتار بطور کلی شامل دو بخش زیر می باشد:
 - بازشناسی گفتار
 - سنتز گفتار



وزارت ارتباطات و فناوری اطلاعات



پردازش گفتار – بازشناسی گفتار

- تبدیل گفتار انسان به متن
- ورودیها شامل فایل صوتی، میکروفون، گفتار تلفنی و ...
- اولین تلاش در دهه چهل میلادی
- موفق ترین رویکرد تشخیص الگو
 - مدل شدن گفتار با استفاده از واحدهای آوایی
 - کلمه
 - هجا
 - سه واجی
- مدل آوایی متناسب با نظریه نوام چامسکی در مورد پردازش گفتار در مغز



وزارت ارتباطات و فناوری اطلاعات



پردازش گفتار – بازشناسی گفتار

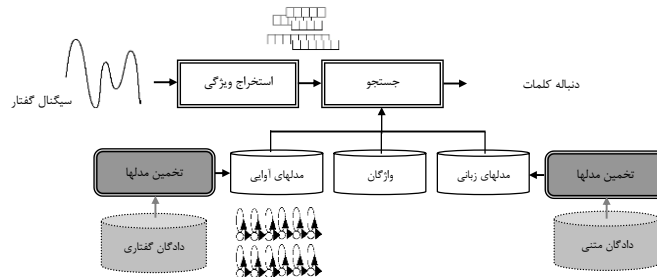
- پیچیدگیهای مطرح در بازشناسی گفتار
 - استقلال و یا وابستگی به گوینده
 - پیوسته و یا گسسته بودن گفتار
 - اندازه واژگان
 - محدودیتهای زبانی
 - لهجه و گویش
 - کارایی در حضور نویز و در محیطهای کاربردی مختلف
 - ابهام آکوستیکی



وزارت ارتباطات و فناوری اطلاعات



پردازش گفتار – بازشناسی گفتار



وزارت ارتباطات و فناوری اطلاعات

پردازش گفتار – بازشناسی گفتار

- استخراج ویژگی
 - هدف آن کاهش حجم محاسبات و حذف افزونگی‌های موجود در سیگنال گفتار
 - روشهای مبتنی بر سیستم شنوایی در مقابل نویز مقاوم‌تر هستند
- مدل‌سازی آوایی
 - مدل کردن گفتار در بازشناسی گفتار
- مدل‌سازی زبانی
 - بکارگیری آن در تبدیل متن به گفتار، OCR، و ...
 - استفاده برای ریخت‌شناسی (Morphological)، ساختاری (Syntax)، معنایی (Semantic)



وزارت ارتباطات و فناوری اطلاعات

پردازش گفتار – بازشناسی گفتار

- دادگان گفتاری و متنی
 - تهیه مدل آوایی نیازمند دادگان متنی است
 - فرآیند آموزش مدل‌های آکوستیکی بایستی به کمک دادگان‌های گفتاری از قبل تهیه شده صورت بگیرد
 - دادگان‌ها شامل حجم قابل توجهی از نمونه‌های آوایی (مثلا ۲۰۰ ساعت گفتار) همراه با تنوع کافی در تعداد گویندگان
 - دادگان متنی نیز دربرگیرنده حجم وسیعی از کلمات (مثلا حدود ۲۰۰ میلیون کلمه) هستند که به کمک آنها اطلاعات آماری شامل احتمال رخداد کلمات پشت سرهم و اطلاعات معنایی استخراج می‌گردد
 - جستجوی واحدهای آوایی و کلمات
 - بعد از آموزش مدل‌ها، هنگام بازشناسی بایستی برای دنباله آوایی بهترین دنباله کلمات پیدا شود
- وزارت ارتباطات و فناوری اطلاعات

پردازش گفتار – بازشناسی گفتار

- معیارهای ارزیابی سیستمهای بازشناسی گفتار
 - میزان دقت
 - میزان تصحیح اشتباه
 - نرخ خطای کلمات
 - سرگستگی
- وزارت ارتباطات و فناوری اطلاعات

پردازش گفتار – تبدیل متن به گفتار

- نرم‌افزاری که متن را در ورودی دریافت و آنرا به گفتار تبدیل می‌کند
- نیازمند الگوریتم‌های پردازش زبان طبیعی و هوش مصنوعی و پردازش گفتار است
- عوامل مهم در کیفیت یک نرم‌افزار تبدیل متن به گفتار
 - بیان و تلفظ صحیح کلمات و عبارات و حتی کلمات غیر متداول
 - قابلیت فهم بالای عبارات بیان شده
 - خوشایند بودن عبارات بیان شده برای شنونده که به عوامل زیادی مانند کیفیت طیفی گفتار سنتز شده، یکنواختی گفتار سنتز شده و میزان شباهت آن با گفتار انسان، پروژودی و ... بستگی دارد.
 - قابلیت هماهنگی با زبان انگلیسی و عدم کاهش کیفیت در هنگام تغییر زبان
 - قابلیت هماهنگ بودن با ابزارهای صفحه خوان



وزارت ارتباطات و فناوری اطلاعات



پردازش گفتار – تبدیل متن به گفتار

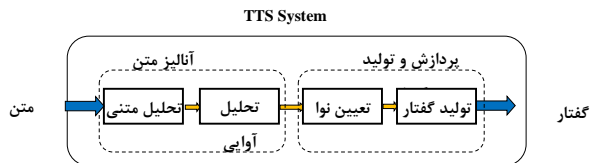
- عوامل مهم در کیفیت یک نرم‌افزار تبدیل متن به گفتار
 - خوشایند گسترده‌تری استفاده نرم افزار برای کاربردهای مختلف مانند متون اقتصادی، سیاسی، اجتماعی، اخبار، رمان، ریاضی و ...
 - سرعت مناسب پردازش و پخش جملات به نحوی که جملات بدون وقفه قابل پخش باشند.
 - عدم نیاز به پردازش بالا به نحوی که کاربر بتواند به راحتی به همراه این نرم افزار با بقیه نرم‌افزارهای موجود در رایانه خود کار کند.
 - طراحی ماژولار به نحوی که امکان بالا بردن کیفیت به مرور زمان و با دریافت فیدبک‌های کاربران وجود داشته باشد.



وزارت ارتباطات و فناوری اطلاعات



پردازش گفتار – تبدیل متن به گفتار



وزارت ارتباطات فناوری اطلاعات



پردازش گفتار – تبدیل متن به گفتار

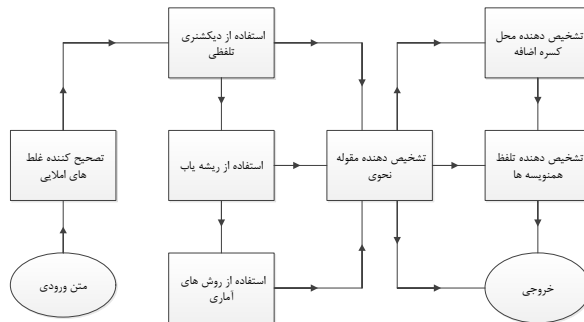
- مشکلات استخراج اطلاعات آوایی از متن ورودی
 - هیچ نگاشت صریحی بین صورت نوشتاری کلمات و صورت آوایی کلمات وجود ندارد.
 - هر کلمه در زبان فارسی ممکن است به صورت‌های مختلف خوانده شود. به این کلمه‌ها هم‌نویسه می‌گویند. این مشکل در زبان فارسی شدیدتر از زبان انگلیسی است.
 - عدم نگارش کسره اضافه موجب افزایش تعداد هم‌نویسه‌ها می‌شود.
 - مشکلات نوشتاری در زبان فارسی بسیار رایج است. به عنوان مثال هر سه کلمه‌ی ای‌صورتی‌که/، ای‌صورتی‌که/، و ایه‌صورتی‌که/ رایج هستند و دارای یک مفهوم و کاربرد می‌باشند.



وزارت ارتباطات فناوری اطلاعات



پردازش گفتار – تبدیل متن به گفتار



وزارت ارتباطات و فناوری اطلاعات



بکارگیری گفتار برای جستجو در وب

- در سال ۱۹۹۷ شرکت گوگل سرویسهای صوتی را آغاز نمود
- در سال ۲۰۰۸ سرویس Autocomplete توسط گوگل ارائه گردید.
- در نوامبر سال ۲۰۰۸ شرکت گوگل کاربرد موبایل خود را برای پذیرش پیامهای صوتی توسعه داد.
- در فوریه سال ۲۰۰۹ گوگل این امکان را به کاربران خود را داد تا تنها با یک لمس صفحه جستجو در گوگل را انجام دهند.
- در سپتامبر سال ۲۰۱۰ شرکت گوگل سرویس Instant را راهاندازی نمود که امکان ارائه نتایج جستجو حتی در هنگام تایپ query را امکان پذیر می‌ساخت.
- در ژوئن سال ۲۰۱۱ به گوگل کروم یک میکروفون اضافه گردید که به کاربر این امکان را می‌داد تا جستجو در گوگل را در دستکتاب با صدای خود انجام دهد.
- در مه سال ۲۰۱۲ گوگل گراف دانش را معرفی نمود که این سرویس یک بانک داده وسیع برای فهم نحوه ارتباط، مکان و وسیله ارتباط کاربر با اینترنت بود. این سرویس این امکان را به گوگل داد تا با استفاده از آن پاسخ به کاربران را تسریع نماید.
- در ژوئن سال ۲۰۱۲ دستیار دیجیتال گوگل این امکان را فراهم نمود تا بدون اینکه کاربران اطلاعاتی را خواسته باشند اطلاعات مورد دلخواه آنها مانند پیش بینی آب و هوا، وضعیت تیم مورد علاقه، ترافیک و دیگر موارد را برای کاربر ارائه نماید.
- در آگوست سال ۲۰۱۳ الگوریتمهای جدیدی را ارائه نمود که بواسطه آن گوگل می‌تواند فهم خود را از مفهوم و معنی کلمات بهبود بخشد.



وزارت ارتباطات و فناوری اطلاعات



بکارگیری گفتار برای جستجو دروب

- با توسعه سرویسهای گوگل و ارائه آنها امکان ترکیب این سرویها با دیگر سرویسها مانند جستجوی صوتی بوجود آمد.
- به عنوان مثال ترکیب گراف دانش با جستجوی صوتی این امکان را فراهم نمود تا با استفاده آن به مکالمه با موتور جستجو بپردازید.
- مشابه فهم جستجو توسط موتور جستجو گراف دانش این امکان را فراهم می کند تا با استفاده از آن چیزهایی مانند کتاب، محل مورد جستجو و غیره را به کاربر معرفی نماید.
- ترکیب سرویسهای مختلف با جستجوی صوتی گوگل این امکان را می دهد تا به کمک آن امکانات بهتری در اختیار کاربران قرار دهد



وزارت ارتباطات فناوری اطلاعات



بکارگیری گفتار برای جستجو دروب

- در حال حاضر علاوه بر سرویس گوگل، iPhone siri و Microsoft Cortona نیز به ارائه خدمات صوتی می پردازند.
- Siri ابتدا در سال ۲۰۱۱ توسط اپل برای ارائه بر روی iPhone 4s عرضه گشت
- از آن زمان دارندگان موبایلهای هوشمند می توانستند از یک دستیار مجازی سوالهای خود را بپرسند.
- در iOS 8 اپل قابلیت Hey, Siri را اضافه نمود.
- با استفاده از این قابلیت کاربر تنها با گفتن Hey, Siri می تواند از این دستیار مجازی سوالهای خود را بپرسد و در این حالت نیازی به تماس با صفحه گوشی نمی باشد.
- شرکت مایکروسافت نیز پس از این شرکتها اقدام به ارائه دستیار هوشمند خود نمود.
- این دستیار هوشمند پس از سری بازیهای Halo عرضه گشت و در گوشیهای هوشمند با سیستم عامل Windows 8.1 و رایانهها با سیستم عامل Windows 10 می توانستند از این دستیار صوتی استفاده کنند.
- این دستیار صوتی Cortana نام دارد و در همگام تنظیم سوالهای زیادی از شما می کند. هرچه به سوالهای این دستیار صوتی بهتر و مناسبتر پاسخ داده شود جوابهای بهتری دریافت می گردد.



وزارت ارتباطات فناوری اطلاعات



بکارگیری گفتار برای جستجو دروب

- تفاوت استفاده از کلمات و یا عبارتها در جستجوی صوتی بسیار اهمیت دارد. این تفاوتها از آنجاییکه در کارایی یک سیستم تاثیر گذار است:
 - جستجوهای صوتی معمولا بیشتر حالت مکالمه دارند
 - تجربه سرویسهای صوتی نشان داده است که در مکالمه بیشتر کاربران مانند یک دوست از سرویس صوتی سوال می‌پرسند. یکی از سوالهایی که کاربران از سرویس صوتی پرسیده‌اند این است که از «کجا می‌توانم قهوه اسپرسو پیدا کنم؟»
 - جستجوهای صوتی بیشتر فوری هستند
 - استفاده کنندگان از جستجوی صوتی بیشتر به دنبال روش و یا ابزاری هستند تا با فوریت بالا موردی را که می‌خواهند در اینترنت بیابند. از این دست استفاده می‌توان به جستجو برای یک رستوران و یا ساعت کار یک فروشگاه اشاره کرد. تجربه‌ای که در استفاده از سرویس صوتی گوگل نشان می‌دهد که کاربران برای یافتن نیاز خود در جستجوی صوتی نیاز فوری دارند
 - جستجوهای صوتی بیشتر وابسته به مکان هستند



وزارت ارتباطات و فناوری اطلاعات



بکارگیری گفتار برای جستجو دروب

- پنج دلیلی که کاربران از سرویس صوتی استفاده می‌کنند (مطالعات گوگل):
 - سوال در مورد جهت و مسیر
 - تماس با فرد خاصی
 - برای نوشتن متن خاصی
 - برای کمک گرفتن در تکالیف روزانه
 - برای پخش موسیقی
- علاوه بر موارد بالا استفاده از سوالهای پایه افزایش یافته است

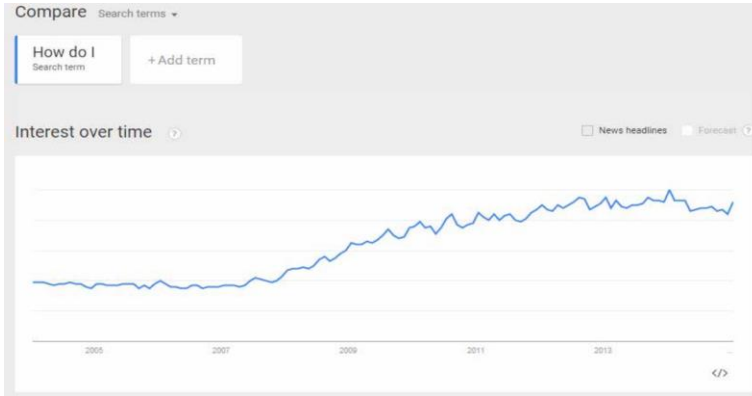


وزارت ارتباطات و فناوری اطلاعات



بکارگیری گفتار برای جستجو دروب

- افزایش استفاده از I how do

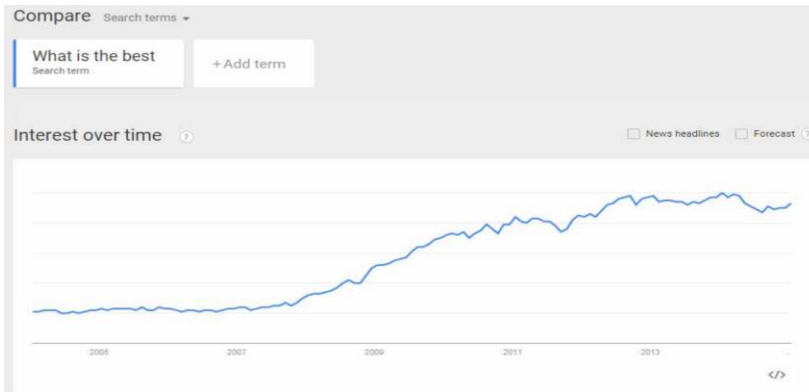


وزارت ارتباطات فناوری اطلاعات



بکارگیری گفتار برای جستجو دروب

- افزایش استفاده از what is the best



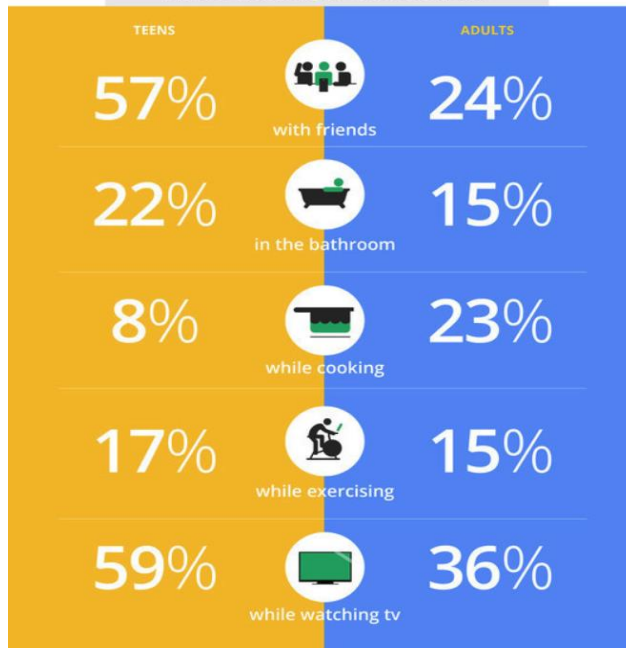
وزارت ارتباطات فناوری اطلاعات



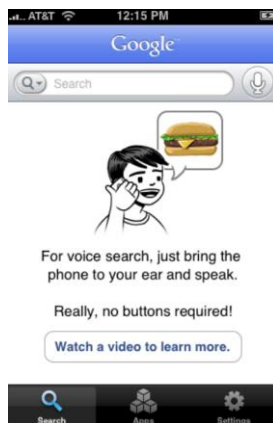


وزارت ارتباطات و فناوری اطلاعات 23

WHEN do we use voice search?



وزارت ارتباطات و فناوری اطلاعات 24



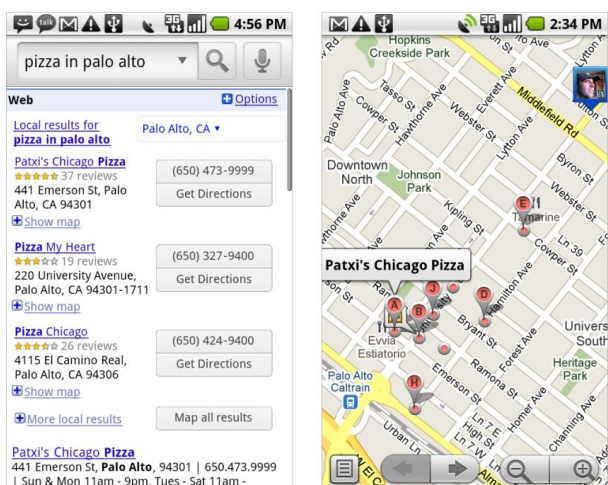
وزارت ارتباطات و فناوری اطلاعات

بکارگیری گفتار برای جستجو در وب

- تلاش گوگل برای یافتن معیارهایی که به تجربه کاربر نزدیکتر باشد منجر به معیارهای زیر گردید:
 - نرخ خطای کلمه
 - کیفیت معنایی
 - ابهام زدایی
 - خارج از واژگان
 - تاخیر



وزارت ارتباطات و فناوری اطلاعات



وزارت ارتباطات و فناوری اطلاعات

جستجوی گفتاری در طرح جویسگر

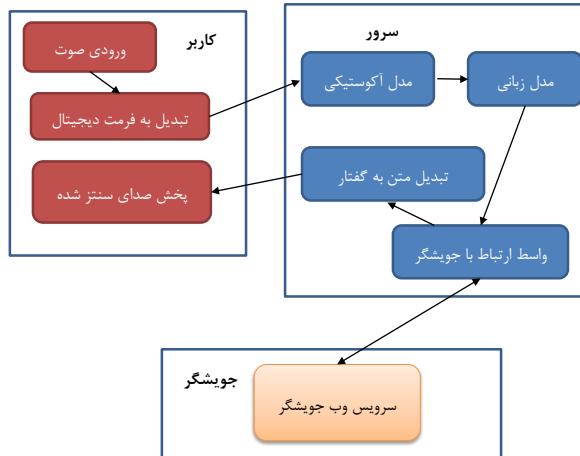
• وضع موجود

- تخصص شرکتهای ایرانی در زمینه پردازش گفتار
- تجربه در زمینه واژگان و دادگان گفتاری
- تجربه در زمینه طراحی نرم افزارهای پردازش گفتار برای تلفنهای همراه



وزارت ارتباطات و فناوری اطلاعات

جستجوی گفتاری در طرح جویشر



وزارت ارتباطات و فناوری اطلاعات

جستجوی گفتاری در طرح جویشر

- ارزیابی میزان دقت جستجوی گفتاری
 - دقت سیستم بازشناسی گفتار در سطح حرف
 - دقت بازشناسی گفتار در سطح کلمه
 - دقت بازشناسی گفتار پیوسته
 - دقت بازشناسی گفتار گسسته
 - دقت مدل زبانی
 - کیفیت گفتار سنتز شده بر اساس معیار MOS
 - کیفیت گفتار سنتز شده بر اساس معیارهای DRT و MRT
 - کیفیت نرم افزار سمت کلاینت در ضبط و پخش صوت



وزارت ارتباطات و فناوری اطلاعات

جستجوی گفتاری در طرح جویشر

- ارزیابی میزان دقت جستجوی گفتاری
 - تست کیفی نرم افزار
 - تست کارایی سرور
 - تست تعداد کاربر همزمان
 - تست کارایی سیستم کیفیت نسبت به پهنای باند مصرفی
 - تست مقاومت در مقابل نویز
 - تست تطبیق پذیری نسبت به نرخ بیت



وزارت ارتباطات و فناوری اطلاعات

RFP

- جستجوی گفتاری جویشرهای بومی
 - سیستم کامل بازشناسی گفتار به همراه تبدیل متن به گفتار فارسی
 - ارائه برای کامپیوترهای رومیزی و گوشیهای تلفن همراه با سیستم عامل اندروید
 - پوشش پرسش و پاسخهای مرسوم کاربران موتورهای جستجوی بومی
 - پاسخگویی برخط پرسش کاربران
 - تبدیل متن به گفتار با کیفیت و خوشایندی بالا
 - قابلیت استفاده با خطوط اینترنتی ثابت و سیار با حداقل سرعت ۱۲۸ کیلو بایت بر ثانیه
 - قابلیت پاسخگویی به تعداد بالای کاربر همزمان



وزارت ارتباطات و فناوری اطلاعات

گامهای آتی

- بهبود عملکرد و توسعه سیستم جستجوی گفتاری وب
 - اضافه شدن زبان انگلیسی و زبانهای دیگر
 - ارتقاء کیفی و دقت بازشناسی و سنتز گفتار
 - ارتقاء عملکرد (پاسخگویی به تعداد بالای کاربران)
 - افزایش حجم دادگان
 - افزایش کیفیت
- ارائه سرویسها جدید صوتی و گفتاری
 - تایپ گفتاری
 - جستجوی صوتی در اسناد



وزارت ارتباطات و فناوری اطلاعات

با تشکر



وزارت ارتباطات و فناوری اطلاعات