



ارائه شده توسط:

سایت ترجمه فا

مرجع جدیدترین مقالات ترجمه شده

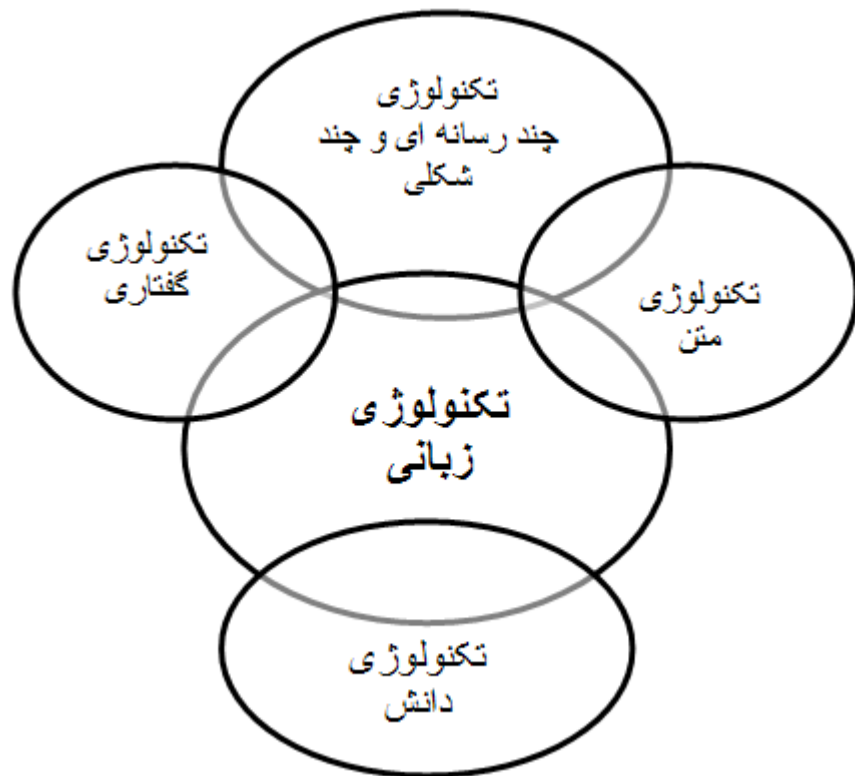
از نشریات معتبر

# تکنولوژی زبان

## -قلمرو

تکنولوژی های زبانی تکنولوژی اطلاع رسانی هستند که متخصص پرداختن به پیچیده ترین رسانه اطلاع رسانی در جهان ما یعنی: زبان انسانی، هستند بنابراین، این تکنولوژیها نیز اغلب تحت اصطلاح تکنولوژی زبان انسانی قرار می گیرند . زبان انسانی به شکل گفتاری و نوشتاری ظاهر میشود. در حالی که گفتار قدیمی ترین و طبیعی ترین شیوه ارتباطات زبانی است، با این حال ،اطلاعات پیچیده و بخش اعظم دانش انسانی در متون نوشتاری نگهداری و منتقل می شود . تکنولوژی های گفتار و متن با استفاده از این دو شیوه درک اقدام به پردازش و ایجاد زبان می کنند . اما زبان جنبه هایی نیز دارد که در میان گفتار و متن مشترک هستند از قبیل کتب فرهنگ لغت، بخش اعظم دستور زبان ، و معنی جملات . بنابراین ، بخش های بزرگی از تکنولوژی زبانی را نمی توان در زمره تکنولوژی گفتار و متن قرار داد . در میان آنها تکنولوژی هایی هستند که زبان را به دانش متصل می کنند . ما نمی دانیم که چگونه زبان ، دانش و اندیشه در مغز انسان بازنمود می یابند . با این حال ، تکنولوژی های زبان ناگزیر به ایجاد سیستم های بازنمود رسمی که زبان را به مفاهیم و وظایف در دنیای واقعی پیوند می دهند، هستند . این [ تکنولوژی ] میانجی و واسطی را به حوزه سریع در حال رشد تکنولوژی دانش فراهم می کند.

ما در برقراری ارتباط ، زبان را با دیگر شیوه های ارتباطات و دیگر رسانه های اطلاع رسانی در هم می آمیزیم. گفتار را با حالات دست و سر و صورت ترکیب می کنیم . متون دیجیتال با تصاویر و صداها ترکیب می شوند . حرکات ممکن است حاوی زبان و شکل گفتاری و نوشتاری باشند . بنابراین تکنولوژی گفتار و متن با بسیاری از تکنولوژی های دیگر که پردازش ارتباطات چند مدلی و اسناد چند رسانه ای را تسهیل می کنند، همپوشانی و تعامل دارند .



برای آشنایی جامع با این حوزه ، خوانندگان این مقاله رابه

ColeR.A., J.Mariani, H.Uszkoreit, G.Varile, A.Zaenen, V.Zue, A.Zampolli (Eds.)  
 (1997)Survey of the State of the Art in Human Language Technology , Cambridge  
 University Press and Giardini.(<http://www.dfki.de/~hansu/HLT-Survey.pdf>)

ارجاع می دهیم.

## 2- کاربرد

اگر چه سیستم های LT موجود از رسیدن به قابلیت انسانی دور هستند ، با این حال کاربرد های احتمالی متعددی دارند. هدف آنها ایجاد محصولات نرم افزاری است که تاحدی دانش زبان انسانی را داشته باشند. قراراست که اینگونه محصولات زندگی مان را تغییر دهند و وجود آنها جهت بهبود تعامل میان انسان با ماشین کاملا ضروری است زیرا مانع اصلی در تعامل میان انسان و کامپیوتر صرفا وجود یک مشکل ارتباطی است . کامپیوترهای امروزی زبانمان را درک نمی کنند اما یادگیری زبان های کامپیوتری دشوار است و به ساختار اندیشه انسانی مربوط نمی شوند . حتی اگر زبانی را که ماشین درک می کند و نیز دامنه

بحث آن بسیار محدود باشد ، باز هم ، استفاده از زبان انسانی می تواند پذیرش نرم افزار و بهره وری کاربران آن را افزایش دهد.

" تکنولوژی دوستانه " باید گوش دهد و صحبت کند میانجی های زبان طبیعی کاربر را به برقراری ارتباط با کامپیوتر به زبان فرانسه ، انگلیسی ، آلمانی ، یا دیگر زبانهای انسانی قادر می سازند. برخی از کاربردهای چنین میانجی هایی عبارتند از پرس و جوهای پایگاههای داده ای ، بازیابی اطلاعات از متون ، سیستم های موسوم به خبره ، کنترل روباتی. ارتباط با کامپیوترها با استفاده از زبان گفتاری اثر ماندگاری بر محیط کاری خواهد داشت . حوزه های کاملاً جدیدی برای کاربرد تکنولوژی اطلاع رسانی باز خواهند شد. هرچند لازم است که زبان گفتار را با دیگر شیوه های ارتباطی از قبیل اشاره کردن با ماوس یا انگشت ترکیب نمود . اگر بتوانیم چنین ارتباط چند مدلی را سرانجام در مدل عمومی موثر تعاون تجسم ببخشیم ، آنگاه در تبدیل ماشین به یک شریک موفق شده ایم . هدف غایی تحقیق دسترسی فراگیر به همه انواع تکنولوژی و به ساختار اطلاع رسانی جهانی از طریق تعامل طبیعی است . در یک سناریوی جاه طلبانه اما نه چندان دور از ذهن ، تکنولوژی زبانی میانجی را برای یک هوش فراگیر که ارائه دهنده کمک در سر کار و در بسیاری از وضعیتهای زندگی روزمره است فراهم می آورد.

ماشین ها همچنین می توانند به افراد برای برقراری ارتباط با همدیگر کمک کنند.

تکنولوژی های زبانی همچنین می توانند به افراد کمک کنند تا با همدیگر ارتباط برقرار کنند. مشکلات ارتباطی میان افرادی که زبانهای مادری متفاوت دارند. در مقایسه با مشکلات میان موجودات بشری و ماشین ها از قدمت بیشتری برخوردارند. همیشه یکی از هدفهای اصلی تکنولوژی زبانی ترجمه کاملاً خودکار میان زبانهای انسانی بوده است . دانشمندان از تجربه تلخ نتایج گذشته دریافته اند که هنوز هم از رسیدن به هدف جاه طلبانه ترجمه متون نامحدود بسیار دورند. با این حال ، آنها توانسته اند سیستم های نرم افزاری را ایجاد کنند که کار مترجمان انسانی را ساده می کنند و بهره وری آنها را بوضوح بهبود می

بخشند. ترجمه های کمتر از تمام خود کار نیز می توانند برای جستجو کنندگان اطلاعات که ناگزیرند در میان مقادیر زیادی متن به زبانهای خارجی اطلاعات مورد نیاز خود را جستجو کنند، کمک بزرگی باشند. جدی ترین گلوگاه تجارت الکترونیک حجم ارتباطات میان موسسات تجاری و مشتریان یا در میان موسسات تجاری می باشد. تکنولوژی زبانی می تواند به مرتب سازی، تصفیه سازی و مسیر سازی نامه های الکترونیک وارده کمک کند. همچنین می تواند به عامل ارتباطی مشتریان جهت جستجوی اطلاعات و نوشتن پاسخ کمک کند. در مواردی که پرسشها از قبل پاسخ داده شده اند، تکنولوژی اطلاعات می تواند پاسخ های مناسب سریع تری را بیابد و بطور خودکار به آنها پاسخ دهد.

### زبان تارو پود شبکه است.

رشد سریع اینترنت و یا WWW و ظهور جامعه اطلاعاتی چالش های هیجان انگیز جدیدی را برای تکنولوژی زبانی بوجود می آورد. اگر چه رسانه های جدید متن، گرافیک، صدا و تصاویر متحرک را با هم ترکیب می کنند، باین حال کل جهان اطلاعات چند رسانه ای را تنها می توان از طریق زبان سازماندهی، نمایه سازی و جهت یابی نمود. برای مرور (browsing)، جهت یابی (narigoting)، تصفیه سازی (filtering) و پردازش اطلاعات بر روی وب، ما به نرم افزاری نیاز داریم که بتواند در محتویات اسناد وارد شود. تکنولوژی زبانی برای مدیریت محتوا پیش شرطی ضروری برای تبدیل ارزش اطلاعات دیجیتال به دانش جمعی می باشد. چند زبانی فزاینده وب چالش دیگری را برای تکنولوژی زبانی تشکیل می دهد. تنها با کمک ابزار چند زبانه برای نمایه سازی و جهت یابی می توان بر وب جهانی تسلط یافت. سیستم های اطلاعات چند زبانه (crosslingual) و مدیریت دانش بر موانع زبانی که سد راه تجارت الکترونیک، آموزش و همکاری بین المللی اند، فائق می آیند.

### 3- تکنولوژی ها

در دنباله این بحث ، انتخابی از مربوط ترین تکنولوژی های زبانی بطور خلاصه نشان داده خواهند شد . با کلیک کردن بر روی اسامی تکنولوژی ها ، می توانید به اطلاعات اضافی دسترسی پیدا کنید.

#### شناسایی گفتار

زبان گفتاری شناسایی می شود و مثلا در سیستمهای دیکته ، به متن ، یا در سیستم های کنترل روبات ، به فرامین ، یا به بازنمود درونی دیگری ، تبدیل می شود.



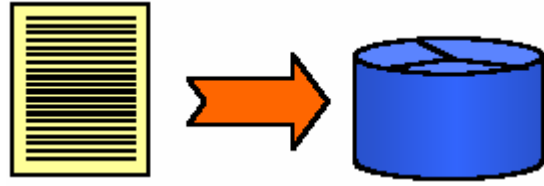
#### ترکیب گفتار

ادها (Utterance) در زبان گفتاری از متن (سیستم های متن - به - گفتار ) یا از بازنمودهای درونی لغات یا جملات ( سیستم های مفهوم - به - گفتار) تولید می شوند.



#### مقوله بندی متن

این تکنولوژی متون را به مقولات اختصاص می دهد. متون ممکن است به بیش از یک مقوله متعلق باشند، مقوله ها ممکن است حاوی مقولات دیگری باشند. تصفیه سازی حالت خاصی از مقوله بندی دارای تنها دو مقوله است.



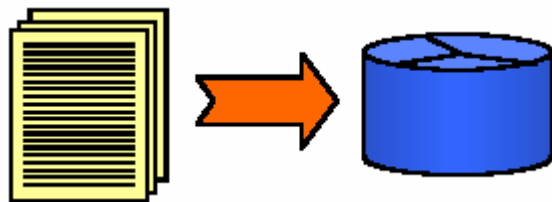
### تلخیص متن

مربوط ترین بخش های یک متن به صورت خلاصه استخراج می شوند. این کار به طول و درازای مورد نیاز خلاصه ها بستگی دارد. در صورتی که لازم باشد که خلاصه به یک پرسش معین مختص باشد، تلخیص سخت تر است.



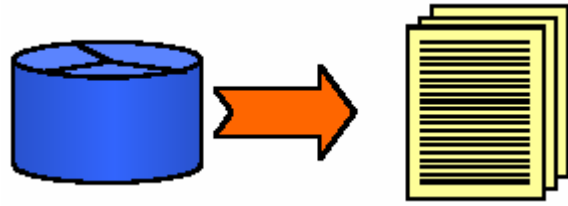
### نمایه سازی متن

به عنوان پیش شرطی برای بازیابی سند، متون در یک پایگاه داده ای نمایه سازی شده ذخیره می شوند. معمولاً متن برای همه شکل‌های کلمه یا - بعد از شرح و تفسیر برای همه کلمات نمایه می شود. گاهی اوقات نمایه سازی با مقوله بندی و تلخیص ترکیب می شود.



### بازیابی متن

متون از یک پایگاه داده ای که بیشتر با یک پرسش یا سند معین مطابقت می کند، بازیابی می شوند. اسناد مورد نظر با توجه به مناسبت مورد انتظارشان مرتب می شوند. نمایه سازی، مقوله بندی، تلخیص و بازیابی اغلب تحت اصطلاح بازیابی اطلاعات قرار می گیرند.



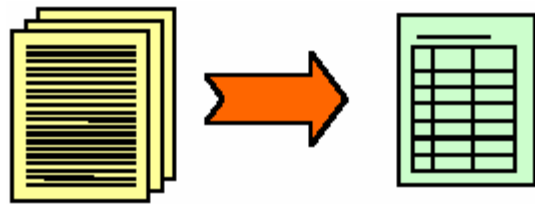
### استخراج اطلاعات

تکه های اطلاعاتی مربوط و مناسب اطلاعات کشف و برای استخراج نشان دار می شوند: قطعات استخراج شده می توانند به شکل های ذیل باشند: موضوع ، هویت های با نام از قبیل اسامی شرکت ، مکان یا شخص ، رابطه های ساده از قبیل قیمتها ، مقاصد، کارکردها و غیره و یا رابطه های پیچیده از قبیل شرح دادن تصادفات ، ادغام شرکتها یا مسابقات فوتبال .



### ترکیب داده ها و استخراج داده های متن

قطعات استخراج شده اطلاعات از چندین منبع در یک پایگاه داده ای ترکیب می شوند. ممکن است مناسباتی که از قبل تشخیص داده نشده اند ، کشف شوند.



### پاسخ به پرسش

پرسشهای زبان طبیعی برای دسترسی به اطلاعات در پایگاه داده ای مورد استفاده قرار می گیرند. پایگاه داده ای ممکن است مبنای داده های ساختاری شده یا مخزنی از ستون دیجیتال باشد که در آن بخش های معین به عنوان پاسخ های بالقوه نشان دار شده اند.





## ایجاد گزارش

گزارش به زبان طبیعی ایجاد می شود که محتوای اساسی پایگاه داده ای یا تغییرات آن را شرح می دهد. گزارش می تواند حاوی اعداد جمع شده ، حداکثر ، حداقل و بنیادی ترین تغییرات باشد.



## سیستم های مباحثه گفتاری

سیستم می تواند گفتگویی را با یک کاربر انسانی انجام دهد که در آن کاربر می تواند اطلاعات را درخواست کند یا خریدها ، شرایط یا دیگر معاملات را هدایت کند.



## تکنولوژی های ترجمه

تکنولوژی هایی که متون را ترجمه می کنند یا به مترجمان کمک می کنند . ترجمه خودکار یا ترجمه ماشینی خوانده می شود . حافظه های ترجمه از مقادیر بزرگی متن همراه با ترجمه های موجود برای دنبال کردن موثر ترجمه های احتمالی کلمات ، عبارات و جملات استفاده می کنند.



## روش ها و منابع

از آنجایی که تحقیق و مدل سازی زبان انسانی یک تلاش بین رشته ای صادقانه می باشد ، لذا روشهای تکنولوژی زبانی از چندین رشته ناشی می شوند : علوم کامپیوتر ، زبان شناسی رایانشی و نظری ، ریاضیات ، مهندسی برق و روان شناسی.

## روش های کلی CS

زبان های برنامه نویسی ، الگوریتم ها برای انواع داده های کلی ، و روش های مهندسی نرم افزار برای ساختارسازی و سازماندهی توسعه نرم افزار و تضمین کیفیت.

## الگوریتم های تخصصی

الگوریتم های آماری بطور مخصوصی در پردازش گفتاری ، بازیابی اطلاعات ، و دستیابی خودکار به مدل های زبانی موفق شده اند . مدل های دیگر در این رده ، شبکه های عصبی و فنون قدرتمند برای بهینه سازی و جستجو می باشند.

## صورت گزایی های منطقی و زبانشناختی

برای پردازش عمیق زبانشناختی، صورت گزایی های دستورزبانی مبتنی بر محدودیت به خدمت گرفته می شوند. صورت گزایی های پیچیده برای بازنمود محتوا و دانش معنی شناختی توسعه داده شده اند .

## دانش زبانشناختی

منابع دانش زبان شناختی برای بسیاری از زبانها مورد استفاده و مفید هستند : دستورزبانهای ریخت شناختی و نحوی ، قواعد تاویل معنی شناختی ، تلفظ و آهنگ (intonation)

## پیکره های زبانی و ابزارهای پیکره زبانی (corpus)

مجموعه های بزرگ مختص به کاربرد یا مجموعه های کلی و عام زبان گفتاری و نوشتاری برای فراهم آوری و آزمون کردن مدل های آماری یا مدل های زبانی مبتنی بر قواعد مورد استفاده قرار می گیرند.<sup>۱</sup>

---

<sup>۱</sup> - منبع: (هانس اوزکورت (Hans Uszkoreit) مترجم: سیاوش مولایی بیرگانی دانشجوی کارشناسی ارشد کتابداری و اطلاع رسانی پزشکی دانشکده مدیریت و اطلاع رسانی دانشگاه علوم پزشکی ایران



این مقاله، از سری مقالات ترجمه شده رایگان سایت ترجمه فا میباشد که با فرمت PDF در اختیار شما عزیزان قرار گرفته است. در صورت تمایل میتوانید با کلیک بر روی دکمه های زیر از سایر مقالات نیز استفاده نمایید:

لیست مقالات ترجمه شده ✓

لیست مقالات ترجمه شده رایگان ✓

لیست جدیدترین مقالات انگلیسی ISI ✓

سایت ترجمه فا ؛ مرجع جدیدترین مقالات ترجمه شده از نشریات معتبر خارجی