



وزارت علوم، تحقیقات و فناوری
پژوهشگاه علوم و فناوری اطلاعات ایران

عموان سخنرانی

تکنیک‌های جستجو، مستی‌شناسی‌ها

المیرا کریمی

دانشجوی دکتری علم اطلاعات و دانش‌شناسی
(پژوهشگاه علوم و فناوری اطلاعات ایران)

اردیبهشت ۹۷

جستجوی اطلاعات

هستی‌شناسی در جستجوی اطلاعات

فعالیت‌های جستجو

تاکتیک‌های جستجو

تاکتیک‌های جستجو و هستی‌شناسی‌ها

تاکتیک‌های جستجو در گنج

جستجوی اطلاعات



➤ فعالیت پیچیده انسانی

➤ آغاز از نیاز اطلاعاتی

➤ مستلزم شکل خاصی از استراتژی

اصلاح پرس و جو

➤ کاربران به صورت مداوم پرسش جستجوی قبلی خود را به امید بازیابی بهتر اصلاح می کنند.
این اصلاحات فرمول‌بندی مجدد یا اصلاح پرس و جو نامیده می شود.

وضعیت حال حاضر سیستم‌ها در تعامل کاربر، مطابقت براساس تایپ کلمات و مطابقت نمایه کلمات

استفاده از اصطلاحات کم و استفاده نکردن از جستجوی پیشرفته

نیاز به استفاده از فرمول‌بندی در جستجو بدون پیچیده کردن سیستم برای تعاملی کردن بازیابی اطلاعات



هستی‌شناسی در جستجوی اطلاعات

- مورد توجه‌ترین ابزار بازنمایی دانش در سال‌های اخیر
- بازنمون دانشی ساختاریافته در مفهوم معین هستند
- فراهم‌آوری تفسیر برای جستجو با توجه به محتوای اسناد
- به کارگیری در سیستم جستجو
- استفاده در رابط کاربر برای تایپ اصطلاحات از سوی کاربر
- استفاده از هستی‌شناسی در توسعه پرس‌وجو



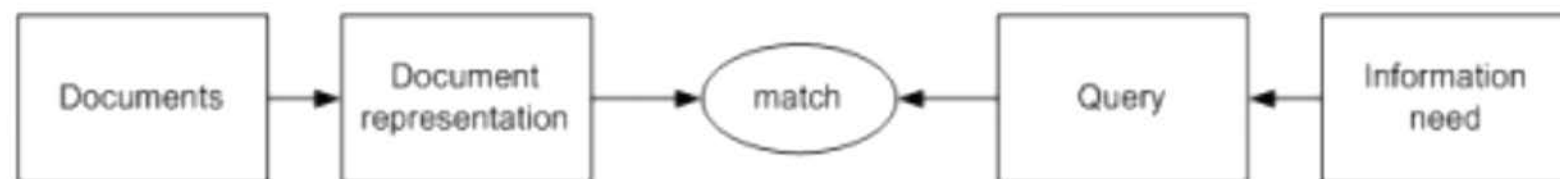
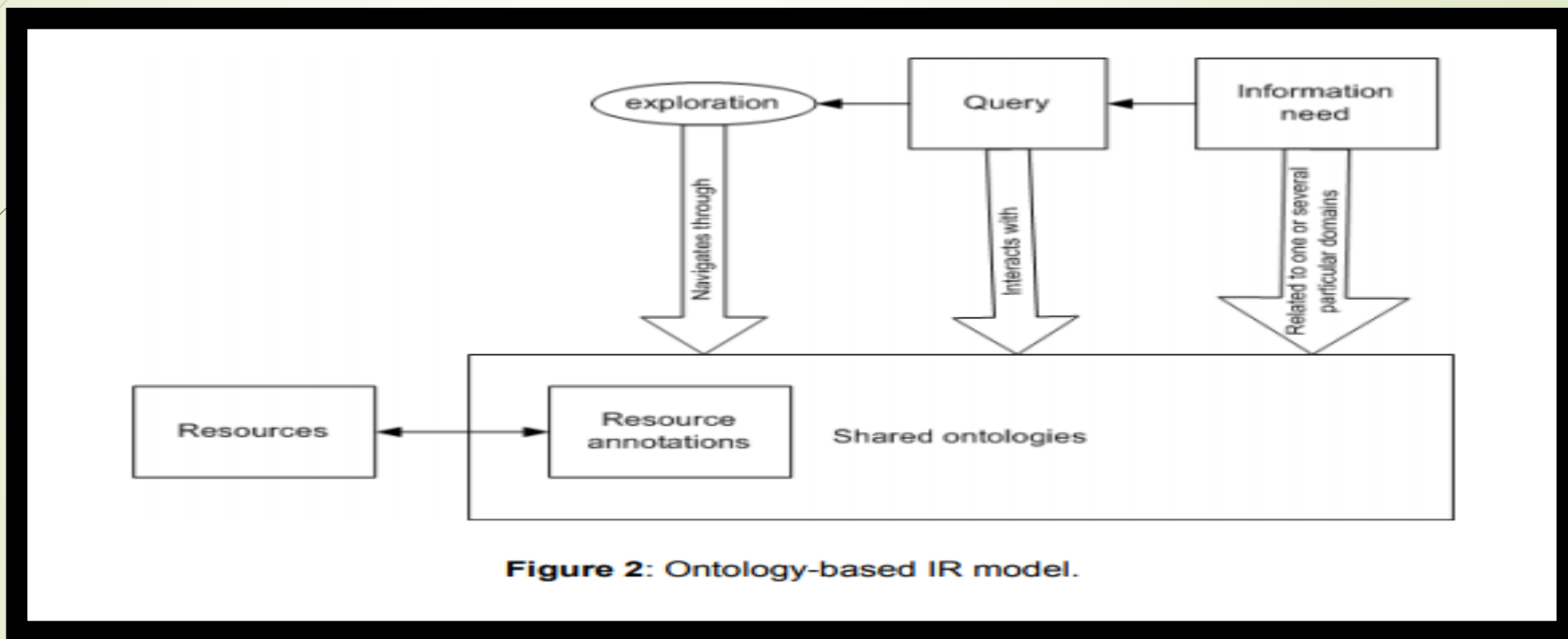


Figure 1: Classical IR model

Figure 1: Classical IR model

بازیابی اطلاعات مبتنی بر هستی‌شناسی



فعالیت‌های جستجو

۴ سطح یا نوع فعالیت‌های جستجو توصیف شده‌اند:



تاکتیک‌های جستجو



تاکتیک‌های نظارتی

تاکتیک Check ➤

تاکتیک Record ➤



تاکتیک‌های ساختاری

SELECT ➤

SURVEY ➤

CUT ➤

STRETCH ➤



تاکتیک‌های فرمول‌بندی

SPECIFY ➔

EXHUST ➔

REDUCE ➔

PARALLEL ➔

PINPOINT ➔



تاکتیک‌های اصطلاحی

SUPER ➤

SUB ➤

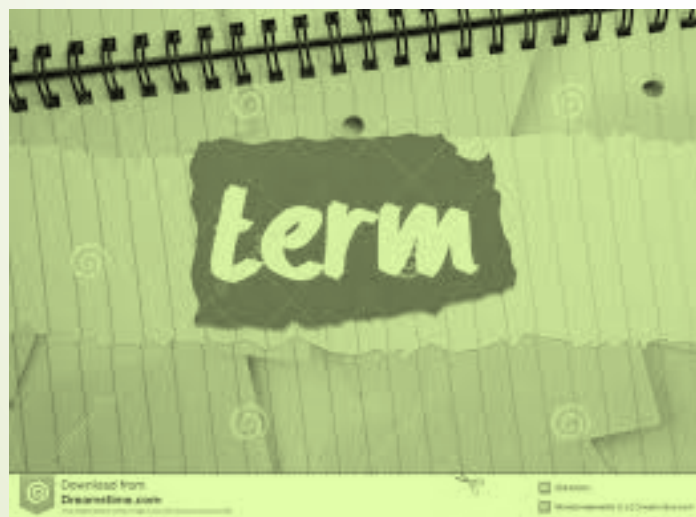
RELATE ➤

REARRANGE ➤

CONTRARY ➤

RESPELL ➤

RESPACE ➤



تاکتیک‌های ایده‌ای



RESCUE ➤

BREACH ➤

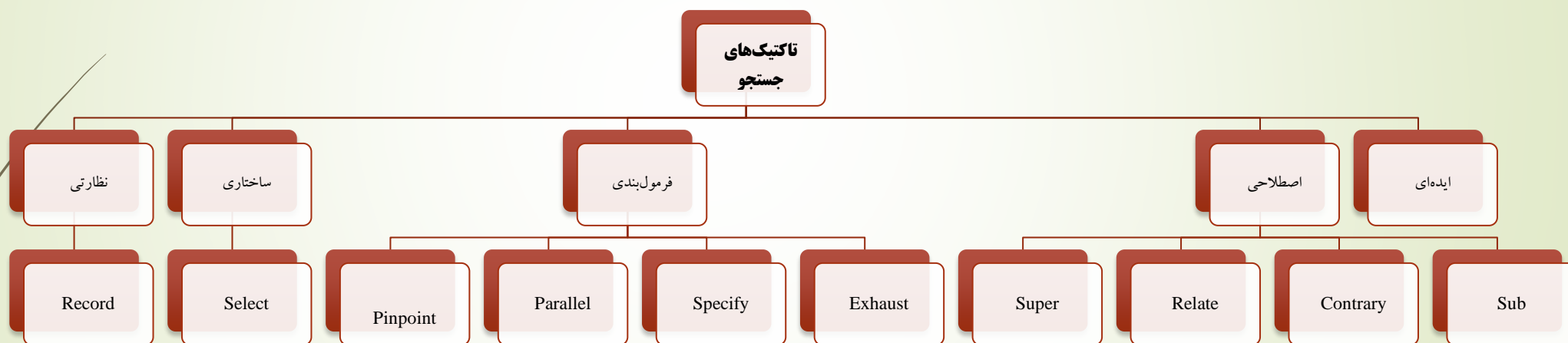
FOCUS ➤

تاکتیک‌های جستجو و هستی‌شناسی‌ها

برای پرس و جوهای مبهم نیاز به تعامل با کاربر وجود دارد.

بیتس (۱۹۹۰) اشاره می‌کند، گاهی کاربر ترجیح می‌دهد که فرایند جستجو را تاحدی، مطابق با استراتژی‌ها و تاکتیک‌های مشخصی کنترل کند. هستی‌شناسی‌ها می‌توانند برای **حمایت مستقیم تاکتیک‌ها** استفاده شوند.

تاکتیک‌های جستجو و هستی‌شناسی‌ها



تاکتیک‌های جستجو در گنج

- بررسی کاربران گنج از طریق تحلیل لاگ‌های جستجو
- بازه زمانی سه ماهه (از اردیبهشت ۹۵ تا مرداد ۹۵)
- تعداد کل جستجوهای دریافتی ده میلیون رکورد
- محدودسازی به جستجوهای حوزه اطلاع‌رسانی (با محدودسازی با اصطلاحات اصطلاحنامه)
- تعداد کل جستجوهای حوزه اطلاع‌رسانی ۱۰۶۶۴۱
- اطلاعاتی که در لاگ‌های دریافت شده از کاربر ارائه شده بود شامل: شماره جستجو، عبارت جستجو، تاریخ جستجو، زمان جستجو، تعداد رکوردهای بازیابی شده و شماره آی پی کاربر بود

تاکتیک‌های جستجو در گنج

- واحد اصلی تحلیل اصلاح پرس و جو
 - به دلیل تمرکز این مطالعه بر فرمول بندی مجدد پرس و جو
 - پرس و جوهای انتخاب شدند که محتوای مورد جستجوی خود را (کلیدواژه‌ها) اصلاح کرده بودند.
- تعریف ما از فرمول بندی مجدد پرس و جو:
- ترکیب ویژگی‌های بهینه از پژوهش پیشین، کلمات مشترک و زمان

تاکتیک‌های جستجو در گنج

جستجوهای اصلاح شده ۵۹۶

تعداد کل تاکتیک‌های مورد استفاده کاربران ۸۰۰ مورد

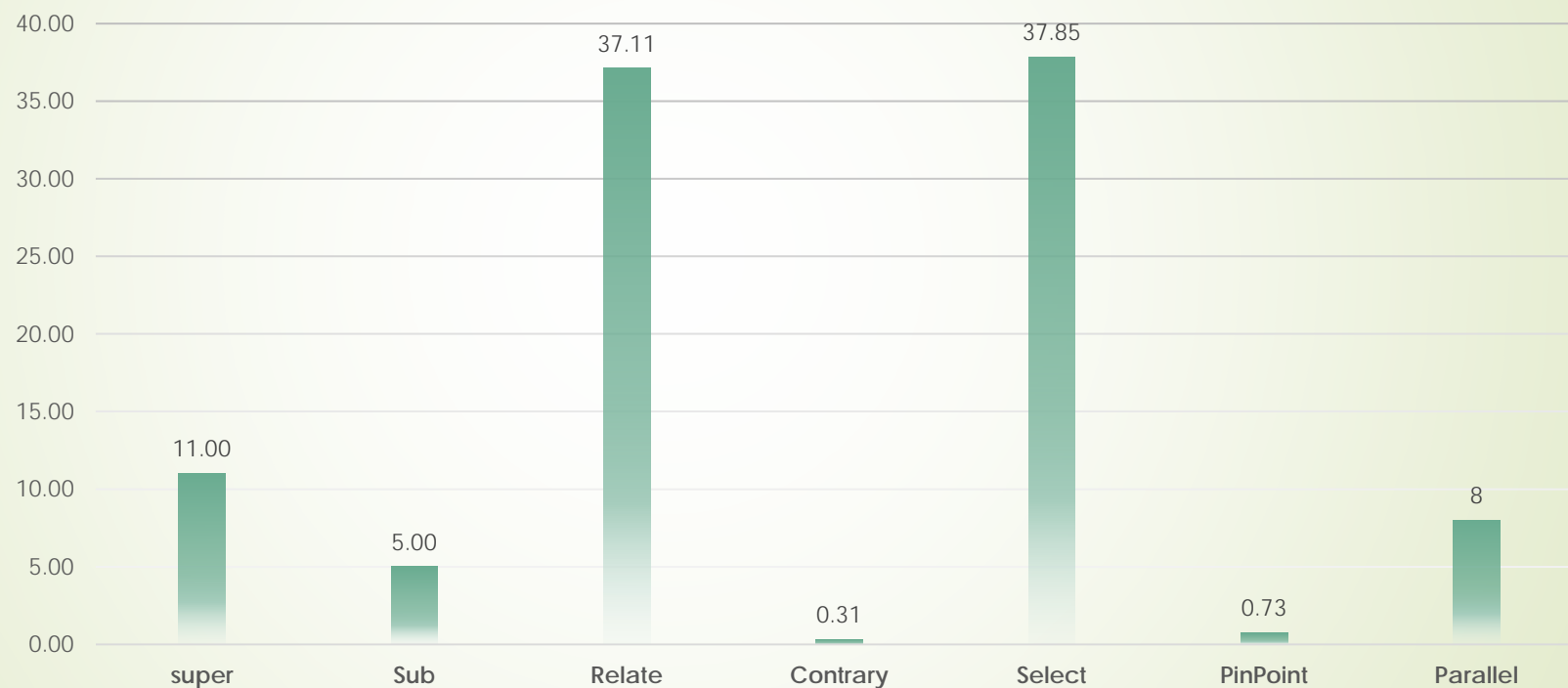
شناسه کاربر	آی پی	تعداد نتایج بازیابی شده	تاریخ	کلیدواژه جستجو شده
01fef46c288769d152cee011d5b63415	۵,۷۴,۶۱,۷۹	۲۵۸	۰۵/۰۹/۲۰۱۶	کتابخانه کودک
01fef46c288769d152cee011d5b63415	۵,۷۴,۶۱,۷۹	۲۵۸	۰۵/۰۹/۲۰۱۶	کتابخانه کودک
01fef46c288769d152cee011d5b63415	۵,۷۴,۶۱,۷۹	۲۵۸	۰۵/۰۹/۲۰۱۶	کتابخانه کودک
01fef46c288769d152cee011d5b63415	۵,۷۴,۶۱,۷۹	۱۶	۰۵/۰۹/۲۰۱۶	کتابدار کودک
01fef46c288769d152cee011d5b63415	۵,۷۴,۶۱,۷۹	۱۶	۰۵/۰۹/۲۰۱۶	کتابدار کودک

تاکتیک‌های جستجو در گنج

وضعیت استفاده کاربران پایگاه اطلاعات گنج از تاکتیک‌های حامی هستی‌شناسی

PARALLEL	PINPOINT	SELECT	CONTRARY	RELATE	sub	super	
۷۲	۷	۳۴۱	۳	۳۳۴	۴۶	۹۷	تعداد
۸	۰,۷۳	۳۷,۸۵	۰,۳۱	۳۷,۱۱	۵	۱۱	درصد

تاکتیک‌های جستجو در گنج



تاکتیک‌های جستجو در گنج

بیشترین تاکتیک مورد استفاده (تاکتیک‌های ساختاری و اصطلاحی)

➤ **Relate**: حرکت سلسله‌مراتبی به سمت اصطلاح هماهنگ

➤ **Select**: تجزیه پرس و جوی جستجوی پیچیده به مشکلات فرعی و کاربر بر روی یک مشکل در یک زمان

کمترین تاکتیک مورد استفاده (تاکتیک‌های اصطلاحی و فرمول‌بندی)

➤ **CONTRARY**: جستجو برای اصطلاح متضاد منطقی از توصیف اطلاعات مطلوب

➤ **PINPOINT**: ایجاد فرمول‌بندی جستجوی دقیق با حداقل‌سازی (یا کاهش) تعداد اصطلاحات برابر، حفظ اصطلاحات توصیفی‌تر

- ▶ طراحی رابط کاربر مبتنی بر آنتولوژی برای جستجو یا کاوش اطلاعات راهی جدید برای جستجوی آینده است که هم موضوعات کاربرپذیری و هم موضوعات فنی مرتبط با تفسیر پرس و جوها را در بر می‌گیرد.
- ▶ نتایج این مطالعه نشان می‌دهد که ایجاد ابزارهای جستجویی که رفتارهای فرمول‌بندی مجدد پرس و جوی پیچیده را که چندین بار در فرایند تعامل IR اتفاق می‌افتد، ساده می‌کند، مهم است.
- ▶ براساس مطالعات این پژوهش، ابزار فرمول‌بندی مجدد تعاملی می‌تواند برای ارتقا و کمک به درگیر شدن کاربر در فرایند فرمول‌بندی مجدد کاربر طراحی شود.
- ▶ با استفاده از تحلیل رفتاری که در این پژوهش انجام شد، می‌توان در جهت ارائه پیشنهاد جستجوی مناسب برای یافتن پاسخ مطلوب کاربر پیش رفت و سیستم را به گونه‌ای بهبود داد که امکان دسترسی سریع و مطلوب کاربر به پاسخ موردنظر فراهم گردد.

سپاس

