

تعریف موتور جستجو: موتور جستجو یا همان جویشگر به برنامه‌ای گفته می‌شود که موضوعات موردنظر کاربران را در قالب‌کلمات کلیدی، درون اسناد و اطلاعات موجود در اینترنت کاوش می‌کند و نتایج را در قالب آدرس محل عرضه می‌نماید.

در موتور جستجو کاربر کلیدواژه‌ای را وارد کرده و برنامه‌ی جستجو در بانک اطلاعاتی خود به جستجو می‌پردازد و سایت‌های مرتبط با موضوع شما را نمایش خواهد داد. البته امروزه بیشتر موتورهای جستجو دارای اطلاعات طبقه‌بندی شده‌ای مانند دایرکتوری‌ها می‌باشند. موتورهای جستجوی عمومی محتویات تمامی وب‌سایت‌ها را جستجو می‌کنند و چکیده‌ای از آن‌ها را در یک پایگاه داده‌ی شاخص‌دهی شده<sup>۱</sup> نگه‌داری می‌کنند.

بزرگترین مشکل ساخت موتورهای جستجو را می‌توان ناهمگونی وحشتناک اسناد و محتویات وب و راضی نگه‌داشتن انبوهی از کاربرانی تجربه و پرتوقع دانست. آندسته از موتورهای جستجویی که بر اساس تطبیق‌کلمات کلیدی با محتویات اسناد کار می‌کنند نتایجی را بر می‌گردانند که ارتباط کمی با موضوع مورد نظر کاربران دارد.

اجزای معماری تشکیل‌دهنده یک موتور جستجو هشت مورد است:

- درون‌کاو - کنترل درون‌کاو - انباری صفحات - مایجول شاخص‌دهی

- index utility - collection analysis module - موتور پرس‌وجو - مایجول رتبه‌بندی

**درون‌کاو یا خزنده<sup>۲</sup>:** هر درون‌کاو یا خزنده از تعداد قابل توجهی برنامه‌ی کوچک تشکیل شده است که هر کدام از صفحه‌ی شروع یا هسته یک وب‌سایت آغاز به کار می‌کنند و پس از دریافت آن صفحه و تحویل به انباری صفحات<sup>۳</sup> لینک‌های درون صفحه‌ی جاری و صفحات پیوندخورده رانیز به همین روال استخراج و ذخیره می‌کنند.

**انباری صفحات<sup>۴</sup>:** بانک اطلاعاتی صفحاتی که توسط مایجول درون‌کاو استخراج می‌شود در انباری صفحات قرار می‌گیرد تا برای شاخص‌دهی و تحلیل محتوایی در اختیار دیگر مایجول‌ها قرار گیرد

**مایجول شاخص‌دهی<sup>۵</sup>:** این مایجول کلمات موجود در صفحات را به همراه URL آن‌ها در یک جدول بسیار عظیم لیست می‌کند.

یک خروجی مایجول شاخص‌دهی، بانک اطلاعاتی شاخص ساختاری<sup>۵</sup> است. این بانک چگونگی پیوند خوردن صفحات را نشان می‌دهد.

---

<sup>۱</sup> indexed database

<sup>۲</sup> crawler

<sup>۳</sup> Page repository

<sup>۴</sup> page repository

<sup>۵</sup> Indexer

ماجول رتبه‌بندی<sup>۷</sup>: نتایج حاصل از جستجوی کلمات کلیدی کاربر را به نحوی رتبه‌دهی می‌کند تا مرتبط‌ترین آیتم در صدر نتایج جستجو قرار گیرد. فرآیند درون‌کاوی وب با پیچیدگی‌های زیر رو به رو است:

- انتخاب صفحات - مدل درون‌کاوی - تازه‌سازی و سرکشی دوره‌ای به صفحات وب

#### معیارهای انتخاب صفحات:

۱- معیار مبتنی بر گرایش‌های کاربران<sup>۸</sup>: در معیار مبتنی بر گرایش‌های کاربران، فراوانی نسبی تمام کلمات کلیدی محاسبه می‌شود. ۲- معیار مبتنی بر شهرت صفحات<sup>۹</sup>: در معیار مبتنی بر شهرت صفحه، تعداد ارجاعات به آن صفحه محاسبه می‌شود یعنی تعداد وب‌سایت‌هایی که به آن صفحه اشاره می‌کنند. ۳- معیار مبتنی بر محل قرار گرفتن صفحات<sup>۱۰</sup>: در معیار مبتنی بر محل قرار گرفتن صفحه منظور از محل قرار گرفتن صفحه آدرس صفحه، ماهیت آدرس از لحاظ [www.com](http://www.com) یا [www.net](http://www.net) یا [www.edu](http://www.edu) و ... است، همچنین میزان فاصله‌ی آن از صفحه‌ی خانگی وب‌سایت است.

دو الگو برای مدل درون‌کاوی چگونگی آغاز و ختم فرآیند استخراج و ذخیره‌سازی صفحات وب کاربرد دارد:

۱- کاوش و توقف<sup>۱۱</sup>: در روش کاوش و توقف، درون‌کاوی پس از ملاقات و دریافت دقیقاً  $k$  صفحه‌ی وب متوقف می‌شود که  $k$  عددی ثابت است. صفحات نیز از پراهمیت به کم‌اهمیت دسته‌بندی می‌شوند.

۲- کاوش و توقف مبتنی بر مقدار آستانه<sup>۱۲</sup>: در روش کاوش و توقف مبتنی بر مقدار آستانه دقیقاً مانند الگوی توقف و کاوش عملیات انجام می‌شود با این تفاوت که صفحاتی دریافت و ذخیره می‌شوند که اهمیت آن‌ها از مقدار آستانه‌ی  $\epsilon$  بیشتر باشد.

#### روش‌های تازه‌سازی و سرکشی به صفحات:

- تازه‌سازی یکنواخت و مدام - تازه‌سازی متناسب با تغییر

شاخص‌های استخراج شده توسط موتورهای جستجو:

- شاخص لینک<sup>۱۳</sup> - شاخص متنی<sup>۱۴</sup>

---

<sup>۶</sup>Structure index

<sup>۷</sup>ranking module

<sup>۸</sup>interest driven

<sup>۹</sup>popularity driven

<sup>۱۰</sup>location driven

<sup>۱۱</sup>crawl and stop

<sup>۱۲</sup>crawl and stop with threshold

در شاخص لینک صفحات موجود در انباره را در قالب یکنواخت جهت دار<sup>۱۵</sup> مدل می شوند. گره ها یا نودهای این گراف، صفحات وب هستند و لبه های گراف ارجاعاتی است که صفحات در قالب hyperlink به یکدیگر دارند. مجموعه ی لینک های به بیرون<sup>۱۶</sup> و به داخل<sup>۱۷</sup> مورد توجه است.

شاخص متنی عبارت است از یک پایگاه اطلاعاتی از کلکلمات کلیدی ممکن در ادبیات هر زبان به همراه مجموعه ی غنی از مخفف ها و استعارات به همراه اندیس صفحاتی که این کلمات در درون آن ها وجود دارند

### ارزش یک صفحه از نظر مایجول رتبه بندی (رتبه بندی اطلاعات مندرج در صفحه):

- دفعات تکرار کلمات کلیدی - ترتیب و مجاورت کلمات کلیدی - محل درج کلمات کلیدی از لحاظ عنوان پاراگراف یا متن معمولی - درج کلمات درون آدرس صفحه در بطن url
- پررنگ بودن کلمات کلیدی - بهره گیری از برچسب های توصیفی<sup>۱۸</sup>
- بهره گیری از برچسب<sup>۱۹</sup> - اطلاعات مندرج در بیرون صفحه - تعداد ارجاعاتی که به صفحه داده شده است
- رده بندی جهانی وب سایت حاوی صفحه از لحاظ طراحی، تعداد بازدیدکننده، جذب ترافیک و ...

بیشترین عوامل رتبه بندی بیرون صفحه تعداد ارجاعات و لینک هایی است که از دیگر صفحات، صفحه ی مورد نظر را نشانه رفته اند.

### الگوریتم های رتبه بندی:

الگوریتم هایی هستند که تصمیم می گیرند که براساس چه کلمات کلیدی، چه وبسایتی در چه صفحه و رده ای از نتایج جستجو قرار بگیرد. الگوریتم های رتبه بندی امروزه بسیار پیچیده هستند و از هزاران پارامتر بهره می برند. در این جا به برخی از مشهورترین پارامترها خواهیم پرداخت. پارامترهای رتبه دهی سه دسته اند:

کلمات (تعداد و موقعیت کلمات) - لینک ها (تعداد و ارجاعات) - آمار کاربران (کلیکی رای کاربر)

**وزن دهی به کلمات:** برای هر کلمه در یک متن یک وزن با الگوی خاصی در نظر گرفته می شود. این وزن بیانگر تاثیر کلمه بر موضوع متن در مقایسه با سایر کلمات به کار رفته است. اهمیت کلمات را می توان بر پایه ی شرایطی مشخص کرد:

- وزن آماری کلمه - مکان قرارگیری کلمه - مفهوم هر کلمه - کاربرد خاص کلمه

<sup>13</sup>link index

<sup>14</sup>text index

<sup>15</sup>directional graph

<sup>16</sup>links outward

<sup>17</sup>incoming link

<sup>18</sup>meta description tag

<sup>19</sup>alt tag

## دانش مورد نیاز برای کار با موتور جستجو:

دانش ذهنی، دانش مورد نیاز برای تبدیل نیاز اطلاعاتی به یک درخواست قابل جستجو می‌باشد.

دانش فنی، مهارت‌های اساسی به‌کارگیری رایانه و ترکیب درخواست‌های وارد شده به‌عنوان عبارت‌های جستجوی خاص می‌باشد.

دانش معنایی، مشخص می‌کند که چگونه و در چه وقتی قابلیت‌های موتور جستجو را باید به‌کار گرفت.

افزایش این سه دانش از طرف کاربر به‌صورت چشم‌گیری باعث افزایش میزان اسناد بازیابی شده می‌شود. در بسیاری از موارد کاربر می‌خواهد چیزی را از موتور جستجو بخواهد که راجع به آن جاهل است به‌همین دلیل رفتار کاربران درحین جستجو تا حد زیادی غیرقابل پیش‌بینی می‌شود.

کار اصلی موتور جستجو سنجش ارتباط اطلاعات ذخیره‌شده و اطلاعات درخواست‌شده است. به‌عبارتی دیگر با ارائه یکسوال، نظام بازیابی باید بررسی کند که آیا اطلاعات ذخیره‌شده مربوط به پرسش است یا نه، و ابهام و استعارات پشت واژگان و نقص بیان مفاهیم با برخی واژگان این ارتباط (ربط) را مشخص‌سازد.

برای برطرف کردن مشکل سوءتفاهم بین ذهن کاربر و الگوریتم‌های موتور جستجو اخیراً الگوریتم‌های پیشرفته‌تری استفاده می‌شود که در آن نظر کاربر به‌عنوان یک پارامتر لحاظ می‌شود. گوگل یکی از موتورهای جستجو است که پیش‌تاز این روش است.

**متا جستجوگرها:** یک موتور جستجوگر سایتی است که به‌طور واسطه بین کاربر و موتورهای جستجو قرار می‌گیرد، پرسش کاربر را دریافت می‌کند و آن را پالایش کرده و با استفاده از سرویس وب موتورهای جستجو، نتایج را از چندین موتور جستجو دریافت و حاصل را ترکیب کرده و به‌کاربر ارائه می‌دهد. استفاده از این روش باعث می‌شود دامنه‌ی جستجو وسیع‌تر شود و نتایج بهتری حاصل شود.

سئو<sup>۲۰</sup> چیست:

---

<sup>۲۰</sup> SEO : Search Engine Optimization

معنی سئو بهینه‌سازی استفاده از موتورهای جستجو است یعنی تولیدکردن صفحاتی که برای موتورهای جستجو جالب هستند. بهینه‌سازی صفحات وب یعنی این که در یک موتور جستجوی بزرگ، بیشترین امتیاز را به دست بیاورد. درواقع یعنی این که چه کارهایی انجام دهیم که وبسایت ما در صفحه‌ی اول نتایج جستجو قرار بگیرد.



نکته: یکی از کارهایی که متخصصان سئو انجام می‌دهند پیگیری تمام تغییرات عملکرد داخل موتورهای جستجو است. بنابراین آن‌ها می‌توانند صفحات وب را بر طبق این تغییرات بهینه‌سازی کنند.

دلیل اهمیت سئو:

- اگر از این علم استفاده نشود بازدیدکنندگان زیادی را از دست خواهیم داد.
- موتورهای جستجو تنها وسیله‌ی شناساندن وبسایت ما در فضای مجازی هستند.
- برتر بودن در موتورهای جستجو مشتریان زیادی را به سوی وبسایت ما دعوت می‌کند.
- ۹۱ درصد کاربران از موتورهای جستجو استفاده می‌کنند.
- ۷۳ درصد معاملات آنلاین از طریق موتورهای جستجو آغاز می‌شود.

**Page Rank:** روش گوگل برای اندازه‌گیری اهمیت صفحات است. زمانی که تمام فاکتورهای دیگر نظیر عنوان صفحه،

تعداد تکرار کلمه در صفحه و ... محاسبه شدند، گوگل از page rank برای تنظیم نتایج جستجو استفاده می‌کند و لذا

وبسایت‌هایی که مهم‌تر هستند به رده‌های بالاتر می‌آیند. [۴]

نحوه رتبه‌بندی در گوگل به این شرح است:

- تمام صفحات مرتبط با عبارت جستجو شده را در بایگانی‌اش می‌یابد.

-صفحات یافته شده را باتوجه به پارامترهای موجود در خود صفحه، رده‌بندی می‌کند.

- وزن‌ها را در متن پیوندهای ورودی صفحه محاسبه می‌کند.

- نتایج جستجو را باتوجه به **page rank** مرتب می‌کند.

**مفهوم page rank:** تئوری گوگل این‌گونه بیان می‌کند که وقتی صفحه‌ی A به صفحه‌ی B پیوند می‌دهد، یعنی این‌که از دیدگاه صفحه‌ی A صفحه‌ی B یک صفحه‌ی مهم است. بنابراین Page Rank اهمیت پیوندهایی را که به صفحه شده است، لحاظ می‌کند. اگر صفحه‌هایی که اهمیت بالایی دارند، به صفحه‌ای پیوند بدهند، آن‌گاه پیوندهای آن صفحه به صفحات دیگر هم اهمیت بیشتری می‌یابند. بنابراین باید گفت که متن پیوند به Page Rank ارتباطی ندارد.

درواقع هرگاه به دنبال اطلاعاتی می‌گردیم با دو سؤال مهم روبه‌رو هستیم:

۱- چگونه جستجو کنیم؟ ۲- کجا جستجو کنیم؟

در صورتی که بتوانیم به این دو سؤال پاسخ مناسب و روشنی بدهیم، به راحتی خواهیم توانست اطلاعات مورد نظر خود را بر روی شبکه بیابیم. اپراتورهای استاندارد هستند که در اکثر سایت‌هایی که امکان جستجو فراهم آورده‌اند، قابل استفاده می‌باشند. **AND:** هنگامی که ما به دنبال صفحاتی می‌باشیم که حاوی چند کلمه یا عبارت معین باشند از این اپراتور استفاده می‌کنیم. به طور مثال زمانی که ما به دنبال اطلاعاتی در مورد اینترنت و امنیت هستیم از عبارت **Internet And Security** استفاده می‌کنیم.

**OR:** زمانی که وجود حداقل یکی از چند کلمه یا عبارات معین در صفحات مورد نظر شما باشد، اپراتور **OR** را در بین این کلمات به کار می‌گیریم. به طور مثال **Internet Or Security**، صفحات سایت‌هایی را در اختیار ما قرار خواهد داد که دارای اطلاعاتی در مورد اینترنت، امنیت یا هر دو باشند.

**+** : این اپراتور به عنوان پیشوند قبل از کلمات مورد نظر به کار می‌رود. درواقع کلماتی که پیش از آن‌ها این اپراتور گذارده شود، لزوماً در نتایج جستجو وجود خواهند داشت. به عنوان مثال: **Internet+Security** صفحات و سایت‌هایی را در نتایج حاصله از جستجو بر می‌گرداند که لزوماً دارای لغت امنیت هستند، اما اینترنت می‌تواند در آن‌ها وجود داشته یا نداشته باشد.

**-** : این اپراتور که همانند اپراتور قبلی به صورت پیشوند به کار می‌رود، عدم وجود کلمه‌ی پسین خود را از سایت جستجو می‌خواهد. یعنی نتایج حاصل از جستجو لزوماً کلمه‌ی پس از این اپراتور را نخواهد داشت. به طور مثال: **Internet-**

Security صفحاتی را در اختیار شما قرار می‌دهد که دارای لغت اینترنت باشد، اما نامی از امنیت در آن صفحات ذکر نشده باشد.

«: هرگاه به دنبال یک عبارت هستیم و می‌خواهیم کلمات به همان شکل و ترتیب در متن سایت‌های نتایج جستجو یافت شوند، از این اپراتور استفاده می‌نماییم. مثلاً زمانی که به دنبال اطلاعاتی در مورد امنیت شبکه‌های کامپیوتری هستیم باید عبارت ((Network Security)) را وارد نماییم.

### بخش‌های مجزای یک موتور جستجوگر:

- عنکبوت<sup>۲۱</sup> - خزنده<sup>۲۲</sup> - بایگانی‌کننده<sup>۲۳</sup> - پایگاه داده<sup>۲۴</sup> - سیستم رتبه‌بندی<sup>۲۵</sup>

عنکبوت: اسپایدر یا روبات<sup>۲۶</sup>، نرم‌افزاری است که کار جمع‌آوری اطلاعات مورد نیاز برای یک موتور جستجوگر را برعهده دارد. اسپایدر به صفحات مختلف سر می‌زند، محتوای آن‌ها را می‌خواند، لینک‌ها را دنبال می‌کند، اطلاعات مورد نیاز را جمع‌آوری می‌کند و آن‌را در اختیار سایر بخش‌های موتور جستجوگر قرار می‌دهد. کار یک اسپایدر، بسیار شبیه به کار کاربران وب است. همان‌طور که کاربران، صفحات مختلف را بازدید می‌کنند، اسپایدر هم درست این کار را انجام می‌دهد با این تفاوت که اسپایدر کدهای HTML صفحات را می‌بیند اما کاربران نتیجه‌ی حاصل از کنار هم قرارگرفتن این کدها را مشاهده می‌کنند. index.html صفحه‌ای است که کاربران آن‌را می‌بینند. [۶]

نحوه‌ی کار اسپایدر: برای این که شما هم بتوانید دنیای وب را از دیدگاه یک اسپایدر ببینید، کافی است که کدهای HTML صفحات را مشاهده کنید.

آیا این دنیای متنی برای شما جذاب است؟ اسپایدر، به‌هنگام مشاهده‌ی صفحات، بر روی سرورها رد پا برجای می‌گذارد.

شما اگر اجازه‌ی دسترسی به آمار دید و بازدیدهای صورت گرفته از یکسایت و اتفاقات انجام شده در آن‌را داشته باشید، می‌توانید مشخص کنید که اسپایدر کدام یک از موتورهای جستجوگر صفحات سایت را مورد بازدید قرار داده است.

خزنده: کراولر، نرم‌افزاری است که به‌عنوان یک فرمانده برای اسپایدر عمل می‌کند. کراولر مشخص می‌کند که اسپایدر کدام صفحات را مورد بازدید قرار دهد.

---

<sup>21</sup>Spider  
<sup>22</sup>Crawler  
<sup>23</sup>Indexer  
<sup>24</sup>Database  
<sup>25</sup>Ranker  
<sup>26</sup>Robot

درواقع کراولر تصمیم می‌گیرد که کدام یک از لینک‌های صفحه‌ای که اسپایدر در حال حاضر در آن قرار دارد، دنبال شود. ممکن است همه‌ی آن‌ها را دنبال کند یا بعضی‌ها را دنبال کند و یا هیچ کدام را دنبال نکند.

در این صورت حتی ممکن است که کراولر، قبلاً برنامه‌ریزی شده باشد که آدرس‌های خاصی را طبق برنامه، در اختیار اسپایدر قرار دهد تا از آن‌ها دیدن کند. دنبال کردن لینک‌های یک صفحه به این بستگی دارد که موتور جستجوگر چه حجمی از اطلاعات یک سایت را می‌خواهد در پایگاه داده ذخیره کند. همچنین ممکن است اجازه‌ی دسترسی به بعضی از صفحات به موتورهای جستجوگر داده نشده باشد. شما به عنوان دارنده‌ی سایت، همان‌طور که دوست دارید موتورهای جستجوگر اطلاعات سایت شما را با خود ببرند، می‌توانید آن‌ها را از بعضی صفحات سایت خود دور کنید و اجازه‌ی دسترسی به محتوای آن صفحات را به آن‌ها ندهید.

**بایگانی کننده:** تمام اطلاعات جمع‌آوری شده توسط اسپایدر در اختیار ایندکسر قرار می‌گیرد. در این بخش اطلاعات ارسالی مورد تجزیه و تحلیل قرار می‌گیرند و به بخش‌های متفاوتی تقسیم می‌شوند. تجزیه و تحلیل بدین معنی است که مشخص می‌شود اطلاعات فوق از کدام صفحه ارسال شده است، چه حجمی دارد و کلمات موجود در آن کدامند، کلمات چندبار تکرار شده‌اند، کلمات در کجای صفحه قرار دارند و ...

در حقیقت ایندکسر، صفحه را به پارامترهای آن خرد می‌کند و تمام این پارامترها را به یک مقیاس عددی تبدیل می‌کند تا سیستم رتبه‌بندی بتواند پارامترهای صفحات مختلف را با یکدیگر مقایسه کند. در زمان تجزیه و تحلیل اطلاعات، ایندکسر برای کاهش حجم داده‌ها از بعضی کلمات که بسیار رایج هستند صرف نظر می‌کند. کلماتی نظیر a, an, the, www, is و ... از این گونه کلمات هستند.

**پایگاه داده:** تمام داده‌های تجزیه و تحلیل شده در ایندکسر، به پایگاه داده ارسال می‌گردند. که در این بخش، داده‌ها گروه بندی، کدگذاری و ذخیره می‌شود. همچنین داده‌ها قبل از آن که ذخیره شوند، طبق تکنیک‌های خاصی فشرده شده تا حجم کمتری از پایگاه داده را اشغال کنند. یک موتور جستجوگر باید پایگاه داده عظیمی داشته باشد و به‌طور مداوم حجم محتوای آن را گسترش دهد و البته اطلاعات قدیمی را هم به‌روزرسانی نماید. بزرگی و به‌روز بودن پایگاه داده‌ی یک موتور جستجوگر برای آن امتیاز محسوب می‌گردد. یکی از تفاوت‌های اصلی موتورهای جستجوگر در حجم پایگاه داده‌ی آن‌ها و همچنین روش ذخیره‌سازی داده‌ها در پایگاه داده است.



بعد از آن که تمام مراحل قبل انجام شد، موتور جستجوگر برای پاسخگویی به درخواست کاربر، ابتدا تمام صفحات موجود در پایگاه داده که به موضوع جستجو شده، مرتبطاند، مشخص می‌شوند. پس از آن سیستم رتبه‌بندی نیز وارد عمل شده و آن‌ها را از بیشترین ارتباط تا کمترین ارتباط مرتب می‌کند و به عنوان نتایج جستجو به کاربر نمایش می‌دهد.

حتی اگر موتور جستجوگر بهترین و کامل‌ترین پایگاه داده را هم داشته باشد اما نتواند پاسخ‌های مرتبطی را ارایه کند، قطعاً یک موتور جستجوگر ضعیف خواهد بود. درحقیقت سیستم رتبه‌بندی قلب تپنده در یک موتور جستجوگر است و تفاوت اصلی موتورهای جستجوگر هم در این بخش قرار دارد. سیستم رتبه‌بندی برای پاسخ‌گویی به سوالات کاربران، پارامترهای بسیاری را در نظر می‌گیرد تا بتواند بهترین پاسخ‌ها را در اختیار آن‌ها قرار دهد. [۶]

در حال حاضر قدرتمندترین سیستم رتبه‌بندی را گوگل در اختیار دارد.

نکته: برای سهولت در بیان مطالب هرگاه صحبت از بایگانی کردن میان می‌آید، مقصود این است که صفحه تجزیه و تحلیل شده و به پایگاه داده‌ی موتور جستجوگر وارد می‌شود.

### داستان شکارچی:

برای آن که تصور درستی از نحوه‌ی کار یک موتور جستجوگر داشته باشیم به بیان داستان زیر می‌پردازیم. داستان ما یک شکارچی دارد که تصمیم به شکار می‌گیرد:

**کار کراولر:** او قصد دارد برای شکار به منطقه‌ی حفاظت شده‌ای در خراسان برود.

**پروتکل Robots:** ابتدا تمام محدودیت‌های موجود برای شکار در این منطقه را بررسی می‌کند.

آیا در این منطقه می‌توان به شکار پرداخت؟ کدام حیوانات را می‌توان شکار کرد؟ حداکثر تعداد شکار چه میزانی است؟

فرض می‌کنیم او مجوز شکار یک آه‌ورا از شکاربانی منطقه دریافت می‌کند.

**کار اسپایدر:** او آه‌ورا را شکار می‌کند و سپس آن را با خود به منزل می‌برد.

**کار ایندکسر:** شکار را تکه تکه کرده، گوشت، پوست، استخوان، و سایر قیمت‌های آه‌ورا فریز کرده و بخش‌های زاید شکار را دور می‌ریزد.

**کار پایگاه داده:** بسته‌های حاصل را درون فریزر قرار داده و ذخیره می‌کند.

کار سیستم رتبه‌بندی: مهمانان سراغ او می‌آیند و او بسته به سلیقه‌ی مهمانان برای آن‌ها غذا طبخ می‌کند. ممکن است عده‌ای کله‌پاچه، عده‌ای آبگوشت و عده‌ای غذاهای دیگری را دوست داشته باشند.

پخت‌و‌پز طبق سلیقه‌ی مهمانان کار سختی است. البته ممکن است همه‌ی آن‌ها آبگوشت بخواهند اما آن‌ها مسلماً خوشمزه‌ترین آبگوشت را می‌خواهند.

در ادامه با مقایسه‌ی مثال شکارچی با نحوه‌ی کار موتورهای جستجو و همچنین قسمت‌های مختلف آن به بیان نکته‌هایی می‌پردازیم که به نوعی جمع‌بندی حاصل از این مباحث می‌باشد.

#### نکته‌ها:

- شکارچی می‌توانست برای شکار کبک یا آهو و یا هر دو، به آن منطقه برود همان‌طور که موتور جستجویی تواند از سرور سایت شما انواع فایل‌های عکس، فایل متنی، فایل اجرایی و ... را درخواست کند.
- شکارچی می‌تواند شب به شکار برود یا روز. موتور جستجوگر هم ممکن است شب به سایت شما مراجعه کند یا روز. بنابراین همواره باید مطمئن باشیم که سایت ما به روز است و موتور جستجوگر می‌تواند در آن به شکار فایل‌ها بپردازد.
- غذای خوشمزه را می‌توانیم با نتایج جستجوی دقیق و مرتبط مقایسه کنیم. اگر شکارچی بهترین شکار را با خود به منزل ببرد اما غذایی خوشمزه و مطابق سلیقه‌ی مهمانان طبخ نکرده، تمام زحمات هدر رفته است.