عنوان: Full Text Search و Rank فیلدهای بازیابی شده

نویسنده: دل محسن

تاریخ: ۰:۵۵ ۱۳۹۲/۰۳/۱۰

آدرس: www.dotnettips.info

گروهها: SQL Server, Full Text Search

حالتی را در نظر بگیرید که بخواهید تعداد زیادی رکورد را که حجم هر رکورد هم قابل ملاحظه هست، نگهداری کنید(مثلا چندین هزار مقاله) و همچنین قابلیت جستجو را در این رکوردها لحاظ کنید به صورتی که بر اساس رکوردهایی که بیشترین تعداد تکرار کلمات مدنظر را دارند مرتب شوند.

شاید اولین راه حل، مطلب آقای سلیم آبادی در اینجا باشه، که تعداد تکرار یک کلمه را در فیلدی در جدول بیان کردند و درست هم هست اما با 2 شرط:

- 1) رکوردهای ما حجم کمی داشته باشند چرا که دستور LIKE پاسخ سریعی را با حجم بالای اطلاعات به ما نمیدهد.
 - 2) رکوردهای ما از خانوادهی char باشند. برای مثال اگر مقالات ما به صورت pdf باشند این کد جواب نمیدهد.

اما راه دوم استفاده از <u>Ful Text Search</u> و دستور CONTAINSTABLE هست که 2 شرط لازم را برای راه حل اول احتیاج نداره. در اینجا فقط نحوهی استفاده از CONTAINSTABLE رو مطرح میکنیم.

CONTAINSTABLE جدولی از موارد یافت شده را بر اساس معیارهایی که ما به اون معرفی میکنیم، ایجاد میکند. این جدول حاوی دو فیلد KEY(کلید فیلد مورد نظر) و RANK(مقداری بین 0 تا 1000) است که میزان همسانی رکورد با معیار ما را مشخص میکند و ما با استفاده از این فیلد میتوانیم رکوردهایمان را مرتب کنیم.

به این کد توجه کنید:

```
SELECT t.Title, p.[RANK]
   FROM Articles AS t
        INNER JOIN CONTAINSTABLE(Articles, Data, 'management' ) AS p
        ON t.Id = p.[KEY]
ORDER BY p.RANK
```

در اینجا کار جستجو انجام شده و بر حسب میزان نزدیکی محتویات رکورد با معیار ما مرتب شده است.

نکته: هیچ فرقی نمیکند که محتویات فیلد مورد نظر شما یک متن ساده، یک فایل word یا حتی pdf باشد. فقط باید تنظیمات Full Text Search درست انجام شود.

نظرات خوانندگان

نویسنده: محسن خان

تاریخ: ۲۲:۴۹ ۱۳۹۲/۰۳/۱۰

برای PDF که عنوان کردید، آیا جستجوی فارسی کار میکند؟ از چه IFilter ایی برای اینکار استفاده کردید؟

نویسنده: دل محسن

تاریخ: ۱۴:۵۰ ۱۳۹۲/۰۳/۱۱

بله کار میکنه. من با Adob IFilter(رایگان) و Foxit IFilter(تجاری هست) کار کردم.

نویسنده: علی

تاریخ: ۶/۰۶/۱۳۹۲ ۱۲:۲۲

سلام.

من با rank کار میکنم توی پروژه هام.

منتها خیلی وقتها رکوردهایی رنک بالایی میگیرند در حالی که رکوردهای مشابه دیگر رنک پایینی میگیرند. البته کمتر از 5 درصد اتفاق میافته

در مورد خود رنک مطلب خوبی پیدا نکردم که اصلا خود مفهوم رنک رو برام توضیح بده

نویسنده: علی

تاریخ: ۲۴/۱۰/۲۴ ۱۹:۲۷

منم توی این رنک بدجور گیر کردم. به جابجایی کامات حساسه . مثلا «خواجه علی» و «علی خواجه ».