

# فاز سوم پروژه بازیابی اطلاعات

نام و نام خانوادگی: شیرین عبادی  
شماره دانشجویی: ۹۷۳۴۰۱۶

1.

از نظر پیاده‌سازی استفاده از **bulk** بسیار آسان‌تر از پیاده‌سازی فاز یک است. با استفاده از **api** می‌توان بسیاری از پیچیدگی‌های مربوط به ساخت شاخص را به سرور الاستیک سپرد، و در این بین تنها داکيومنت‌ها فرستاده می‌شوند. اما در فاز یک تمامی مراحل ساخت ایندکس از توکنایز کردن تا **append** کردن داکيومنت‌ها بر عهده توسعه‌دهنده است و بنابراین کار دشوارتر است.

از نظر زمانی نیز **bulk** مدت زمان **5.13** ثانیه و ایندکسینگ فاز یک مدت زمان **30.47** ثانیه زمان برده است.

۲.

الف) پرسمان دشوار : « "هواداران باشگاه" عراق !ایران»

نتایج حاصل از الستیک سرچ:

[نتیجه ۱](#)

[نتیجه ۲](#)

[نتیجه ۳](#)

نتایج حاصل از فاز یک:

[نتیجه ۱](#)

[نتیجه ۲](#)

[نتیجه ۳](#)

نتیجه اول و دوم هر دو یکی است اما نتیجه سوم موتور الستیک مرتبطتر رفتار کرده است.

ب) پرسمان نادر : «اورشلیم !صهیونیست»

هیچ یک از موتور ها سندی بازنگرداند!

ب) پرسمان نادر : «صهیونیست !اورشلیم»

نتایج حاصل از الستیک سرچ:

[نتیجه ۱](#)

[نتیجه ۲](#)

[نتیجه ۳](#)

نتایج حاصل از فاز یک:

[نتیجه ۱](#)

[نتیجه ۲](#)

[نتیجه ۳](#)

نتایج حاصل از هر دو موتور یکسانند!

به طور کلی می توان نتیجه گرفت که دقت بازیابی هردو موتور تقریباً نزدیک به هم است. البته الویت هرکدام در صورتی که یکی از موارد نباشد متفاوت است. (مانند پرسمان الف) اما به طور کلی رفتار هر دو نزدیک به هم است. از نظر سرعت موتور الستیک بهتر عمل می کند. البته این وابسته به مواردی همچون سرعت اینترنت و .. نیز هست. از طرفی

فاز یک روی یک سیستم لوکال اجرا می‌شود اما الستیک روی سرور الستیک که می‌توان گفت سرعت بالاتری دارد. در چنین شرایطی سرعت بازیابی توسط الستیک بالاتر است.

۳.

ترجیح بنده استفاده از موتور الستیک است چرا که بدلیل عدم دخالت انسان در ایندکسینگ آن احتمال خطا در آن کمتر است. بعلاوه سرعت بالاتری نیز دارد.

### 'لیک برتر فوٹبال',

```
{'text': 'لیک برتر فوٹبال', 'score': 3.684386e-06},  
{'text': 'لیک برتر فوٹسال', 'score': 1.1221841e-06},  
{'text': 'لیک برتر فوٹبال', 'score': 7.1267203e-07},  
{'text': 'لیک برتر فوٹبالی', 'score': 6.989063e-07},  
{'text': 'لیک برتر فوٹال', 'score': 9.195401e-08}}]
```

=====

### 'ٹورنمنت شش جانبہ',

```
{'text': 'ٹورنمنت شش جانبہ', 'score': 2.214439e-06},  
{'text': 'ٹورنمنت شش جانبی', 'score': 9.388285e-07},  
{'text': 'ٹورنمنت شش جاذبہ', 'score': 6.737793e-07},  
{'text': 'ٹورنمنتی شش جانبہ', 'score': 6.628849e-07},  
{'text': 'ٹورنمنت شش جانبہ', 'score': 6.144986e-07}}]
```

=====

### 'طبیعی نژادی',

```
{'text': 'طبیعی نهادی', 'score': 0.000113283226},  
{'text': 'طبیعی نژاد', 'score': 8.068119e-05},  
{'text': 'طبیعی نمادی', 'score': 6.851754e-05},  
{'text': 'طبیعی نقادی', 'score': 5.464236e-05},  
{'text': 'طبیعی نژادی', 'score': 5.176822e-05}}]
```

=====

### 'اردوی طیم امید',

```
{'text': 'اردوی طیم امید', 'score': 5.4468444e-07},  
{'text': 'اردوی طیم امیر', 'score': 4.346943e-07},  
{'text': 'اردوی طیم امین', 'score': 2.7609326e-07},  
{'text': 'اردوی طیم امی', 'score': 3.8065412e-08},  
{'text': 'اردوی طیم امیل', 'score': 3.8065412e-08}}]
```

=====

### 'جام ملب های آسٹیا',

```
{'text': 'جام ملب های آسٹیا', 'score': 1.25268e-07},  
{'text': 'جام ملب های آسٹنا', 'score': 4.4106642e-08},  
{'text': 'جام ملب های آسٹیل', 'score': 1.7457772e-08},  
{'text': 'جام ملب های آریا', 'score': 1.26198865e-08},  
{'text': 'جام ملب های آلیا', 'score': 9.906219e-09}}]
```

=====

### 'کنزہ سیاسی آمریکا',

```
{'text': 'کنگرہ سیاسی آمریکا', 'score': 1.8001992e-05},  
{'text': 'کارہ سیاسی آمریکا', 'score': 1.0186398e-05},  
{'text': 'کنارہ سیاسی آمریکا', 'score': 9.621039e-06},  
{'text': 'کورہ سیاسی آمریکا', 'score': 7.972548e-06},
```

{'text': 'کنده سیاسی آمریکا', 'score': 6.399634e-06}}  
=====

#### 'انقلاب اسلامی ایران',

{'text': 'انقلاب اسلامی ایران', 'score': 0.00029608927},  
{'text': 'انتخاب اسلامی ایران', 'score': 0.00014258877},  
{'text': 'انقلابی اسلامی ایران', 'score': 0.00012659801},  
{'text': 'انقلاب، اسلامی ایران', 'score': 6.851434e-05},  
{'text': 'انتصاب اسلامی ایران', 'score': 5.142784e-05}}  
=====

#### 'فدراسیون فوتبال ایران',

{'text': 'فدراسیون فوتبال ایران', 'score': 0.00021502444},  
{'text': 'فدراسیون، فوتبال ایران', 'score': 3.576952e-05},  
{'text': 'فدراسیون فوتبال ایران', 'score': 2.5018382e-05},  
{'text': 'فدراسیونی فوتبال ایران', 'score': 2.3344137e-05},  
{'text': 'فدراسیون فوتبال ایران', 'score': 1.2801158e-05}}  
=====

#### 'لایحه مجلس خبرگان',

{'text': 'لایحه مجلس خبرگان', 'score': 3.207954e-05},  
{'text': 'لایحه مالی خبرگان', 'score': 1.1077829e-05},  
{'text': 'لایحه مجلس خبرگان', 'score': 6.8262448e-06},  
{'text': 'لایحه مخلص خبرگان', 'score': 2.4110616e-06},  
{'text': 'لایحه مالی خبرگان', 'score': 2.3572647e-06}}  
=====

#### 'نخبگان دانشگاهی',

{'text': 'نخبگان دانشگاهی', 'score': 0.00039060856},  
{'text': 'ناوگان دانشگاهی', 'score': 0.00016786794},  
{'text': 'نخبگانی دانشگاهی', 'score': 0.00015978346},  
{'text': 'نخبگان، دانشگاهی', 'score': 0.00010711628},  
{'text': 'نگهبان دانشگاهی', 'score': 4.2403135e-05}}  
=====

#### 'نمایند مجل',

{'text': 'نماینده مجل', 'score': 4.8595244e-05},  
{'text': 'نمایند. مجل', 'score': 1.2255474e-05},  
{'text': 'نمایند مجل', 'score': 8.531589e-06},  
{'text': 'نمایند، مجل', 'score': 3.955441e-06}}  
=====

#### 'فضاسازی کازب',

{'text': 'فضاسازی کادر', 'score': 0.00016055968},  
{'text': 'فعالسازی کادر', 'score': 5.9505375e-05},  
{'text': 'فضاهایی کادر', 'score': 5.067181e-05},

```
{'text': 'فضاسازی کاذب', 'score': 4.0450068e-05},  
{'text': 'فضاسازی، کادر', 'score': 3.023273e-05}}]
```

=====

**صلاح ایرادات شورای تگهپان**,

```
{'text': 'صلاح ایرادات شورای تهران', 'score': 3.4275183e-07},  
{'text': 'صالح ایرادات شورای تهران', 'score': 2.8467574e-07},  
{'text': 'صلاح ایراداتی شورای تهران', 'score': 2.069271e-07},  
{'text': 'صلاح ایرادهای شورای تهران', 'score': 1.1226785e-07},  
{'text': 'صلال ایرادات شورای تهران', 'score': 1.11945326e-07}}]
```

=====

**تهریف های تاریخی**,

```
{'text': 'تاریخ های تاریخی', 'score': 3.9336683e-05},  
{'text': 'تبریک های تاریخی', 'score': 2.9583667e-05},  
{'text': 'تعریف های تاریخی', 'score': 2.7400572e-05},  
{'text': 'تحریف های تاریخی', 'score': 1.638109e-05},  
{'text': 'تشریف های تاریخی', 'score': 7.764495e-06}}]
```

=====

**قلیت های دینی**,

```
{'text': 'قلی های دینی', 'score': 9.913008e-06},  
{'text': 'قلیل های دینی', 'score': 3.9711854e-06},  
{'text': 'قلیچ های دینی', 'score': 3.6520464e-06},  
{'text': 'قمیت های دینی', 'score': 3.2916585e-06},  
{'text': 'قلت های دینی', 'score': 2.2310808e-06}}]
```

=====

**تدارکان لازم**,

```
{'text': 'تدارکاتی لازم', 'score': 0.00034567664},  
{'text': 'تدارکات لازم', 'score': 0.00018709387},  
{'text': 'تدارکات، لازم', 'score': 7.3267685e-05},  
{'text': 'تدارکی لازم', 'score': 7.1104914e-05},  
{'text': 'تدارکات\ لازم', 'score': 5.226709e-05}}]
```

=====

پس از استفاده از توکن‌های وارونه مشاهده شد که نتیجه نهایی دچار بهبود شد. به‌طور مثال در مورد صلاح ایرادات شورای نگهبان، صلاح به اصلاً تغییر یافت. اما همچنان اشکالاتی بر آن وارد است.



الف) پرسمان دشوار و کم تکرار تک کلمه‌ای: «فایزر»

الستیکسرچ:

نتیجه ۱

نتیجه ۲

فاز دو:

نتیجه ۱

نتیجه ۲

نتیجه هردو یکسان است.

ب) پرسمان دشوار و کم تکرار چندکلمه‌ای: «طرح منع تردد»

الستیکسرچ:

نتیجه ۱

نتیجه ۲

نتیجه ۳

فاز دو:

نتیجه ۱

نتیجه ۲

نتیجه ۳

در این مرحله نتیجه سوم با هم تفاوت دارد که نتیجه حاصل از فاز دو مرتبطتر است. در مورد سرعت عمل نیز، سرعت الاستیکسرچ به مراتب بهتر از سرعت فاز دو است.