

# Sketch Engine — Infrastructure, Support and Customer Services

Jan Bušta



[jan.busta@sketchengine.co.uk](mailto:jan.busta@sketchengine.co.uk)

5<sup>th</sup> Sketch Engine Workshop  
Bolzano, July 14, 2014

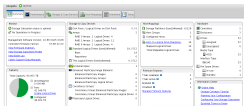
## Meet users' requirements.

Simple facts:

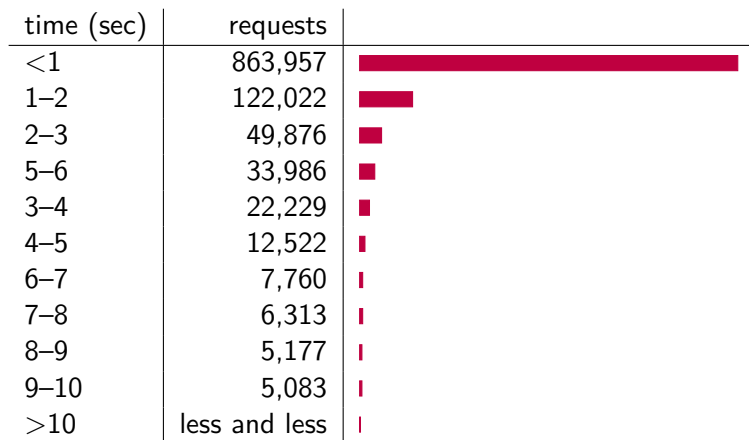
- **Thousands** of users.
- About **3 TB** of data provided to users per month (two times more every 6 months).
- **2,315,959 requests** a month.
- **240** preloaded corpora.
- **8,583** users' corpora.
- Handling more than **100 TB** of data.
- Availability at **99,83%**.
- **All requirements are fast growing...**

# Servers

- 5 IBM Blade servers for High-Availability Clusters.
  - Remote administration.
  - Preventing data corruption.
  - Automatic recovery.
- Main services, beta, API, bug tracking, long job and corpora development server.
- Storage system with hard disk replacement without stopping services.
- Online monitoring system with real-time notification.
- Up-to-date software to keep your data protected and available.

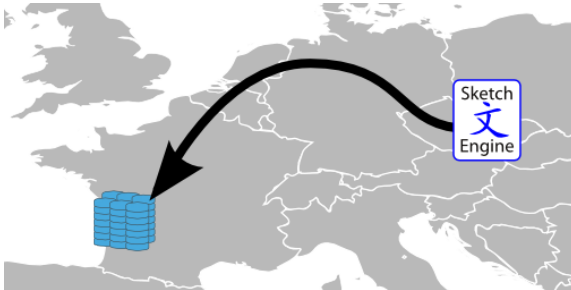


# Request processing



# Backup

- Backup twice a week.
- Geographically remote location of backup server.
- Ready to provide the services in 5 minutes after the request.
- Providing services during primary servers maintenance.

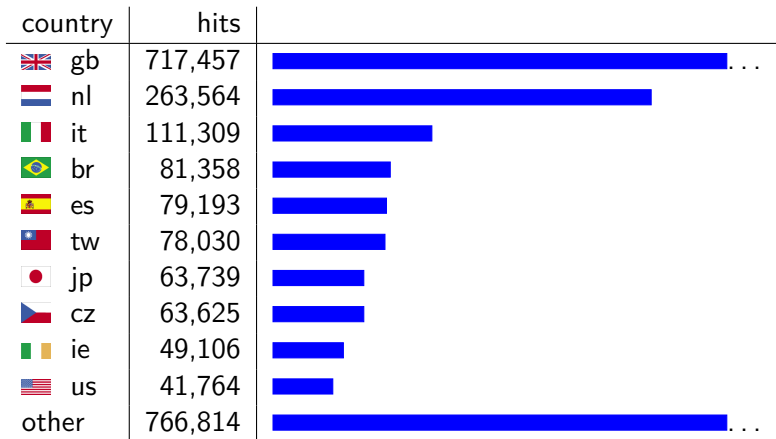


# Local Installations











- Solution for ambitious customers and big companies.
  - Servers & data are placed in area of customer's wish on rent or customer's servers.
  - Meets legal limitation of data access.
  - Sketch Engine can be customized for specific purposes.



## Usage ordered by country (April 2014)



# Corpora querying statistics (April 2014)

corpus	language	queries	
bawe	English	203,307	
esTenTen	Spanish	110,001	
itTenTen	Italian	98,397	
base	English	64,709	
BNC	English	62,594	
aclarc	English	51,472	
enTenTen (2012)	English	45,430	
frTenTen	French	16,474	
estonianRC	Estonian	15,450	
eseuTenTen (2011)	Spanish	13,733	

# Support Service

- Bug tracking system. (Preferred.)
  - ⇒ **62%** of bugs are fixed during the same day.
  - ⇒ **94%** solved in the first week after reporting.
- User support via e-mail. (Less reliable.)
- Documentation system (readings, examples, tutorials)  
<http://www.sketchengine.co.uk/documentation>
- One-off data processing and development requests.

Sketch Engine infrastructure  
is always growing  
to provide a fine service.

Thank you!

Thank you for your attention!