Architecture of a Search Engine (Karaoke Version)
Architecture of a Search Engine (Karaoke Version)

Indexing Process

Data Storage

Retrieval Process
Architecture of a Search Engine (Karaoke Version)

Indexing Process

- Acquisition
- conversion to plain text, and unified encoding

Data Storage

- Document store
- Index

Retrieval Process
Architecture of a Search Engine (Karaoke Version)

Indexing Process

- Acquisition
  - conversion to plain text, and unified encoding

- Text analysis
  - index terms, features, classification, meta data

Data Storage

- Document store
- Index

Retrieval Process
Architecture of a Search Engine (Karaoke Version)

Indexing Process
- **Acquisition**
  - conversion to plain text, and unified encoding
- **Text analysis**
  - index terms, features, classification, meta data
- **Indexing**
  - document statistics, model formation, index update, distribution

**Data Storage**
- **Document store**
- **Index**
- **Stats**

**Retrieval Process**
Architecture of a Search Engine (Karaoke Version)

Indexing Process

Acquisition
- conversion to plain text, and unified encoding

Text analysis
- index terms, features, classification, metadata

Indexing
- document statistics, model formation, index update, distribution

Data Storage

Document store

Index

Stats

Bulk indexing

User Interface

Retrieval Process

query
Architecture of a Search Engine (Karaoke Version)

Indexing Process

- Acquisition
  - conversion to plain text, and unified encoding
- Text analysis
  - index terms, features, classification, meta data
- Indexing
  - document statistics, model formation, index update, distribution

Data Storage

- Document store
- Index
- Stats

User Interface

- User query

Retrieval Process

- Retrieval
  - en.wikipedia.org › wiki › Query
  - www.dictionary.com › browse › query
  - Query - Wikipedia
  - In general, a query is a form of questioning, in a line of inquiry.
  - Query | Dictionary.com
  - to ask or inquire about: No one queried his presence. to question...
Acquisition
- conversion to plain text, and unified encoding

Text analysis
- index terms, features, classification, meta data

Indexing
- document statistics, model formation, index update, distribution

Indexing Process

Document store

Index

Stats

Bulk indexing

Data Storage

Query analysis

User Interface

Retrieval Process
Indexing Process

Acquisition
- conversion to plain text, and unified encoding

Text analysis
- index terms, features, classification, meta data

Indexing
- document statistics, model formation, index update, distribution

Data Storage

Log
- queries, clicks, and users

Document store

Index

Stats

Bulk indexing

Retrieval Process

Query synthesis

Query analysis
- Query
  - query meaning
  - query definition
  - query string
  - query synonym

User Interface

Retrieval

en.wikipedia.org › wiki › Query
Query - Wikipedia
In general, a query is a form of questioning, in a line of inquiry.

www.dictionary.com › browse › query
Query | Dictionary.com
to ask or inquire about: No one queried his presence. to question ...
Architecture of a Search Engine (Karaoke Version)

Indexing Process

Acquisition
- conversion to plain text, and unified encoding

Text analysis
- index terms, features, classification, meta data

Indexing
- document statistics, model formation, index update, distribution

Data Storage

Log
- queries, clicks, and users

Document store

Index

Stats

Evaluation

User Interface

Retrieval Process

Retrieval
- d1 7.9
d2 6.8
d3 6.2

User Interface

Query
- query meaning
query definition
query string
query synonym

Retrieval

Queries, clicks, and users

Bulk indexing

IR:II-82 Introduction © HAGEN/POTTHAST/STEIN 2022