

# Internet of things (IOT) and information retrieval: an introduction

Mohammed Rahmat Ali

Researcher Scholar, Hyderabad-500008, Telangana INDIA

**Abstract:** *Internet of Things a real world concepts which is acquiring the world with its application, now in some recent years has the rapid development in Internet Technology (the Forum) which makes smart homes a reality to the real world as expected by the public convince. Smart home system or also known as smart home technology creates a more comfortable, safe, human oriented and intelligent environment which solves the problems faced by those who have a busy schedule and gets some time to spend at home and grow all over the world. To solve this problem, users can rely on automated machines and gadgets such as smart cards smart phones and some smart devices. These gadgets are connected to internet and uses cloud computing to send and receive signals or information from the cloud. The data we use can be obtained by matching some key values using the concept of information retrieval. The main component of a search engine is a web crawler that has the task of organizing websites and a retrieval system that have the task of retrieving text documents files folders and other audio video files that respond to user requests.*

## I. INTRODUCTION

Internet of Things (IoT) is a new face to the world for a better, smart and secure living in terms of technological advancements. Internet of Things (IoT) is a smart device which carries out many important tasks which a human by itself is not capable of doing it all alone and when it is connected to a internet network, makes it intelligent enough of performing any given task with a given part of time. When and IOT device is connected to internet it gives the access to the network to send and receive the information to and from the network. Usually network will be connected to the cloud, due to its feature of versatile.

Data generated by IOT devices connected to the Internet is enormous, according to an assumption according to the data available network and memory storage, by the end of 2020 each car will generate two petabyte of data each year, and one aircraft will generate forty terabytes of data every day. The resulting data cannot be transmitted instantly within the network, but with easy memory devices and data storage. Moreover, the data later is uploaded externally into the cloud data warehouse.

### WHAT IS INTERNET OF THINGS?

Internet of Things (IoT) collects sensitive data and information on the Internet and circulate among the network and also it allows the devices to interconnect and exchange the data with one another and brings user information in a more secure way. Inter of things connects human to human interaction, human to things

interaction, and things to things interaction to collect the data. Internet connects traditional devices to multiple devices to interact with external environments over the Internet. There is no limit to interconnect and exchange the data on IoT. It enables physical objects to communicate with the Internet and to be interactive, acceptable and essential to accessing and transmitting information.

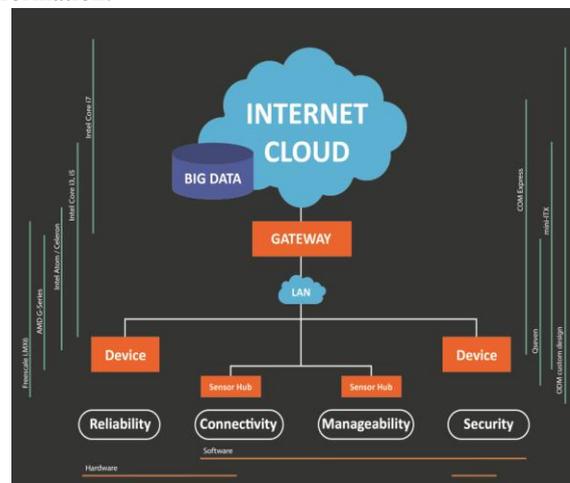


Fig 1: Architecture of IoT

### WHAT IS INFORMATION RETRIEVAL SYSTEM

Information Retrieval is referred as, the search for information based on some specific user information requirements. Information Retrieval is defined as information and scientific research in documents, or the documents describing metadata or databases, whether they are independent or contact databases or hypertext network databases on Internet or intranet. This term Information Retrieval is used to refer to searching for text, sound, images, video, and data. The IR system consists of data sets or information, one or more indexes, the query interface, the search system, and the results interface.

An example of re-application of classic information is traditionally focused on finding ways and mechanisms to find information in files and metadata on files, as well as in databases, on the Internet, and other resources. Information retrieval is often described as the process of identifying a set of documents on a specific topic or meeting certain questions. It is important to mention that infrared is trying to find available information and not to search for new knowledge as it is in extracting data.

The information retrieval process includes the following recurring steps:

1. Definition of information requirements and formal announcement of consultation on information.
2. Identify all possible database information (source recovery).
3. Extract information from specific data sources
4. Know the extracted information and evaluate the recovery results.

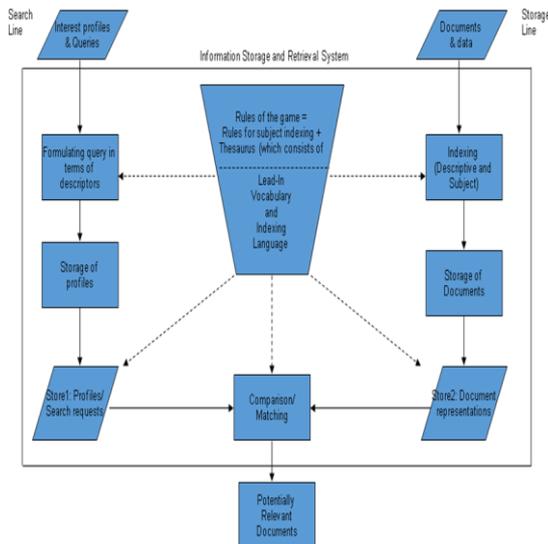


Fig 2: Architecture of Information retrieval System

## II. QUERY PROCESSING IN IRS

Query processing in Information retrieval system constitute of a step of steps. Every step of the processing reduces the complexity of the query and increases the readability of the query. The following are the steps in query processing in Information Retrieval.

- User requirements are defined through the user interface, in the form of text requests
- Questions are analyzed and changed through a set of text operations. The same process was previously used for content indexed by the infrared system
- The following query changes the previously processed request to a system-wide representation.
- This question is run at the top of the document source for a collection of relevant documents. The query may be processed quickly thanks to the previously created index structure of documents in the document source.
- Collections of recoverable documents are arranged: Documents are classified according to the user-defined budget.
- The user then reviews a set of classified documents for useful information. This can refer to a subset of documents that are certainly interesting and thus provide feedback to the system.

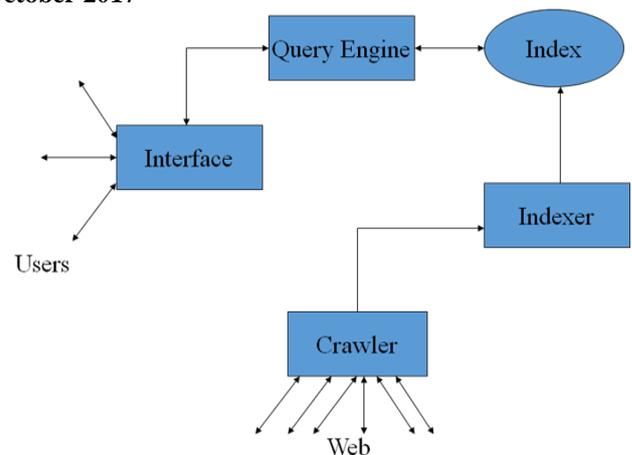


Fig 3: Query Processing in Information retrieval System.

## III. INFORMATION RETRIEVAL SERVICE

Organizations primarily concerned with the provision of online information retrieval services through contracts or payments. Information generally includes multiple topics and is derived from other major sources. The organizations concerned are mainly classified by activities such as credit reporting, direct mail advertising, securities trading services, etc., which also establish databases, according to their core activity. Organizations involved mainly collect, remodel or edit databases from major sources for distribution through information retrieval services classified as.

Database Data Recovery Service

Online information retrieval services.

Online data recovery services.

Remote data retrieval services.

## IV. SEARCHING ON THE INTERNET OF THINGS (IOT)

Entities in the real world (names, places, people, tangible and intangible things) will be important and equal to the current search for important document content in the Web based services. There is also an increase in the importance and value of the information generated from the sensor, status, assets, capabilities, functions and services that can provide things in the real world by communicating them. In Information Retrieval systems uses bot (Process Technologies) is a very complex subject that imposes great demands on design because the information is very dynamic, it is distributed and anticipates some great potential things. Some existing approaches, depending on their scope, are discussed as:

- Only the information contained in the related instruments,
- Things in a personal way,
- Local,
- In major cities, such as smart cities, or
- Global area.

Another difference can be made by rank or potential results as sensor data, sensors only, group of sensors,

objects or physical objects or restrictions based on intelligence space.

#### V. QUALITY OF SERVICES FOR IOT

For measuring the quality of information retrieval and to maintain the quality of information being retrieved and also to improve the quality there are many techniques base on the data and dimension of the data and type of data. The basic two concepts for quality of information retrieval are precision and recall.

##### A. PRECISION

Precision is defined as the information retrieved from the query to the relevance to the data being searched.

$$\text{Precision} = \frac{\text{Number of relevant document retrieved}}{\text{Total number of document retrieved}}$$

##### B. RECALL

Recall is defined as the information that is relevant to the query which is retrieved

$$\text{Recall} = \frac{\text{Number of relevant document retrieved}}{\text{Total number of document relevant.}}$$

##### C. MULTI-THREADED SPIDER

Network delays are common when downloading individual pages. It is best to have many threads running in parallel, each page requesting different hosts. Word addresses can be divided to ensure that the same requests are distributed among different hosts, to maximize usage and avoid overloading from any server. For example, Google's first spider contains many tracking tools that are compatible with about 300 themes each, and can download more than 100 pages per second.

##### D. FOCUSED SPIDER

More pages can be explored first. There are two types of methods: subject-oriented it is driven by links. For the first, if the required description of the subject or sample page is required given that the spider algorithm can classify link lines with similarities (e.g., metric cosmetics) from the source page and / or the main text to describe the subject. For the latter, spider can it detects degrees in grades and outside of each page and finds the vcd turns to select a common page that contains many incoming links (authorities), or prefer a large summary page external links (axis).

##### E. KEEP UP-TO-DATE SPIDER PAGES

Dynamic web: there are many new pages, updated pages, deleted pages, and so on. Search engines need to check sites to be regularly scratched for updates and deletions. Spider you can find the html header information (for example, another Meta tag for update) to determine if it is page changed, and the entire page is updated only if necessary. I can continue with all that often the page refreshes and returns to the most dynamic page in the history. Box we recommend that you update

frequently accessed pages to optimize for freshness pages.

#### VI. VISUALIZATION OF INFORMATION RETRIEVAL

Visualizations of information retrieval performance include:

- Graphs which chart precision on one axis and recall on the other
- Histograms of average precision over various topics
- Receiver operating characteristic (ROC curve)
- Confusion matrix

#### VII. PERFORMANCE EVALUATION

Quality is assessed using the matching approach through the precision calculation, which is the assessment and technique used to retrieve the information to measure the accuracy of the search engine in relation to the integrity of the results returned. The average accuracy results indicate that the detection method outperforms all other methods of detecting the services of the previous generation in terms of accuracy.

#### VIII. CONCLUSION

As technology grows and the use of Internet of Things (IoT) increases, it is likely that this technology will be used soon. Internet of Things (IoT) is a concept that includes many techniques, which connects objects and or people anytime, anywhere. Information Retrieval is a resource-access activity related to the needs of resource-gathering information. IoT technology is a promising new technology made from variety technologies, which bring changes in recovery methods. However, the current recovery method does not guarantee a good result in the Internet of Things (IoT) environment, as it does not considered IoT characteristics.

#### REFERENCES

- [1] <http://ieeexplore.ieee.org/document/7508211/>
- [2] [http://www.ihomelab.ch/fileadmin/Dateien/PDF/Publikationen/2013/20131104\\_07\\_csist%20Minsk.pdf](http://www.ihomelab.ch/fileadmin/Dateien/PDF/Publikationen/2013/20131104_07_csist%20Minsk.pdf)
- [3] [http://www.site.uottawa.ca/~diana/csi4107/IR\\_draft.pdf](http://www.site.uottawa.ca/~diana/csi4107/IR_draft.pdf)
- [4] <http://www.ijptonline.com/wp-content/uploads/2017/07/29642-29647.pdf>
- [5] [http://dcs.gla.ac.uk/workshops/iASC2014/papers/iasc2014\\_caro.pdf](http://dcs.gla.ac.uk/workshops/iASC2014/papers/iasc2014_caro.pdf)
- [6] [http://onlinepresent.org/proceedings/vol142\\_2016/16.pdf](http://onlinepresent.org/proceedings/vol142_2016/16.pdf)
- [7] <https://arxiv.org/abs/1607.06884>
- [8] <http://www.tandfonline.com/doi/pdf/10.7305/automatika.54-4.414>
- [9] [https://link.springer.com/chapter/10.1007/978-981-10-5041-1\\_38](https://link.springer.com/chapter/10.1007/978-981-10-5041-1_38)
- [10] <file:///C:/Users/rahma/Downloads/9783642393136-c2.pdf>

- [11] [https://www.researchgate.net/publication/220230695\\_A\\_General\\_Model\\_of\\_Query\\_Processing\\_in\\_Information\\_Retrieval\\_Systems](https://www.researchgate.net/publication/220230695_A_General_Model_of_Query_Processing_in_Information_Retrieval_Systems)
- [12] [https://en.wikipedia.org/wiki/Information\\_retrieval](https://en.wikipedia.org/wiki/Information_retrieval)
- [13] <https://nlp.stanford.edu/IR-book/pdf/01bool.pdf>
- [14] [http://www.doc.ic.ac.uk/~nd/surprise\\_97/journal/vol4/hks/inf\\_ret.html](http://www.doc.ic.ac.uk/~nd/surprise_97/journal/vol4/hks/inf_ret.html)
- [15] <http://libguides.lamk.fi/informationretrieval/process>
- [16] <https://www.igi-global.com/dictionary/information-retrieval/14470>
- [17] [http://hlwiki.slais.ubc.ca/index.php/Information\\_retrieval](http://hlwiki.slais.ubc.ca/index.php/Information_retrieval)