RAMAPURAM
SRM
INSTITUTE OF SCIENCE & TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)

UNIT – 5

18CSC305J / ARTIFICIAL INTELLIGENCE

RAMAPURAM
SRM
INSTITUTE OF SCIENCE & TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)

18CSC305J-Artificial Intelligence

Unit- V

- Expert System Architecture
  Pros and cons of Expert system
- Rule based systems
  Frame based expert system
- Case study
- NLP – levels of NLP
- Syntactic and Semantic Analysis
  Information Retrieval
- Information Extraction
  Machine Translation
- NLP Applications

- Advance topics in Artificial Intelligence- Cloud Computing and Intelligent agent
- Business Intelligence and Analytics
- Sentiment Analysis
- Deep Learning Algorithms
- Planning and Logic in intelligent Agents

RAMAPURAM
SRM
INSTITUTE OF SCIENCE & TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)

Expert Systems

Expert Systems - Objectives

- Learn the meaning of an expert system
- Understand the problem domain and knowledge domain
- Learn the advantages of an expert system
- Understand the stages in the development of an expert system
- Examine the general characteristics of an expert system

Objectives

- Examine earlier expert systems which have given rise to today's knowledge-based systems
- Explore the applications of expert systems in use today
- Examine the structure of a rule-based expert system
- Learn the difference between procedural and nonprocedural paradigms
- What are the characteristics of artificial neural systems

What is an expert system?

"An expert system is a computer system that emulates, or acts in all respects, with the decision-making capabilities of a human expert."

Professor Edward Feigenbaum

Stanford University

- An expert system compared with traditional computer :

  Inference engine + Knowledge = Expert system

  (Algorithm + Data structures = Program in traditional computer)

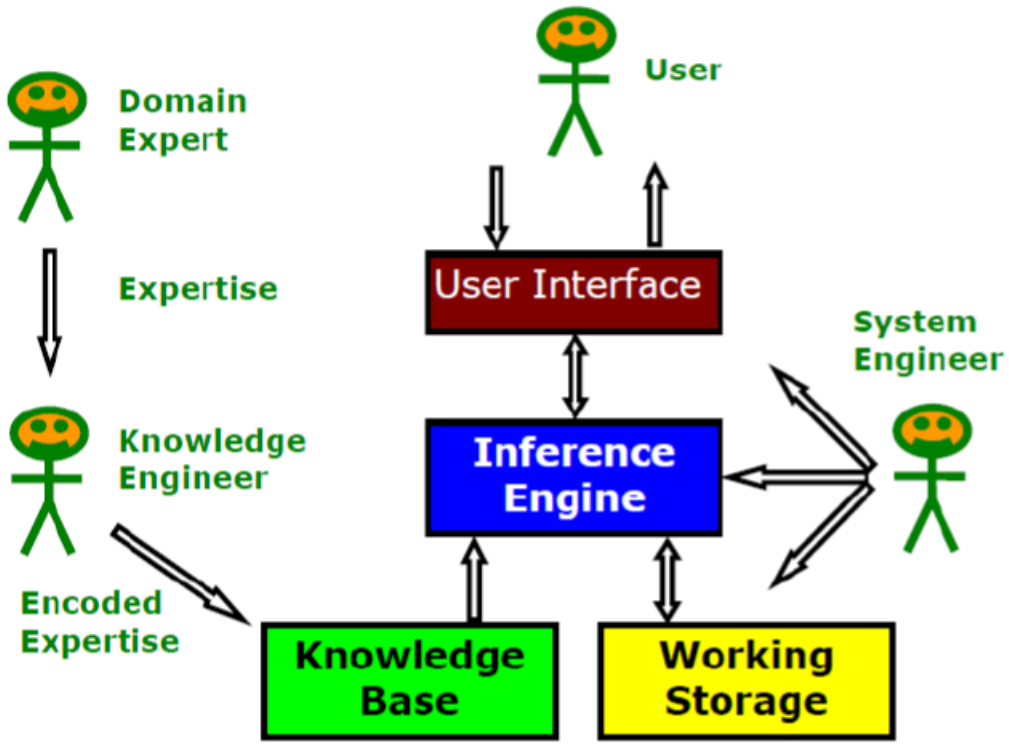
- First expert system, called DENDRAL, was developed in the early 70's at Stanford University.

Architecture of Expert Systems

Components of Expert Systems

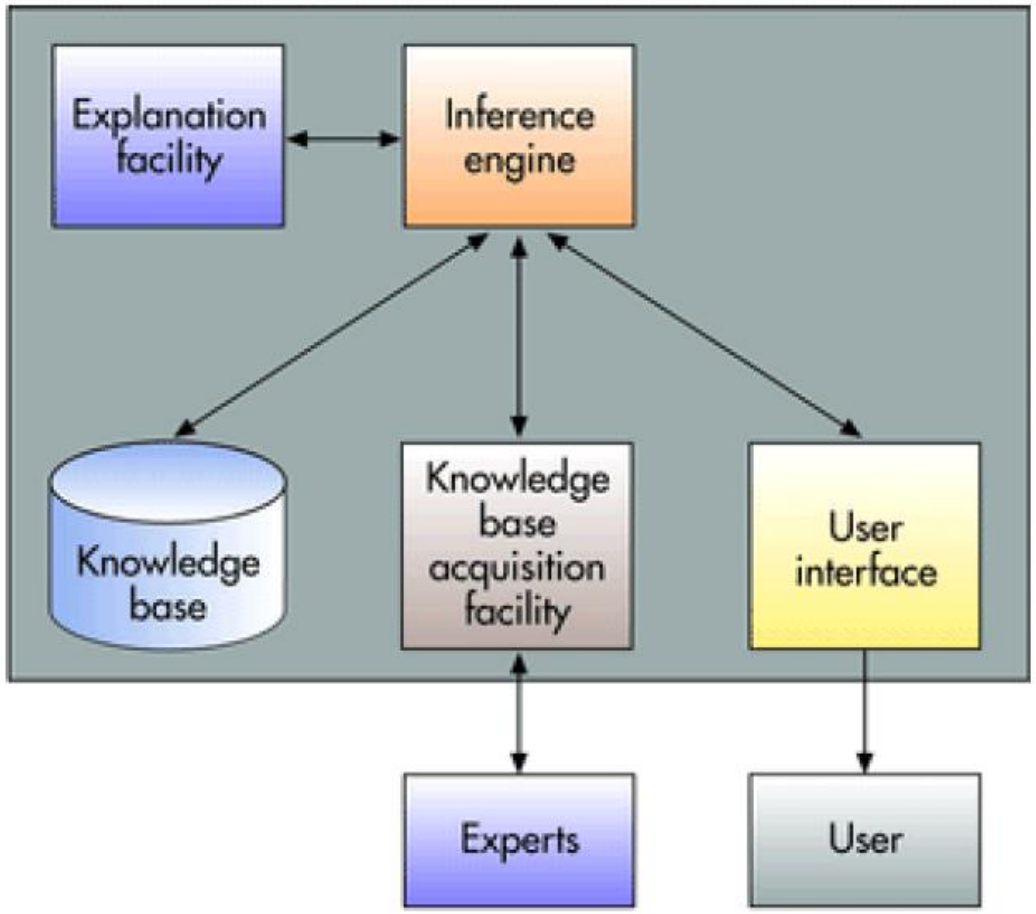
The components of ES include –

- Knowledge Base
- Interface Engine
- User Interface



Components of Expert System

Architecture of Expert Systems



Knowledge Base
Stores all relevant information, data, rules, cases, and relationships used by the expert system.
Uses
•Rules
•If-then Statements
•Fuzzy Logic

Inference Engine
Seeks information and relationships from the knowledge base and provides answers, predictions, and suggestions the way a human expert would.
Uses
•Backward Chaining
•Forward Chaining

Architecture of Expert Systems

Explanation Facility
Allows a user to understand how the expert system arrived at certain conclusions or results.
For example: it allows a doctor to find out the logic or rationale of the diagnosis made by a medical expert system

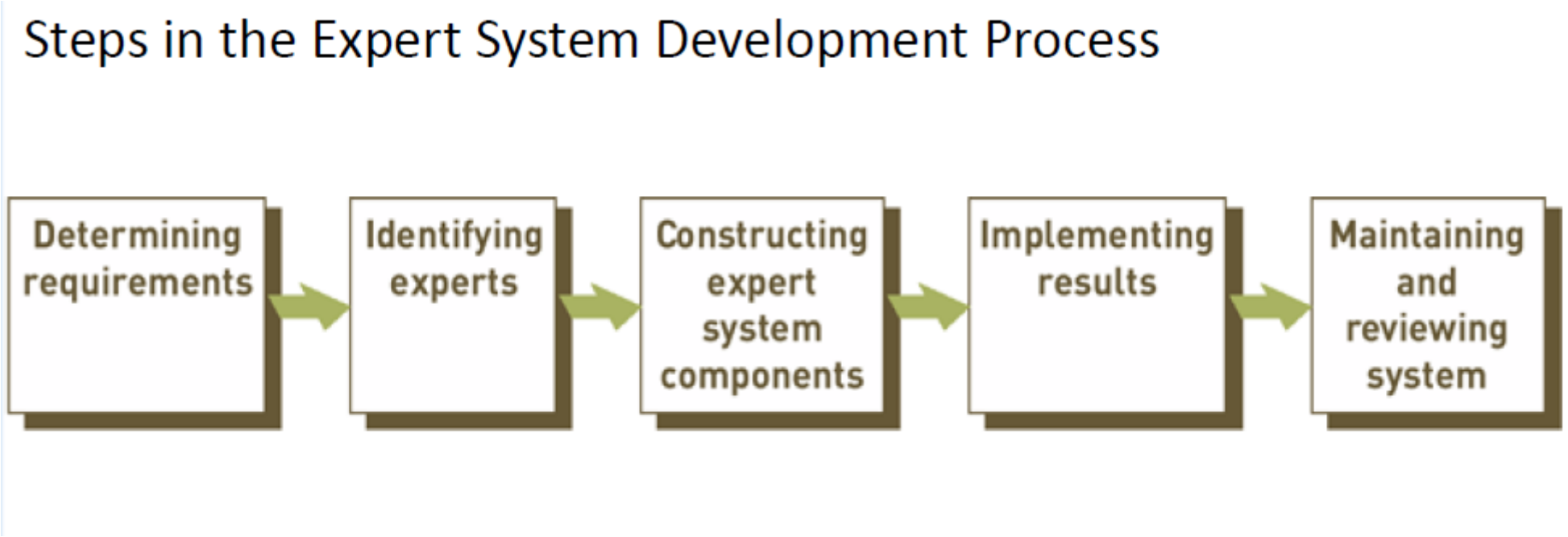
Knowledge acquisition facility
Provide convenient and efficient means of capturing and storing all the components of the knowledge base.
Acts as an interface between experts and the knowledge base.

User Interface
Specialized user interface software employed for designing, creating, updating, and using expert systems.
The main purpose of the user interface is to make the development and use of an expert system easier for users and decision makers

General Methods of Inferencing

- Forward chaining (data-driven)– reasoning from facts to the conclusions resulting from those facts – best for prognosis, monitoring, and control.
  - Examples: CLIPS, OPS5
- Backward chaining (query/Goal driven)– reasoning in reverse from a hypothesis, a potential conclusion to be proved to the facts that support the hypothesis – best for diagnosis problems.
  - Examples: MYCIN

Expert Systems Development



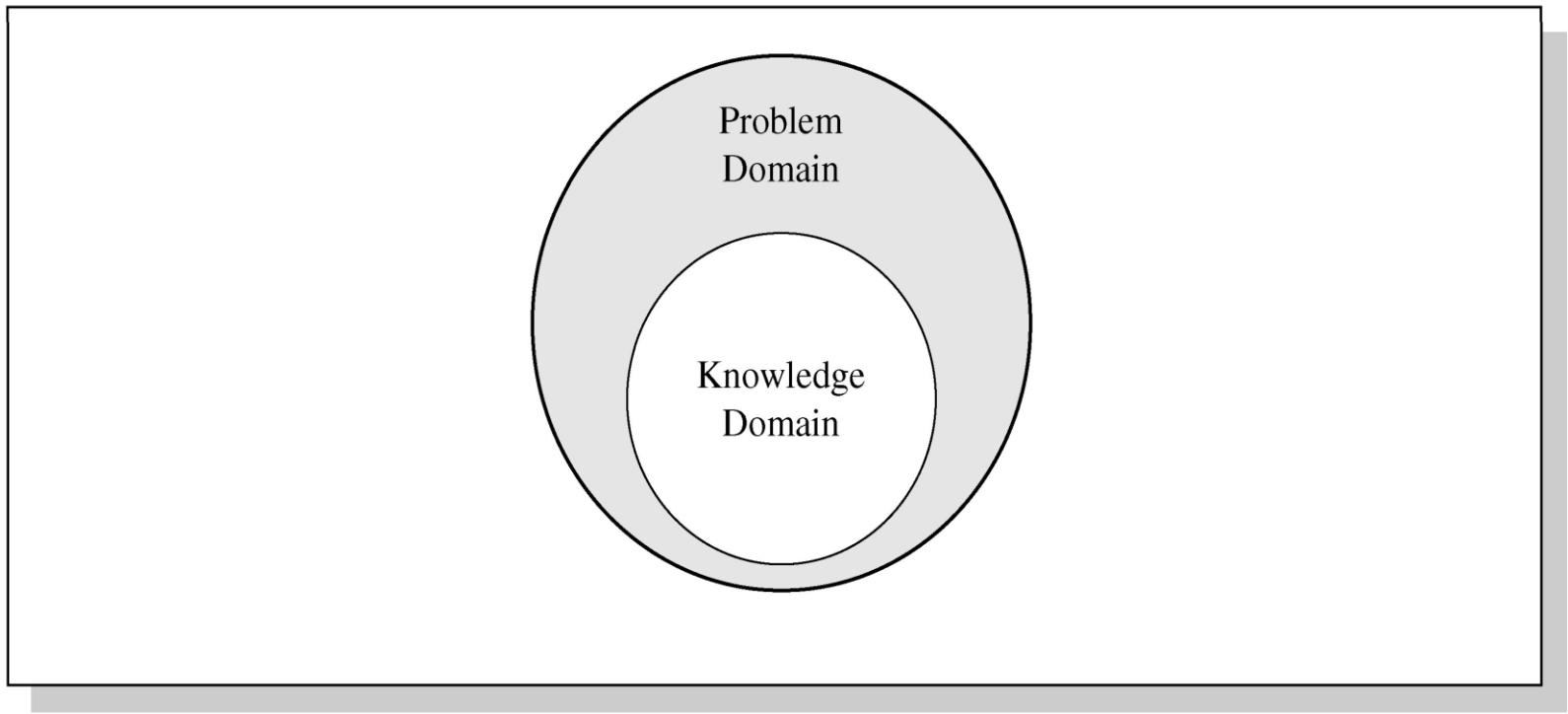
Expert system technology may include:

- Special expert system languages – CLIPS
- Programs
- Hardware designed to facilitate the implementation of those systems

Problem Domain vs. Knowledge Domain

- An expert's knowledge is specific to one problem domain – medicine, finance, science, engineering, etc.
- The expert's knowledge about solving specific problems is called the knowledge domain.
- The problem domain is always a superset of the knowledge domain.

Problem and Knowledge Domain Relationship



Advantages of Expert Systems

1. A consistent output
2. Quick and fast response
3. Location/date/day/time independent
4. Can be made generalised. (A change in application would result in looking out for a human expert related to that field if we are not relying on expert systems. But with expert systems, having generalisation, the process gets less complicated.)
5. Efficient utilisation of the knowledge (Human experts may forget some aspects, whereas expert systems consider all the rules and scenarios.)
6. Simple future enhancements with additional information of the knowledge and the rules (but with humans, it is difficult)
7. Easy maintenance of system.

Disadvantages of Expert Systems

The disadvantages of expert systems are mentioned below:

1. They cannot handle new dynamic situation.
2. The systems cannot be adaptive based on the decisions taken earlier.
3. Limited set of knowledge will leave them with limited set of decision outcomes.
4. Development cost could be high depending on the purpose they are used for.

Representing the Knowledge

The knowledge of an expert system can be represented in a number of ways, including IF- THEN rules:

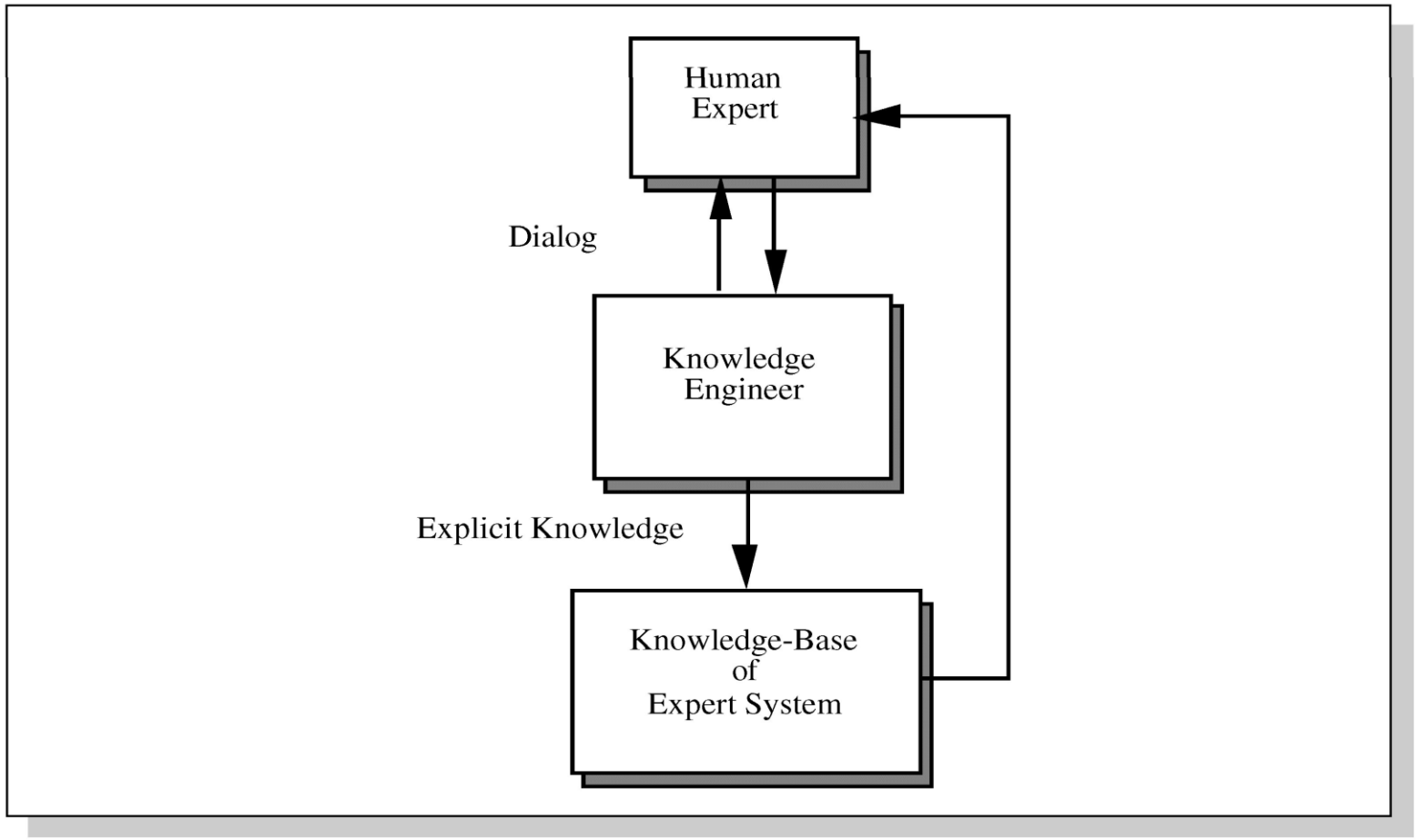
IF you are hungry THEN eat

Knowledge Engineering

The process of building an expert system:

1. The knowledge engineer establishes a dialog with the human expert to elicit knowledge.
2. The knowledge engineer codes the knowledge explicitly in the knowledge base.
3. The expert evaluates the expert system and gives a critique to the knowledge engineer.

Development of an Expert System



The Role of AI

- An algorithm is an ideal solution guaranteed to yield a solution in a finite amount of time.
- When an algorithm is not available or is insufficient, we rely on artificial intelligence (AI).
- Expert system relies on inference – we accept a "reasonable solution."

Uncertainty

- Both human experts and expert systems must be able to deal with uncertainty.
- It is easier to program expert systems with shallow knowledge than with deep knowledge.
- Shallow knowledge – based on empirical and heuristic knowledge.
- Deep knowledge – based on basic structure, function, and behavior of objects.

Early Expert Systems

- DENDRAL – used in chemical mass spectroscopy to identify chemical constituents
- MYCIN – medical diagnosis of illness
- DIPMETER – geological data analysis for oil
- PROSPECTOR – geological data analysis for minerals
- XCON/R1 – configuring computer systems

Broad Classes of Expert Systems

Class	General Area
Configuration	Assemble proper components of a system in the proper way.
Diagnosis	Infer underlying problems based on observed evidence.
Instruction	Intelligent teaching so that a student can ask *why*, *how,* and *what if* questions just as if a human were teaching.
Interpretation	Explain observed data.
Monitoring	Compares observed data to expected data to judge performance.
Planning	Devise actions to yield a desired outcome.
Prognosis	Predict the outcome of a given situation.
Remedy	Prescribe treatment for a problem.
Control	Regulate a process. May require interpretation, diagnosis, monitoring, planning, prognosis, and remedies.

Problems with Algorithmic Solutions

- Conventional computer programs generally solve problems having algorithmic solutions.
- Algorithmic languages include C, Java, and C#.
- Classical AI languages include LISP and PROLOG.

Considerations for Building Expert Systems

- Can the problem be solved effectively by conventional programming?
- Is there a need and a desire for an expert system?
- Is there at least one human expert who is willing to cooperate?
- Can the expert explain the knowledge to the knowledge engineer can understand it.
- Is the problem-solving knowledge mainly heuristic and uncertain?

Languages, Shells, and Tools

- Expert system languages are post-third generation.
- Procedural languages (e.g., C) focus on techniques to represent data.
- More modern languages (e.g., Java) focus on data abstraction.
- Expert system languages (e.g. CLIPS) focus on ways to represent knowledge.

Production Rules

- Knowledge base is also called production memory.
- Production rules can be expressed in IF-THEN pseudocode format.
- In rule-based systems, the inference engine determines which rule antecedents are satisfied by the facts.

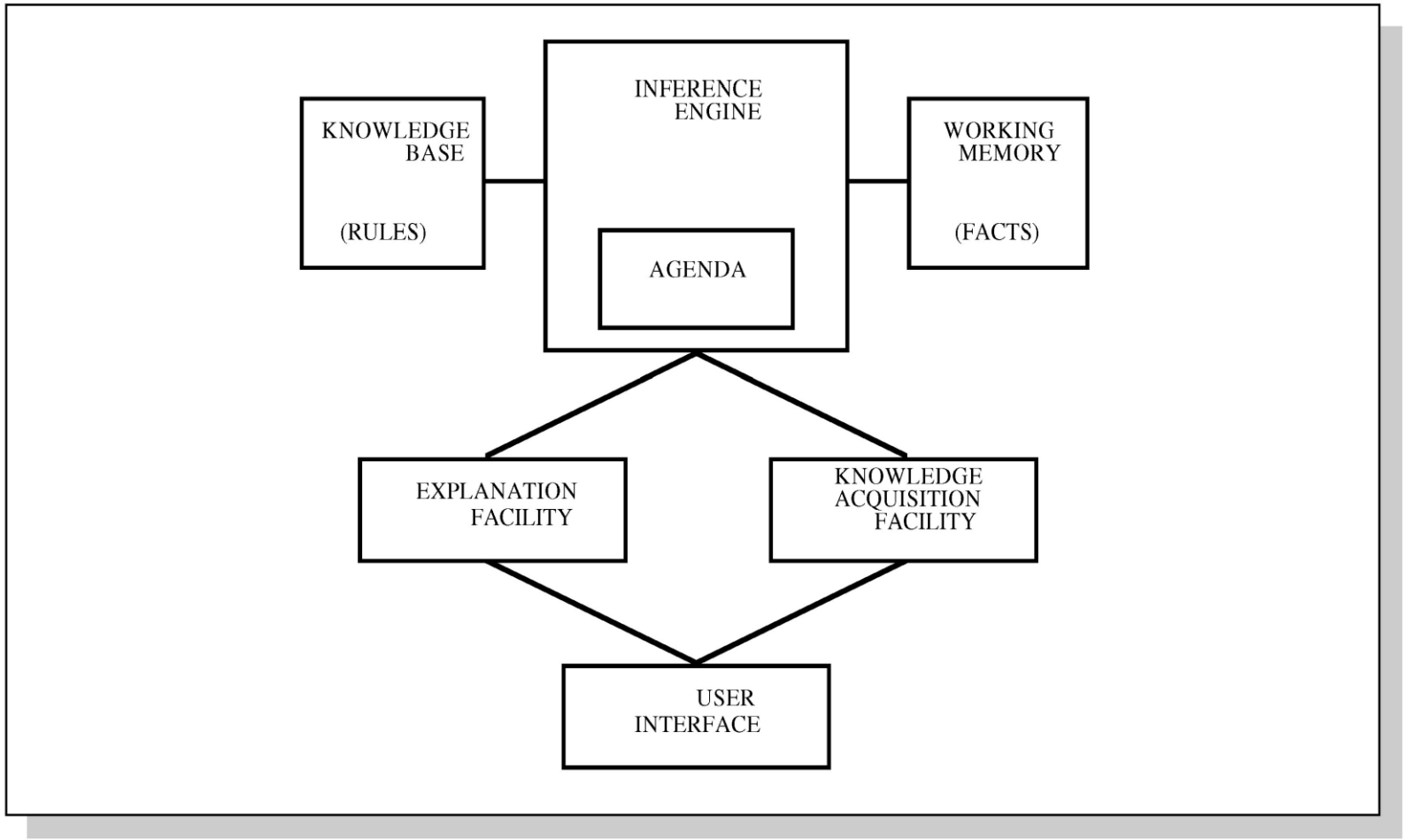
Rule-Based Expert System

- A rule based expert system is the simplest form of artificial intelligence and uses prescribed knowledge based rules to solve a problem
- The aim of the expert system is to take knowledge from a human expert and convert this into a number of hardcoded rules to apply to the input data
- In their most basic form, the rules are commonly conditional statements (if a, then do x, else if b, then do y)
- These systems should be applied to smaller problems, as the more complex a system is, the more rules that are required to describe it, and thus increased difficulty to model for all possible outcomes

Example:
A very basic example of rule based expert system would be a program to direct the management of abdominal aneurysms The system would input the diameter of an aneurysm Using conditional arguments, the input diameter would be stratified to recommend whether immediate intervention was required, and if not what appropriate follow up is recommended

Note: with problems related to radiological images, often preprocessing of the images is required prior to the expert system being applied.

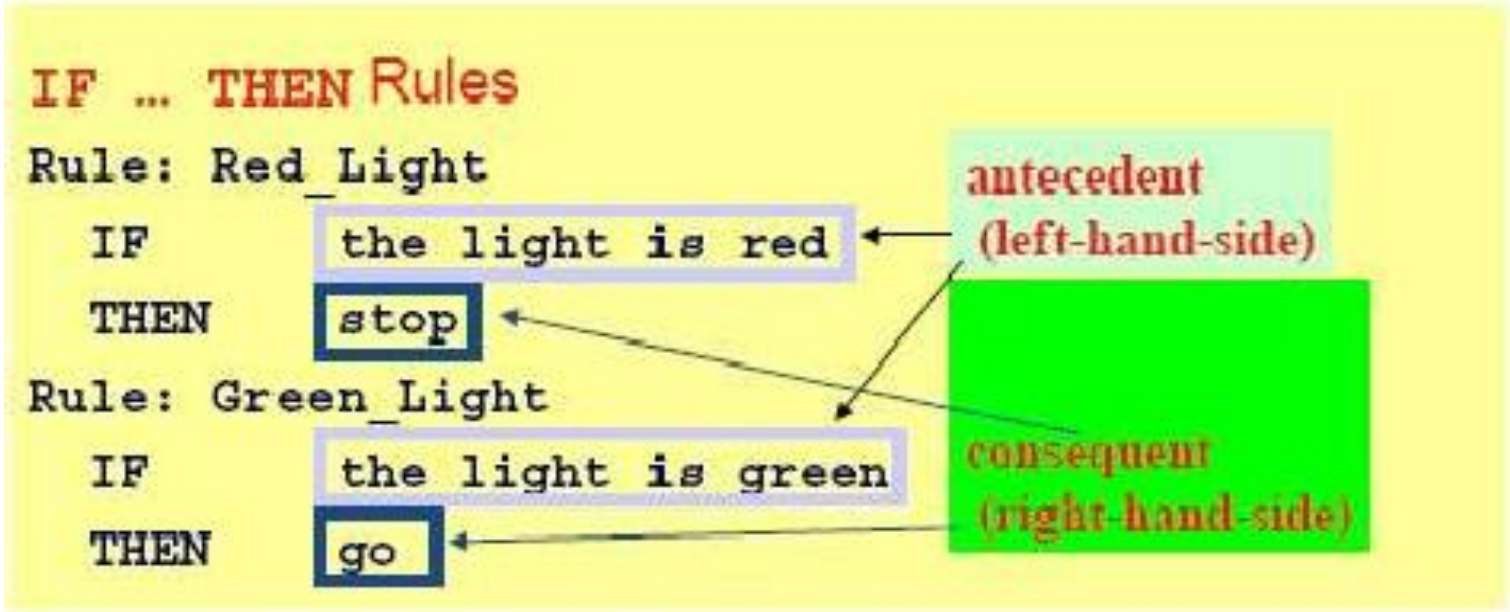
Structure of a Rule-Based Expert System

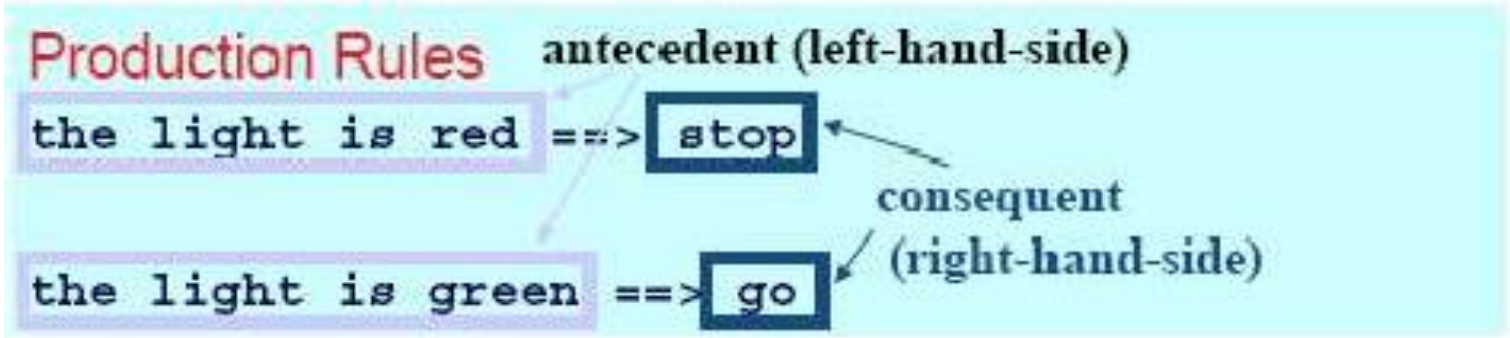


Rule-Based ES

- knowledge is encoded as **IF** ... **THEN** rules
  - these rules can also be written as *production rules*
- the inference engine determines which rule antecedents are satisfied
  - the left-hand side must "match" a fact in the working memory
- satisfied rules are placed on the agenda
- rules on the agenda can be activated ("fired")
  - an activated rule may generate new facts through its right-hand side
  - the activation of one rule may subsequently cause the activation of other rules

Example Rules

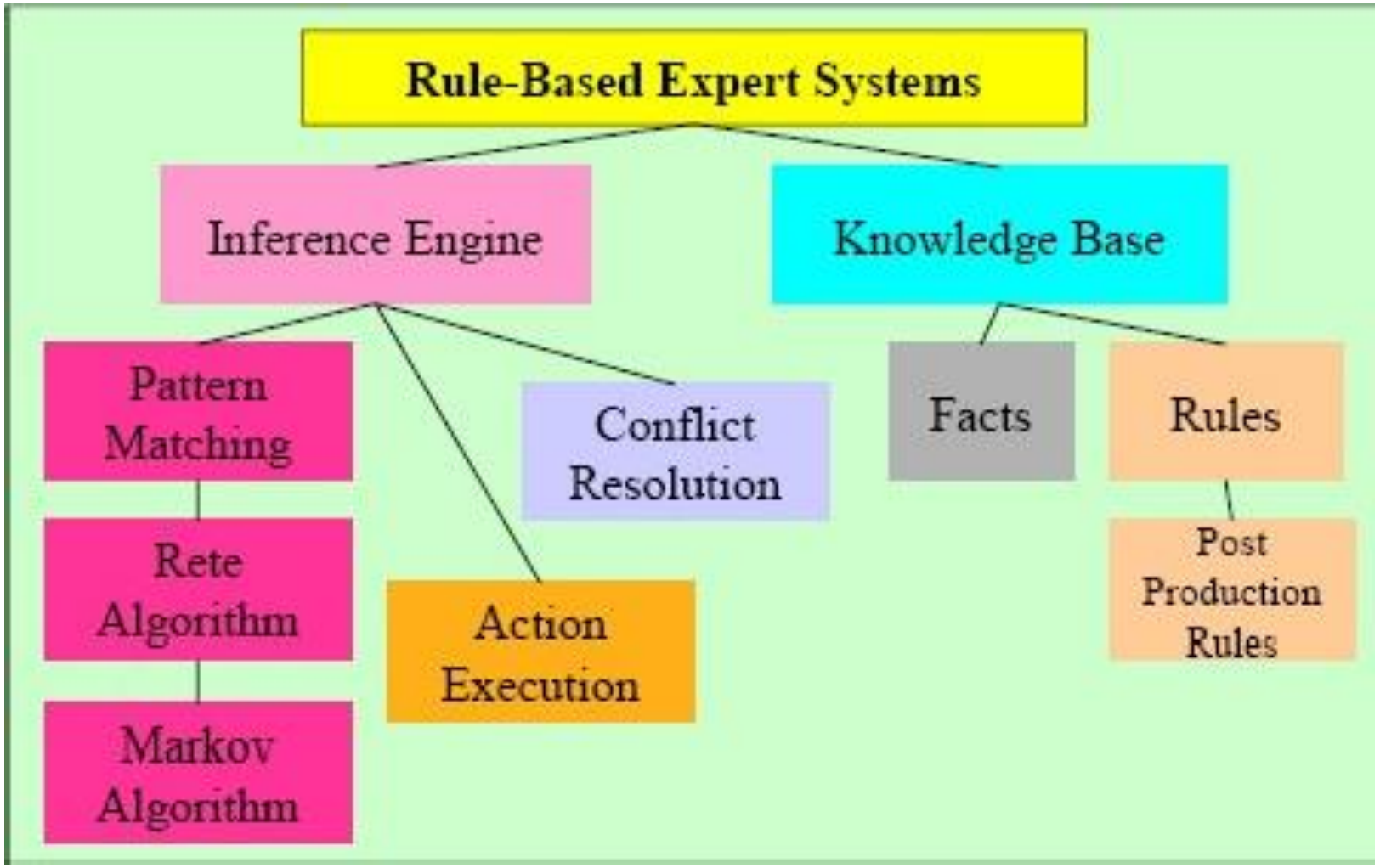




Inference Engine Cycle

- The inference engine determines the execution of the rules by the following cycle:
  - conflict resolution
    - select the rule with the highest priority from the agenda
  - execution (Act)
    - perform the actions on the consequent of the selected rule
    - remove the rule from the agenda
  - match
    - update the agenda
      - add rules whose antecedents are satisfied to the agenda
      - remove rules with non-satisfied agendas
- the cycle ends when no more rules are on the agenda, or when an explicit stop command is encountered

Foundation of Expert Systems



Markov Algorithm

- An ordered group of productions applied in order or priority to an input string.
- If the highest priority rule is not applicable, we apply the next, and so on.
- inefficient algorithm for systems with many rules.
- Termination on (1) last production not applicable to a string, or (2) production ending with period applied
- Can be applied to substrings, beginning at left

Markov Algorithm

(1) αxy → yαx
(2) α → ^
(3) ^ → α

Rule	Success or Failure	String
1	F	ABC
2	F	ABC
3	S	α ABC
1	S	B α AC
1	S	BC α A
1	F	BC α A
2	S	BCA

Table 1.11 Execution Trace of a Markov Algorithm

Rete Algorithm

- Markov: too inefficient to be used with many rules
- Functions like a net – holding a lot of information.
- Much faster response times and rule firings can occur compared to a large group of IF-THEN rules which would have to be checked one-by-one in conventional program.
- Takes advantage of temporal redundancy and structural similarity.
- Looks only for changes in matches (ignores static data)
- Drawback is high memory space requirements.

Frame-Based Expert System

- The expert systems which make use of frames for the knowledge are called frame-based expert systems.
- What is a frame? – A frame is a data structure with typical knowledge about the object or concept.
- Frame has its name and set of attributes
- Example : A car frame can have make, type, color and so on as slots/attributes in the frame
- Each slot/ attribute has unique value associated to it

Frame-Based Expert System

We can have the following included in the slot

1. Frame Name
2. Relationship with other frames
3. Values or Ranges
4. Procedural information

Class: Car
Make:
Type:
Colour:
Airbags:
Auto-transmission:
Windows:

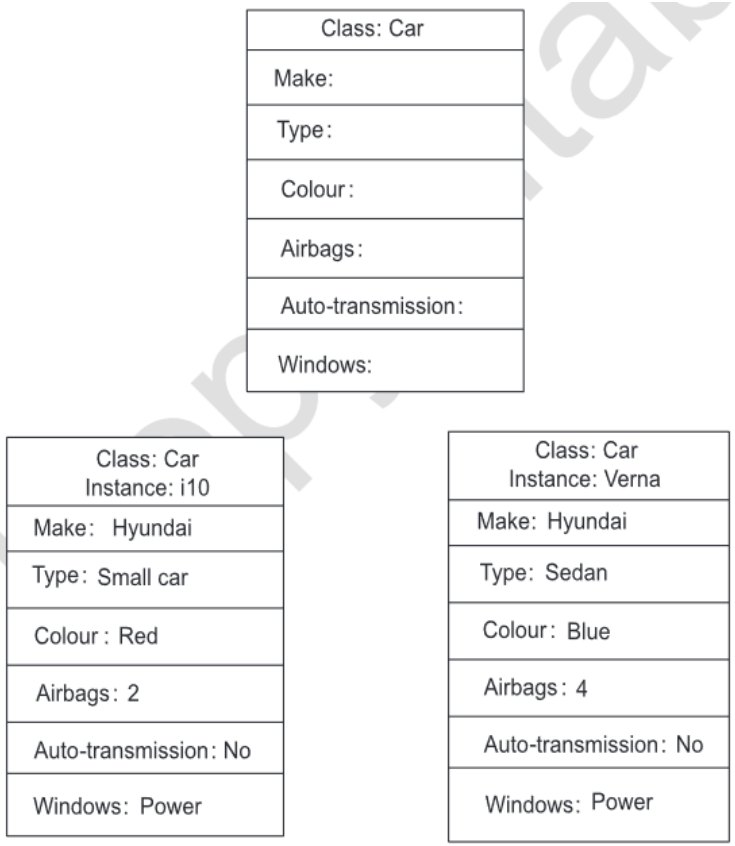
Class: Car Instance: i10
Make: Hyundai
Type: Small car
Colour: Red
Airbags: 2
Auto-transmission: No
Windows: Power

Class: Car Instance: Verna
Make: Hyundai
Type: Sedan
Colour: Blue
Airbags: 4
Auto-transmission: No
Windows: Power

Represents the frame structure for Class and Instance

Frame-Based Expert System

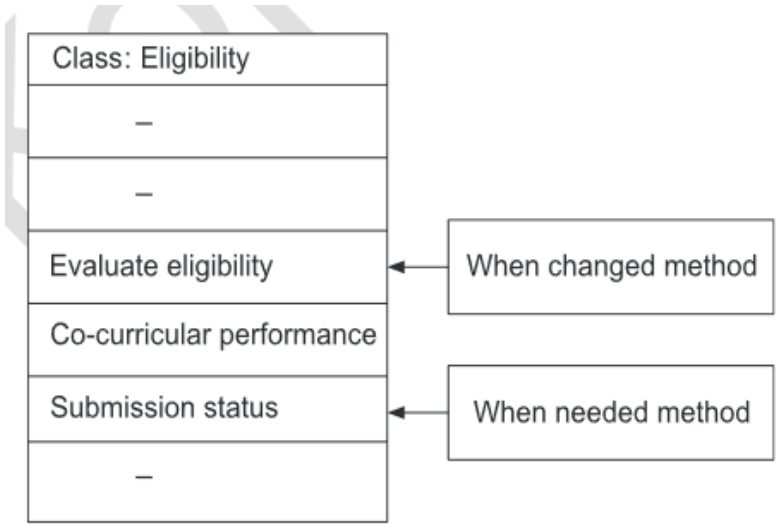
- Relationship – hierarchy, or to be specific, the inheritance is depicted
- Values & Ranges – Represent the actual/default values or specified ranges
- Procedural information – the slot is attached to a procedure that is executed when any event is triggered such as change in value for the slot
- Instance Frame refers to an object and Class frame refers to group



Frames : Class and Instance

Working of Frame-Based Expert System

- Method – A method is a procedure that is executed when requested
- Demon – makes use of if-then structure
- Frames that have the knowledge representation, the methods or the demons essentially add actions to them.
- As an example, how the process of an expert system works to check the eligibility of a student appearing for an exam is explained in the picture



When needed and when required methods : Snapshot

Guidelines to build a Frame-Based Expert System

Guidelines to Build a Frame-based Expert System

1. The first step that is involved in designing any expert system is the scope and the specificity of the problem at hand. There has to be utmost clarity about these two factors.
2. The next step is the identification of classes, instances and the attributes.
3. Since the events or to be specific, the methods are the key role players in evaluation, the display of the system has to be presented in the most simple and transparent manner.
4. The next task is to have the methods of when changed and when needed. This is dependent to the design in step 3, where the events are decided.
5. Rules are to be also defined along with the methods.
6. Finally, a full-proof evaluation of the built system, and if required, expansion should be handled.

A detailed example is handled in the case study for the frame-based expert system, provided further in the chapter.

RAMAPURAM
SRM
INSTITUTE OF SCIENCE & TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)

MYCIN

MYCIN

- MYCIN was an early expert system that used artificial intelligence to identify bacteria causing severe infections.
- recommend antibiotics, with the dosage adjusted for patient's body weight
- The MYCIN system was also used for the diagnosis of blood clotting diseases.
- MYCIN was developed over five or six years in the early 1970s at Stanford University.
- It was written in Lisp

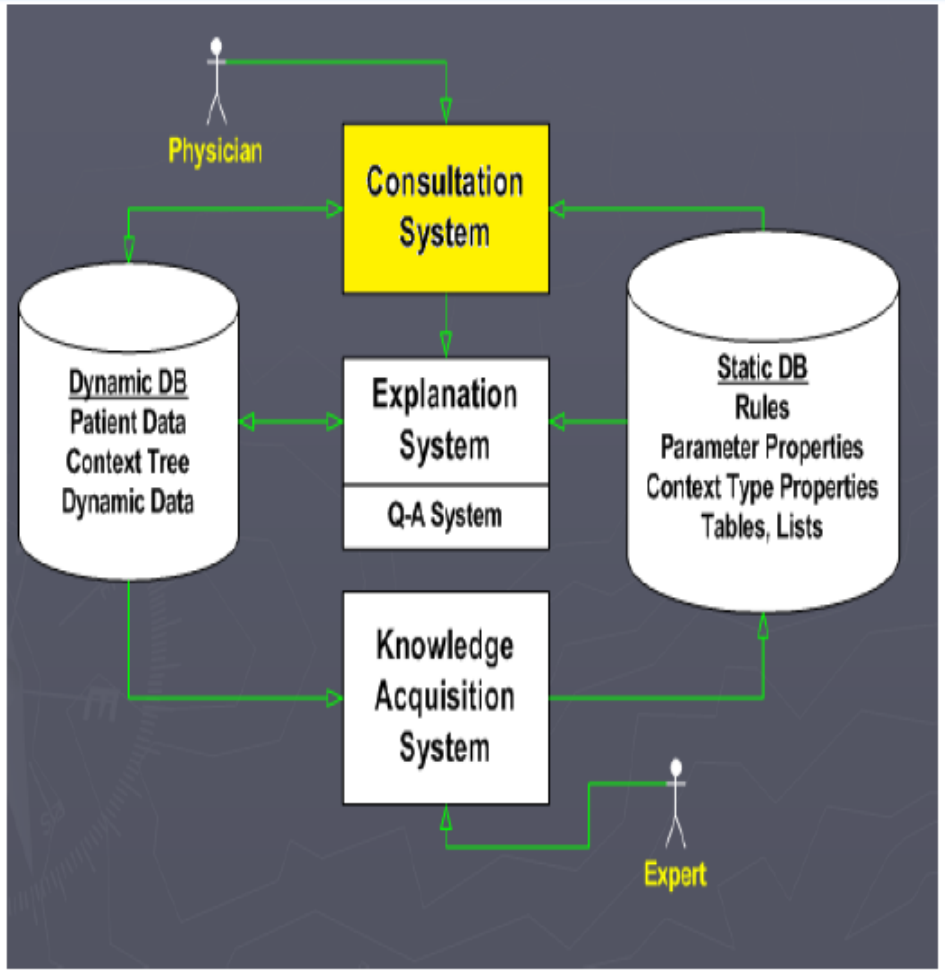
MYCIN

- MYCIN was a stand alone system that required a user to enter all relevant information about a patient by typing in responses to questions MYCIN posed.
- MYCIN operated using a fairly simple inference engine, and a knowledge base of ~600 rules.
- It would query the physician running the program via along series of simple yes/no or textual questions.

Tasks and Domain

- Disease DIAGNOSIS and Therapy SELECTION
- Advice for non-expert physicians with time considerations and incomplete evidence on:
  - –Bacterial infections of the blood
  - –Expanded to meningitis and other ailments
  - –Meet time constraints of the medical field

Consultation System

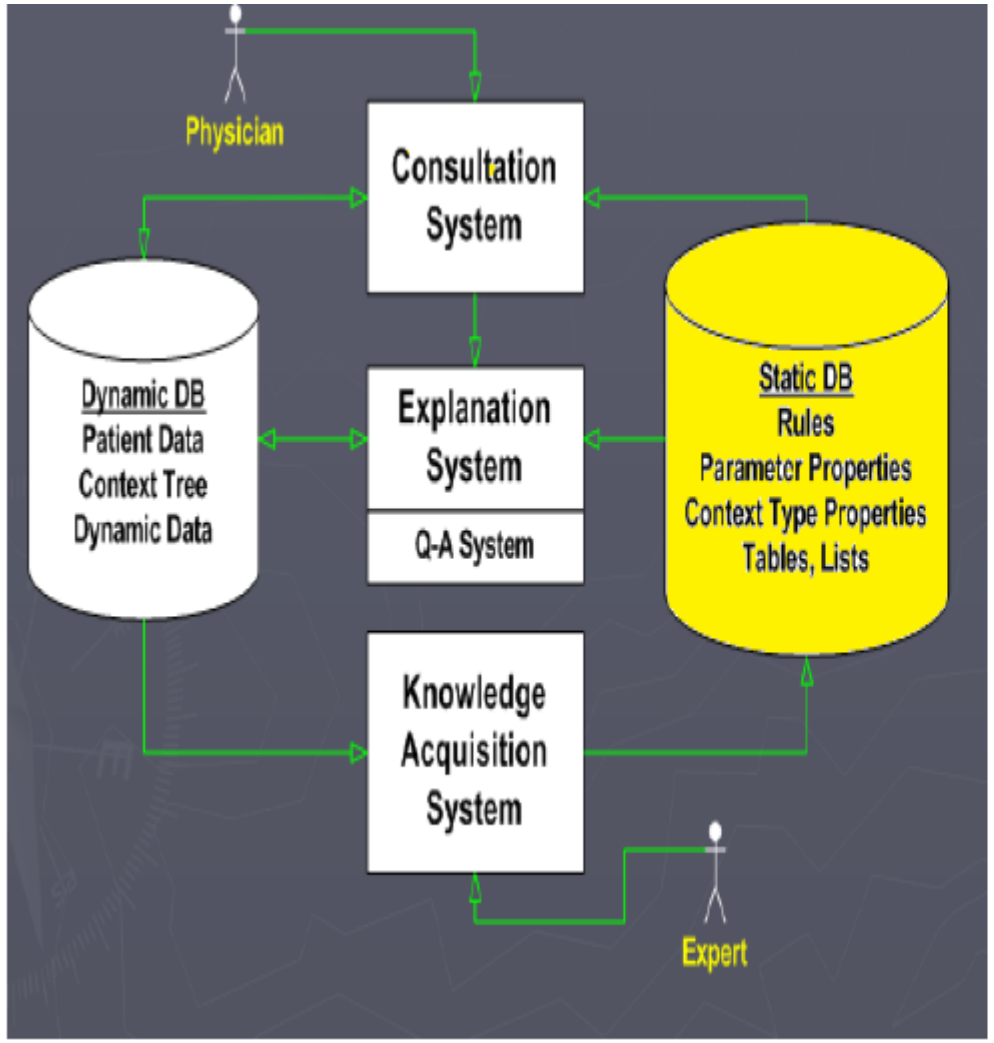


- Performs Diagnosis and Therapy Selection
- Control Structure reads Static DB (rules) and read/writes to Dynamic DB (patient, context)
- Linked to Explanations
- Terminal interface to Physician

Consultation “Control Structure”

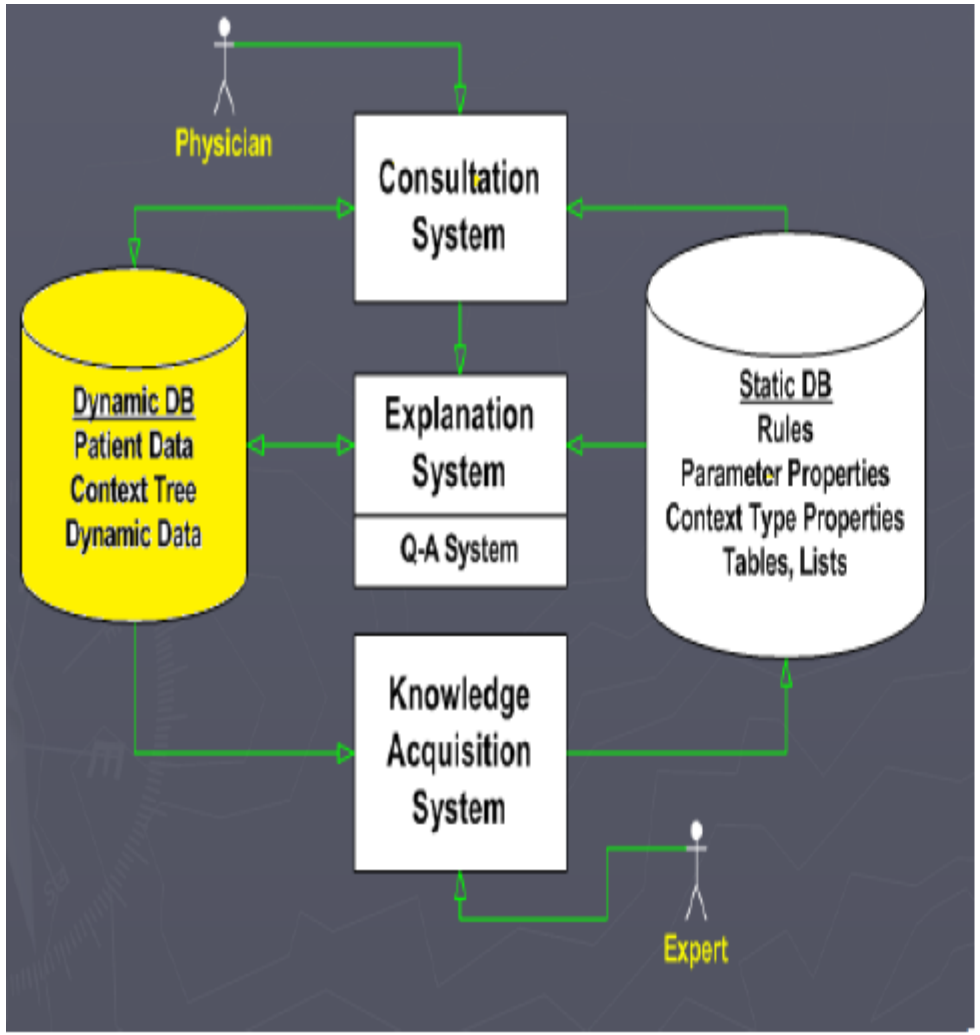
- Goal-directed Backward-chaining Depth-first Tree Search
- High-level Algorithm:
    1. Determine if Patient has significant infection
    2. Determine likely identity of significant organisms
    3. Decide which drugs are potentially useful
    4. Select best drug or coverage of drugs

Static Database



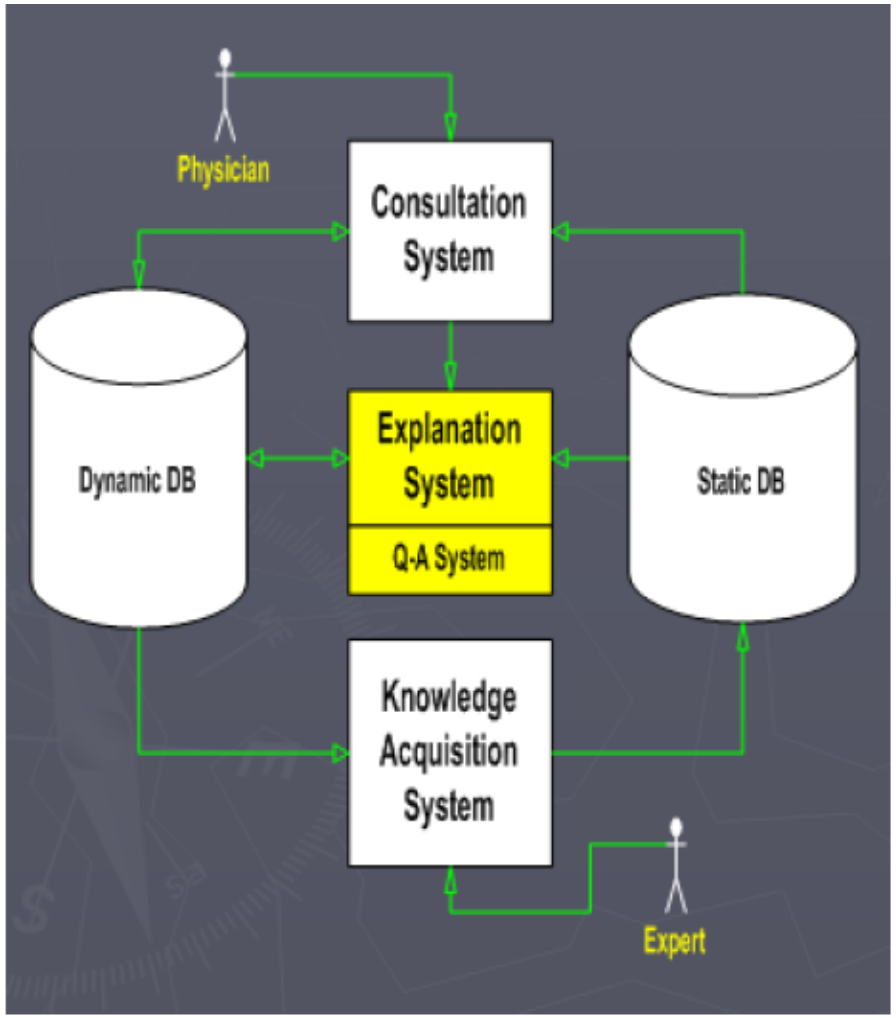
- Rules
- Meta-Rules
- Templates
- Rule Properties
- Context Properties
- Fed from Knowledge Acquisition System

Dynamic Database



- Patient Data
- Laboratory Data
- Context Tree
- Built by Consultation System
- Used by Explanation System

Explanation System



- Provides reasoning why a conclusion has been made, or why a question is being asked
- Q-A Module
- Reasoning Status Checker

Xcon

•The R1 (internally called XCON, for eXpertCONfigurer) program was a production rule based system written in OPS5 by John P. McDermott of CMU in 1978.

–configuration of DEC VAX computer systems

•ordering of DEC's VAX computer systems by automatically selecting the computer system components based on the customer's requirements.

•XCON first went into use in 1980 in DEC's plant in Salem, New Hampshire. It eventually had about 2500 rules.

•By 1986, it had processed 80,000 orders, and achieved 9598% accuracy.

•It was estimated to be saving DEC $25M a year by reducing the need to give customers free components when technicians made errors, by speeding the assembly process, and by increasing customer satisfaction.

•XCON interacted with the sales person, asking critical questions before printing out a coherent and workable system specification/order slip.

•XCON's success led DEC to rewrite XCON as XSEL a version of XCON intended for use by DEC's salesforce to aid a customer in properly configuring their VAX.

XCON: Expert Configurer

Stages of Expert System building

- Identification:
  Problems, data, goals, company, people…
- Conceptualization:
  Characterize different kinds of concepts and relations
- Formalization:
  Express character of search
- Implementation:
  Build the system in executable form
- Testing and Evaluation:
  Does it do what we wanted?
- Maintenance
  Adapt to changing environment or requirements

RAMAPURAM
SRM
INSTITUTE OF SCIENCE & TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)

Natural Language Processing (NLP)

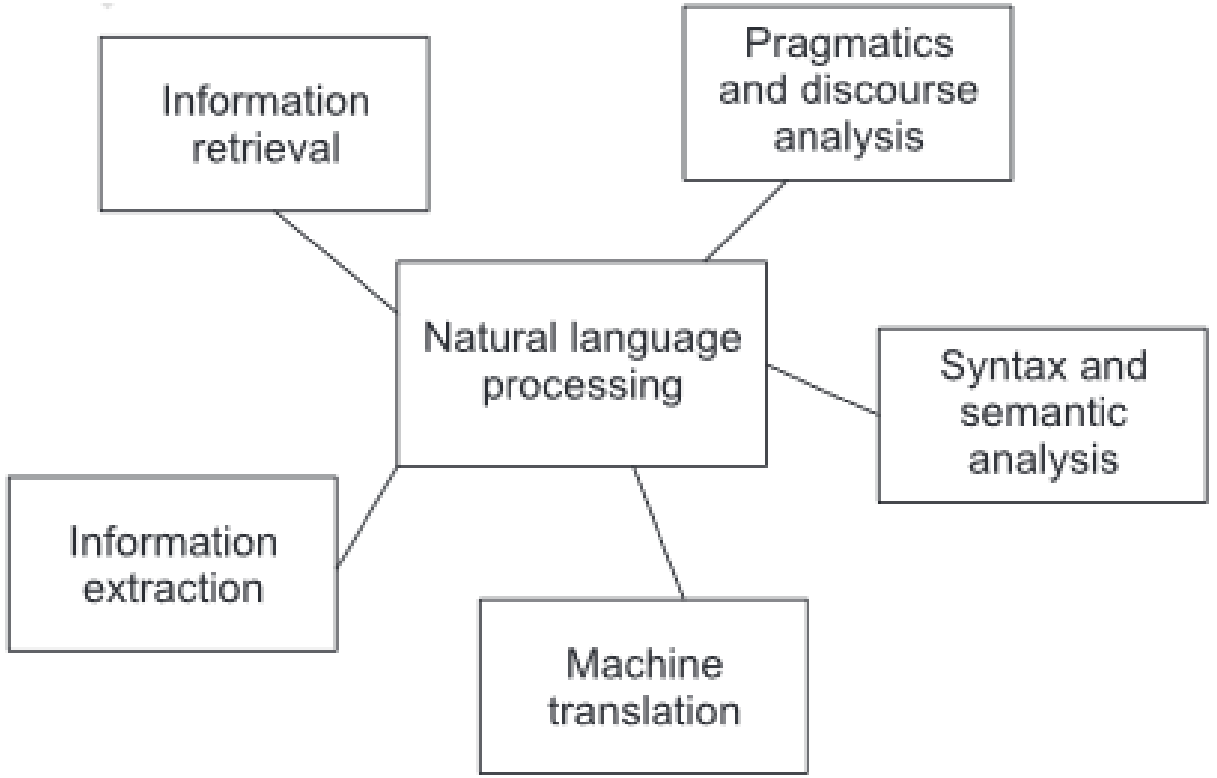
Natural language

- **Natural languages** are languages that living creatures use for communication,
- A machine is considered to be really intelligent only when it can understand and interpret or speak the matter of natural language.
- The capability to understand, interpret and communicate through natural language is a very important criteria of intelligent behavior.

Why Natural language processing?

- Huge amount of data?
  - Internet=at least 2.5 billion pages
- Applications for processing large amounts of texts.
  - Classify text into categories
  - Index and search large texts
  - Automatic translation
  - Speech understanding: Understanding phone conversation
  - Information extraction: Extract useful information from resumes
  - Automatic summarization
  - Question answering
  - Knowledge acquisition: knowledge from expert
  - Text generations/dialogs
- All these requires natural language expertise.

NLP Tasks


Information retrieval
Pragmatics and discourse analysis
Natural language processing
Syntax and semantic analysis
Information extraction
Machine translation

Levels of NLP

1. ***Morphology:*** It is the analysis of individual words that consist of morphemes—the smallest grammatical unit. Generally, words with 'ing', 'ed' change the meaning of the word. This analysis becomes necessary in the determination of tense as well.

2. ***Syntax:*** Syntax is concerned with the rules. It includes legal formulation of the sentences to check the structures. (Some aspects are covered in compiler's phase of syntax analysis that you must have studied). For example, 'Hari is good not to.' The sentence structure is totally invalid here.

3. ***Semantic:*** During this phase, meaning check is carried out. The way in which the meaning is conveyed is analysed. The previous example is syntactically as well as semantically wrong. Now, consider one more example, i.e., 'The table is on the ceiling.' This is syntactically correct, but semantically wrong.

4. ***Discourse integration:*** In communication or even in text formats, often the meaning of the current sentence is dependent on the one that is prior to it. Discourse analysis deals with the identification of discourse structure.

5. ***Pragmatics:*** In this phase, analysis of the response from the user with reference to what actually the language meant to convey is handled. So, it deals with the mapping for what the user has interpreted from the conveyed part and what was actually expected. For a question like 'Do you know how long it will take to complete the job?', the expected answer is the number of hours rather than a yes or no.

6. ***Prosody:*** It is an analysis phase that handles rhythm. This is the most difficult analysis that plays an important role in the poetry or *shlokas* (chants involving the name of God) that follow a rhythm.

7. ***Phonology:*** This involves analysis of the different kinds of sounds that are combined. It is concerned with speech recognition.

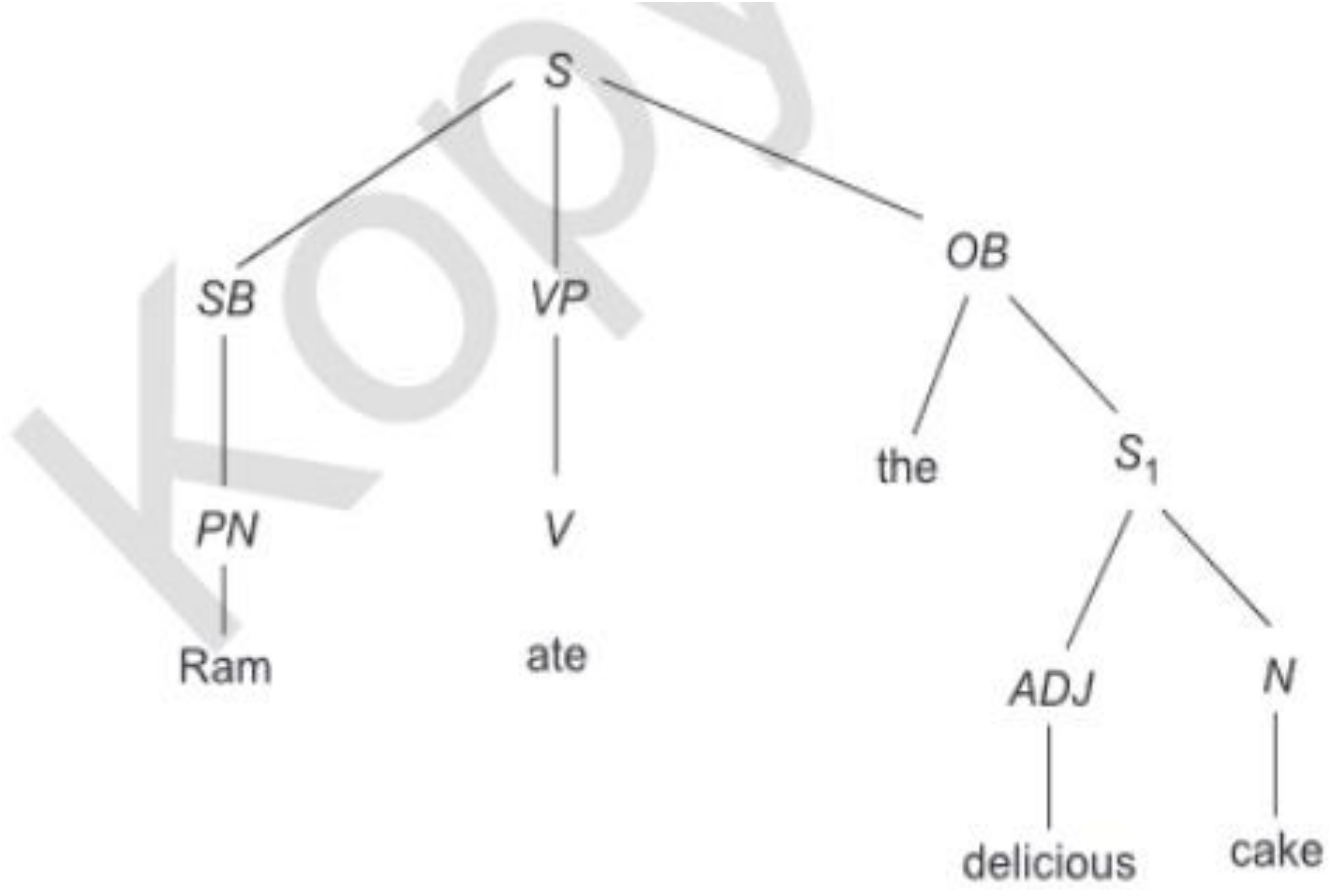
Syntactic Analysis

- Rules of syntax (grammar) specify the possible organization of words in sentences and allows us to determine sentence's structure(s)
  - "John saw Mary with a telescope"
    - John saw (Mary with a telescope)
    - John (saw Mary with a telescope)
- Parsing: given a sentence and a grammar
  - Checks that the sentence is correct according with the grammar and if so returns a **parse tree** representing the structure of the sentence

Syntactic Analysis - Grammar

- `sentence -> noun_phrase, verb_phrase`
- `noun_phrase -> proper_noun`
- `noun_phrase -> determiner, noun`
- `verb_phrase -> verb, noun_phrase`
- `proper_noun -> [mary]`
- `noun -> [apple]`
- `verb -> [ate]`
- `determiner -> [the]`

Syntactic Analysis - Parsing



Syntactic Analysis – Complications (1)

- Number (singular vs. plural) and gender
  - **sentence-> noun_phrase(n),verb_phrase(n)**
  - **proper_noun(s) -> [mary]**
  - **noun(p) -> [apples]**
- Adjective
  - **noun_phrase-> determiner,adjectives,noun**
  - adjectives-> adjective, adjectives
  - adjective->[ferocious]
- Adverbs, …

Syntactic Analysis – Complications (2)

- Handling ambiguity
  - Syntactic ambiguity: “fruit flies like a banana”
- Having to parse syntactically incorrect sentences

Semantic Analysis

- Syntax analysis is doing the parsing activity
- But we need to understand the meaning of the words and it is done by semantic analysis
- For example,
  - ‘Keep the book on the table’ – Here table refers physical object
  - ‘Learn the table of number 23’ – here table refers mathematics concept of table

Lexical Processing

- In lexical processing, the meaning of the tokens is found out
- Word sense disambiguation: Understanding the meaning of a particular word in the context
- It is concerned with the sense where it would be operational
- It would be done with the help of semantic marker
- Semantic marker: ‘Keep’ in sentence 1
- Semantic marker: ‘Learn’ in sentence 2

Semantic grammars

- Example,
- ‘The pen is on the ceiling’
- Solution is,
- S -> **Action** the **Food**
- Action -> eat|drink|shallow|chew – Set of words
- Food -> burger|sandwich|coke|pizza – Set of words

Case Grammar

- Case grammar is also called as Fillmore grammar
- Elements of the sentence are:
  - Object (thing on which it is acted)
  - Agent/actor (Someone who carries out the action or the event)
  - Dative (Someone who is affected by the event)
  - Location (place where the event/action occurs)
  - Time (date or time at which the action/event takes place)
- Example,
  - Rohit will meet Kunal at Mall

Conceptual Dependency

- ATRANS: Transfer of some relationship for verb (e.g.) ’give’
- PTRANS: Transfer of location of an object, say the verb ‘go’
- GRASP: This primitive occurs with verb like ‘throw’
- SPEAK: This primitive has some sound production, say the verb ‘speak’ or ‘say’

Information retrieval

Let us first understand the problem of information retrieval. So, what is it? Suppose you want to retrieve information about first players who scored 200+ runs in a single inning in ODI. So, you are in discussion with your friend who knows about cricket. First, you will get information about all players who scored 200+ runs, and then you will get years in which they scored. This simple retrieval will, conclude that Sachin Tendulkar is the first player to score 200+ runs. Similarly, we come across various scenarios in everyday life which demand information retrieval.

A very common example that we have discussed in the introduction is of the search engines. The search engine performs information retrieval (IR). We type in a query and the information required by us is retrieved. Figure 12.8 depicts the basic IR system.

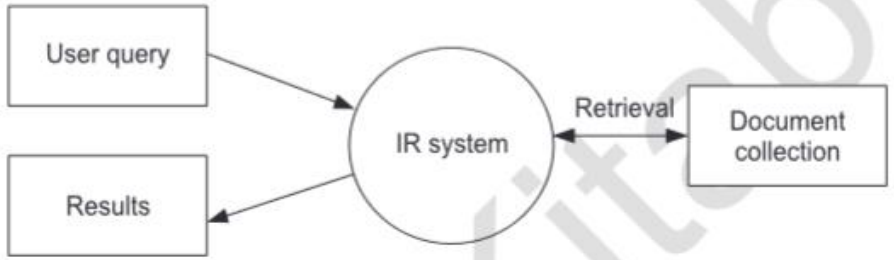


Figure 12.8 An IR system—Getting relevant documents.

We have talked about the search, but the question here is—is it possible to have an optimal way that can be exploited by the search engines for retrieval? What could it be and how would the things work here?

Information retrieval - Models

Boolean Model

In Boolean model, the tokens are treated in the form of 1's and 0's. The presence or absence of a term is marked accordingly. So, the document is represented in the form of t_1 ^ t_2 … t_n, where t_i forms a term. The query can be t_1 ^ ~t_2. The Boolean mapping is found to be the simplest one that was used in earlier systems. With the Boolean model, the requirement is to find the probability of relevance R for a document d and a query q. It is represented as follows:

$P(R \mid d, q)$, where R can be true or false.

But it has a lot of drawbacks. Since it is dependent on the bit presence, the ordering becomes an issue. (In what way the relevant ones are to be presented?). Further, matching similar words, for example, if the terms are cardigan and sweater, then it becomes another issue. Most importantly, a question raises up here—is it possible as an end user to formulate the Boolean query? Though it is pretty complex, still, many probabilistic approaches have been developed with Boolean models.

Information retrieval - Models

Bag of Words

The frequency of a word is taken into account instead of ordering. Consider the following texts:

Rohan drives Audi. (text 1)

Sameer drives BMW. He drives Mercedes also. (text 2)

(At a very basic level, to understand the concept, the words identified here are Rohan, drives, Audi, Sameer, he, Mercedes, also). The dictionary or the bag of words has these tokens. So, a vector for the text 1 and 2 would be

[1, 1, 1, 0, 0, 0, 0] for text 1

[0, 2, 0, 1, 1, 1, 1] for text 2

The count in the vectors is the frequency of occurrence of the identified terms in the text.

Information retrieval - Models

Vector Space Model

In vector space model, the documents and the query are represented as *vectors*. A vector comprises dimensions that are the terms used to index it. The model is given below:

Vector comprising the terms $<t_1, t_2 \ldots, t_n>$

The document and query are formed as $<p_1, p_2, p_3 \ldots p_n>$, where p_x represents the weight of the term x in the document. The query is represented as $<q_1, q_2, q_3, \ldots, q_n>$.

Relevance of a document d to the query q is computed based on the similarity between them. There are various functions to do so. The different ways to compute the similarity measures are cosine, dot product, dice, Jaccard and so on.

Let us proceed with the matrix representation in vector space model. Figure 12.9 shows the matrix.

	Terms			
Documents	p_{11}	p_{12}	...	p_{1n}
	p_{21}	p_{22}	...	p_{2n}
	⋮			
	p_{m1}	p_{m2}	...	p_{mn}
Query	q_1	q_2	...	q_n

Information retrieval - Models

Term-frequency, document frequency and inverse document frequency: The concept of term frequency and inverse document frequency is highlighted now. Most often what is observed is when we refer to the term, the weights are the occurrence of the terms in the matrix. So, the *term frequency* is defined as the frequency of occurrence of the term in the document. It is represented as *tf*. The document frequency tells about the number of documents that have a specific term. It is represented as *df*. Inverse document frequency determines the unevenness of the term in the distribution throughout the documents. This is represented as *idf*. The weights are now computed as follows:

P_{11} = Weight of 1st term in document 1

So, to make it more generalised, we have (Remember the term frequencies are normalised by maximum frequency count of that term.)

Weight of a term in $d = wt(t, d) = tf(t, d) * idf(t)$

where $tf(t, d)$ is the occurrence of the term t in d.

$idf(t)$ = log (Total number of documents/Documents containing t term)

Probabilistic Model

Bayesian probability is the most common mechanism that is used in probabilistic inferring. The major drawback of using the Bayesian approach is that it requires prior knowledge. But it is able to address the uncertainty factor that could occur while submitting the query.

Information Extraction

In information retrieval, we have studied about getting the information—the search engines being the common example. Now, what does information extraction (IE) do? Is it not the case that information extraction and retrieval are the same? Do remember that in information retrieval, the data is not at all modified or changed. Simply, the relevant data is presented to the user. In IE, template matching is carried out. The IE module could make entries in the databases. So, there is a pre-defined fixed format in which the text entries are carried out. In short, IE makes things structured from unstructured inputs. With IE, the information from the documents is extracted into the templates. Are you feeling like there is some similarity between retrieval and extraction? IR gets relevant documents, whereas IE gets relevant information from the documents. Though there could be an overlap between them. IE could lie between IR and text understanding. Figure 12.12 depicts the basic IE system.

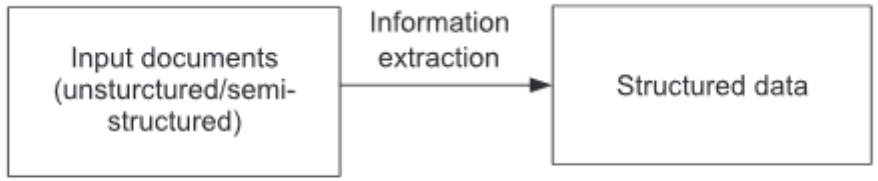


Figure 12.12 Information extraction.

Information Extraction

Let us take an example to explain IE. An employment agency has to send mails to the clients whose profile suit some ABC company's requirements. From the CVs of these clients, only the mail-id needs to be extracted to send the date for interview. In another example, a company wants to categorise candidates based on their skill set. In this case, the information about the skills is retrieved from their CVs. Here, terms like Java, .Net etc. are used for template matching. This is also a task of IE.

There are various aspects of IE such as what is to be extracted?, what is the input data? what methods are to be used?, How is the output expected? Let us try to address them. To extract, it could be any particular entity, or any specific table or relations. Whereas, the input data could be a simple set of documents, web pages and so on, which is unstructured. The methods applied can be handwritten patterns, learning-based methods, rule-based methods or statistical methods. Figure 12.13 shows the basic modules of IE.

What is the job of each of them? Tokeniser, as we are already familiar, does the word separation. In *lexical and morphological analysis*, part of speech tagging is done. Identifying the parts of speech or learning the word sense is carried out here. In *parsing*, syntax analysis is done, whereas in *domain analysis*, merging the partial results or co-reference is carried out.

FRUMP, LSP were the early IE systems. Later on, systems for web-based information extraction were Crunch, Content Seeker, etc. At present, wrappers are used on a large scale. They are called *information extraction procedures* that extract some content or bits from the text. Meta-crawlers are the type of wrappers. Currently, the focus of the research is on investigating and inventing different techniques for text summarisation as well as on cross-media and language-based IE systems.

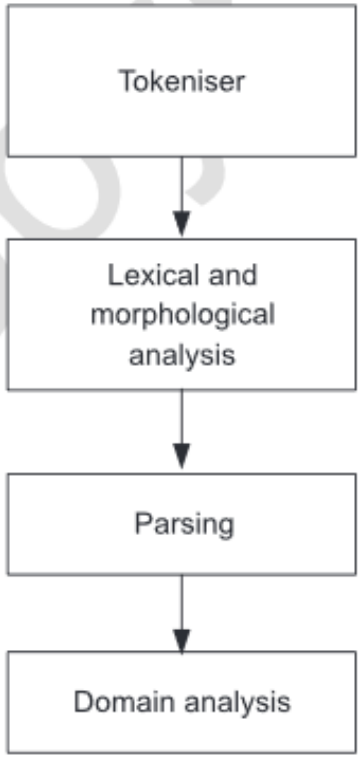


Figure 12.13 Modules of IE.

- Advance topics in Artificial Intelligence- Cloud Computing and Intelligent agent
- Business Intelligence and Analytics
- Sentiment Analysis
- Deep Learning Algorithms
- Planning and Logic in intelligent Agents

Advance topics in Artificial Intelligence- Cloud Computing and Intelligent agent

Cloud computing

The practice of using a network of remote servers hosted on the Internet to:

- store,
- manage,
- and process data,

rather than a local server or a personal computer.

Cloud computing and AI(contd.)

Cloud computing and AI

While artificial intelligence (A.I.) has struggled to gai footholds in other niches, it is finding its place in th world of cloud computing, a sort of revolution within th revolution that could rapidly change the face of businesse using cloud computing solutions over the next few years.

Cloud computing and AI(contd.)

In three areas of cloud computing, A.I. is taking long strides.

Those areas are

- Parallel processing
- Machine Learning-ML Algorithms
- Big Data

What's parallel processing and how it work in cloud

- Parallel processing means more than one microprocessor handling parts of the same overall task. Parallel processing essentially means that multiple processors shoulder the load. To have multiple processors working on the same problem at the same time, there are two big things you need:

  - Latency
  - Bandwidth

What's parallel processing and how it work in cloud(contd.)

Latency

it refers to the amount of time it takes for a processor to send results back to the system. The longer the wait, the longer it will take the entire system to process the problem.

Bandwidth

Bandwidth is a more common term, referring to how much data a processor can send in a given length of time.

ML algorithms for cloud applications

Machine learning (ML) is a type of **artificial intelligence** (**AI**) that allows software applications to become more accurate in predicting outcomes without being explicitly programmed

For cloud applications Machine Learning algorithms are built

ML algorithms for cloud applications(contd.)

ML algorithms for cloud applications involve:

- **Cognitive computing**(to combine different patterns together; i.e. voice, imagery or other such data; for mimicking human behavior)
- **Chatbots and virtual assistants** (they are getting smarter every time they have a conversation)
- **Internet of things-IoT** (It connects every potentially “smart” machine in the world to the cloud and add that massive amount of data to the conversation)

How AI uses big data

As business enterprises increasingly need a massive data-crunching champion, cloud computing companies have begun to deploy Artificial Intelligence as a service (AIaaS). Once AIaaS is deployed, it can begin crunching data at a faster rate than any single microprocessor or human mind could ever hope to compete with.

AI has not come to take over our world, but to improve the way we harness technology to make everything better.

Consider the surface of AI finally scratched. ??

RAMAPURAM
SRM
INSTITUTE OF SCIENCE & TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)

Business Intelligence and Analytics

So, how does AI actually work in the business world? let’s try to understand what artificial intelligence is and why it is so important for today’s business corporations.

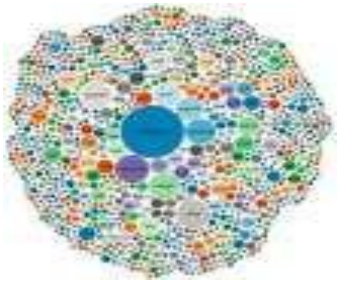
What is Business Intelligence (BI)?

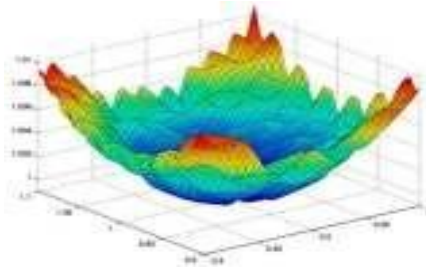
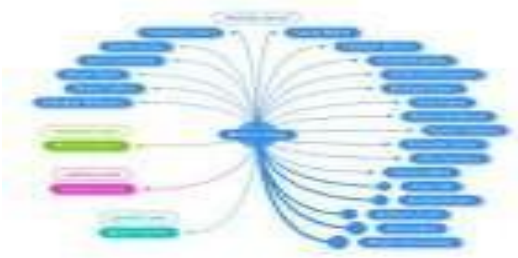
Business intelligence systems are used to maintain, optimize and streamline current operations. BI improves and maintains operational efficiency and helps businesses **increase organizational productivity**. Business intelligence software confers many benefits, notably **powerful reporting and data analysis** capabilities. Using BI's rich visualization mechanisms, managers are able to generate intuitive, readable reports that contain relevant, actionable data.

Popular business intelligence solutions include; SAP BusinessObjects, QlikView, IBM Cognos, Microstrategy, etc.

https://selecthub.com/business-intelligence/business-intelligence-vs-business-analytics/

What is Business Analytics (BA)?



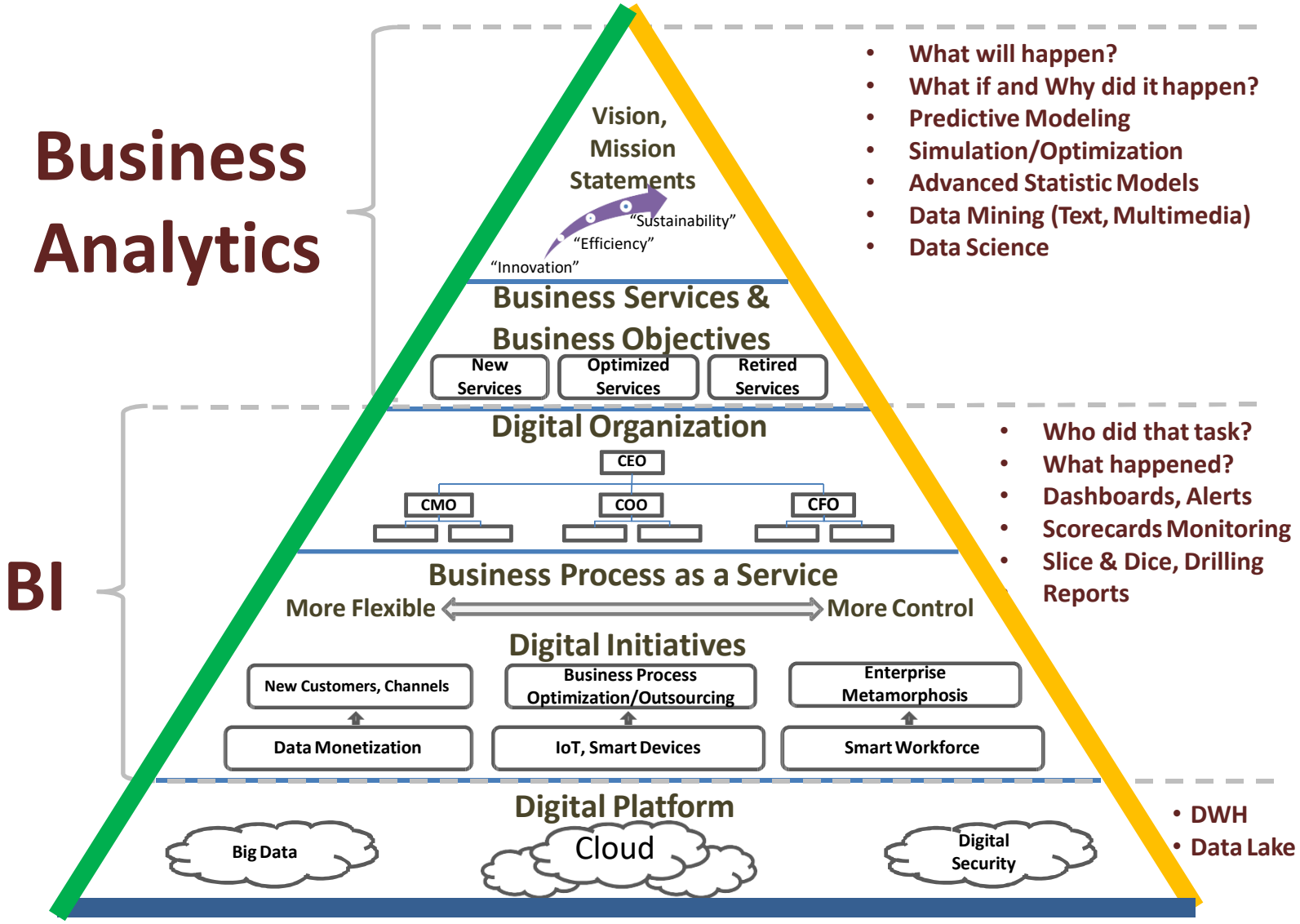



Like business intelligence, BA collects and analyzes data, employs **predictive analytics** and generates **richly visualized reports**, helping identify and address an organization's weak points. That's where similarities end. Business analytics software is used to explore and analyze historical and current data. It utilizes **statistical analysis, data mining** and quantitative analysis to identify past business trends.

Popular business analytics solutions include; SAP Business Analytics Suite, Pentaho BA, Birst BI and Tableau BIg Data Analytics.

https://selecthub.com/business-intelligence/business-intelligence-vs-business-analytics/

Business Analytics vs. Business Intelligence



Choosing between Business Intelligence (BI) and Business Analytics (BA)

While superficially similar, the difference between business intelligence vs business analytics is clear:

- BI uses past and **current data to optimize the present for current success**.
- BA uses the past and analyzes the present to **prepare businesses for the future**.

Choosing the solution for your business depends on your aims.

- If you are satisfied with your business model as a whole and mainly wish to improve operations, **increase efficiency and meet organizational goals, business intelligence may be an optimal solution**.
- If you intend to **change your business model and need to know where to start, business analytics might be the best option**.

https://selecthub.com/business-intelligence/business-intelligence-vs-business-analytics/

Choosing between Business Intelligence (BI) and Business Analytics (BA)

Business Intelligence (BI)

BI has the added advantages of targeting a business's weak areas and providing actionable solutions to those problems. Business Intelligence software is an excellent solution for managers who want to improve decision making and understand their **organization's productivity, work processes and employees. And, with that understanding, improve their business from the ground up**.

Business Analytics (BA)

If your organization is a new entity, or in the midst of significant changes, business analytics software is a serious contender. BA uses historical data, current information, and projected trends to ensure your business makes the right changes. Business analytics is the solution if you want to **analyze your company, your market, and your industry with the dual goals of optimizing current performance and predicting business trends to help you remain competitive in the future**.

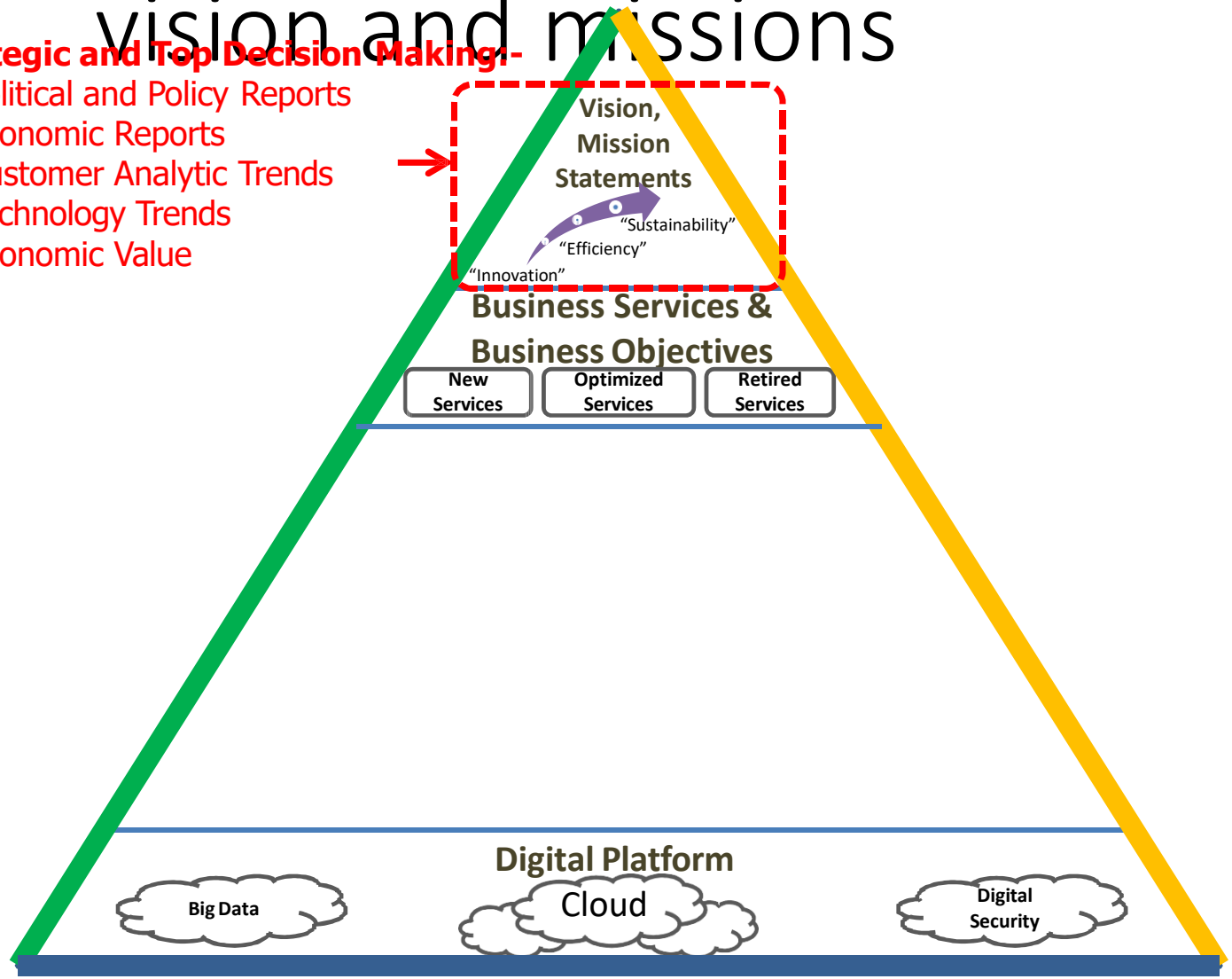
Most businesses want a combination of current success and future preparation. Alone or together, business analytics and business intelligence can help you take your business where you want it to go.

https://selecthub.com/business-intelligence/business-intelligence-vs-business-analytics/

1. (Re)Identifying your vision and missions

Strategic and Top Decision Making:-

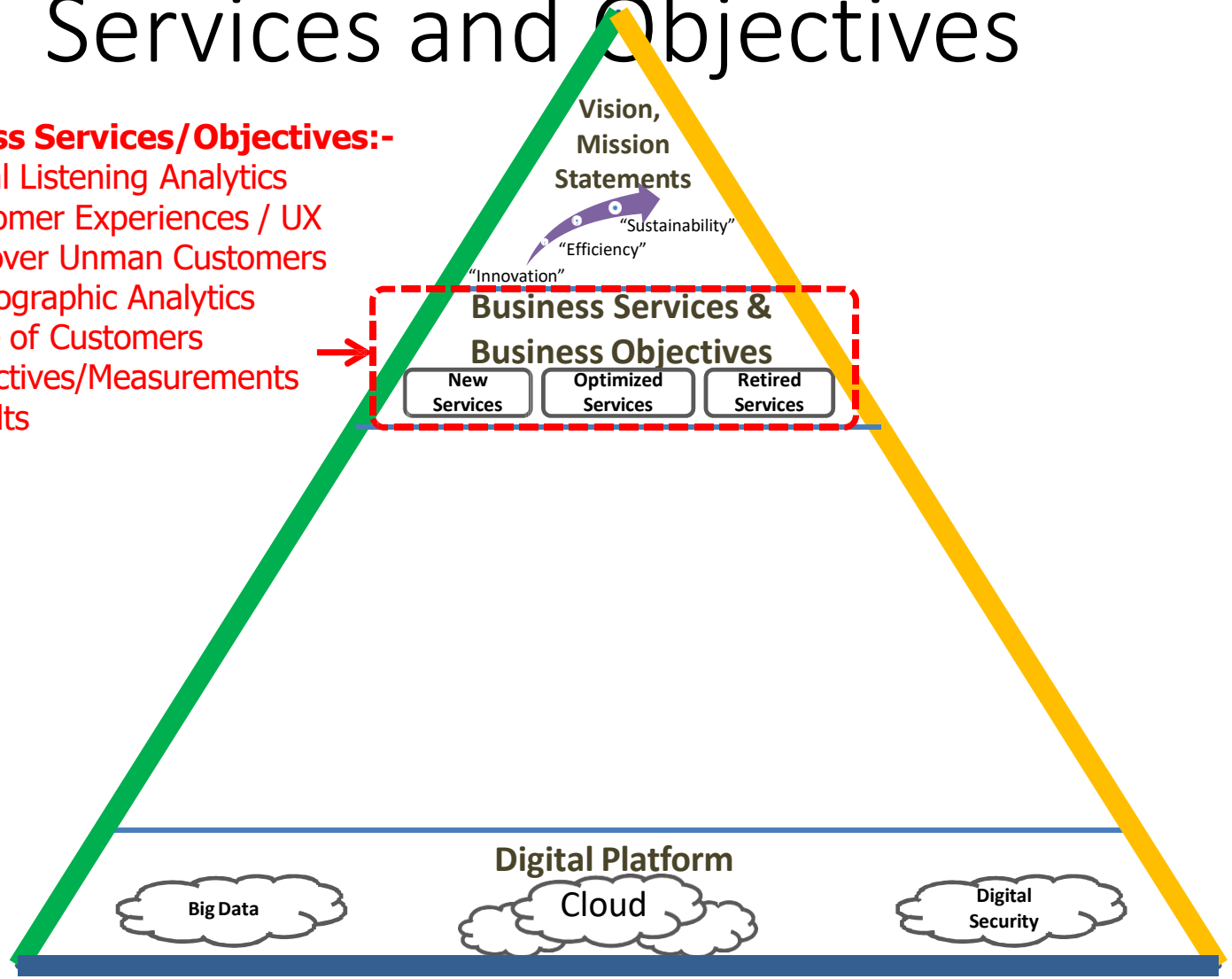
- Political and Policy Reports
- Economic Reports
- Customer Analytic Trends
- Technology Trends
- Economic Value



2. Identifying Business Services and Objectives

Business Services/Objectives:-

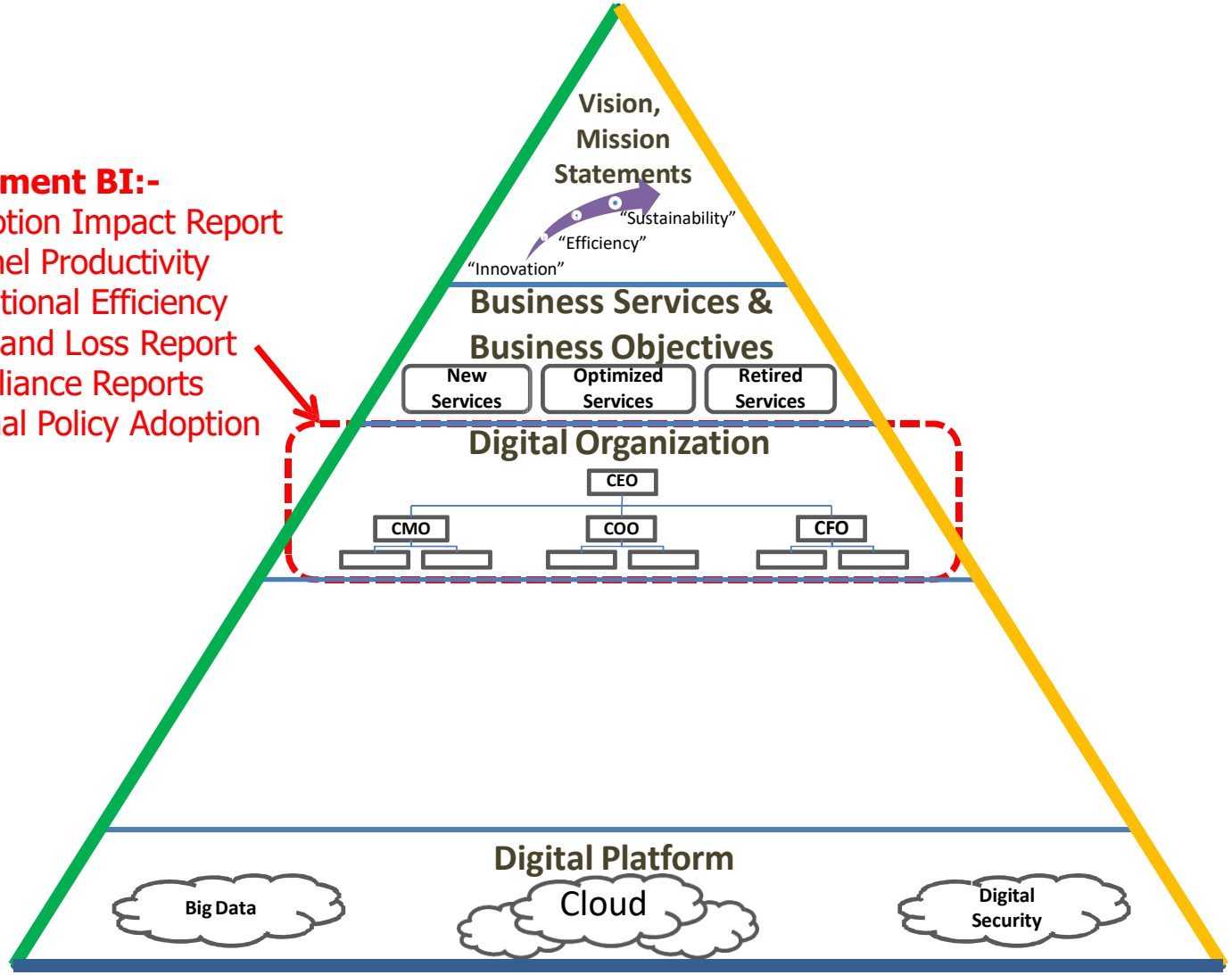
- Social Listening Analytics
- Customer Experiences / UX
- Discover Unman Customers
- Demographic Analytics
- Voice of Customers
- Objectives/Measurements Results



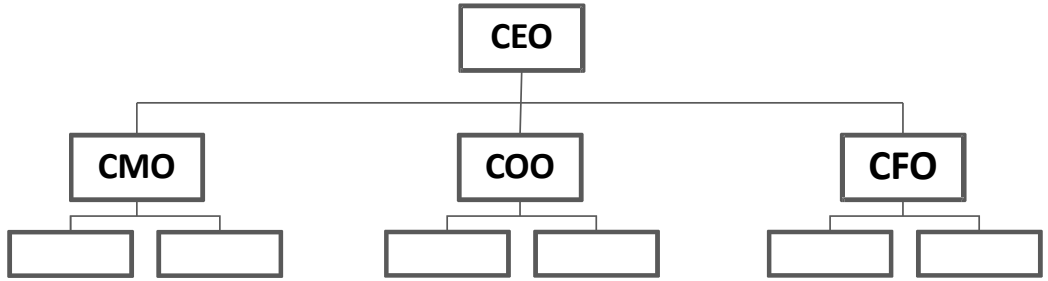
3. Identifying BI for Management Level

Management BI:-

- Promotion Impact Report
- Channel Productivity
- Operational Efficiency
- Profit and Loss Report
- Compliance Reports
- Internal Policy Adoption

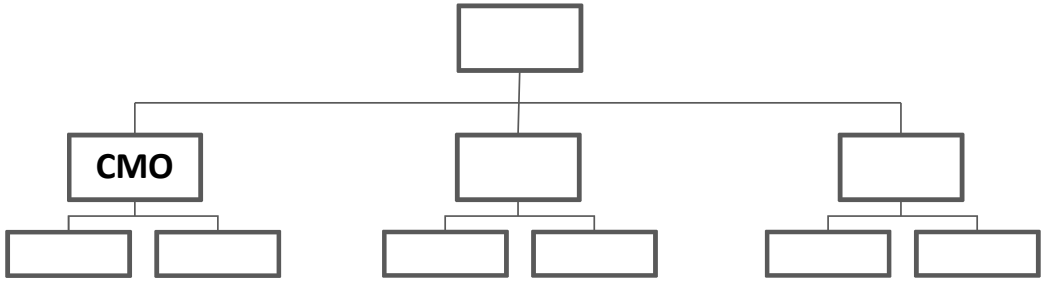


Digital Organization



- CEO : combine all successes from all C-Level
- CMO: innovation for new products offering
- COO : operation and automation
- CFO : finance, budgeting, HR, Audit, QA and IT
- Put the right skill on the right role
- **Promote paperless policy organization**

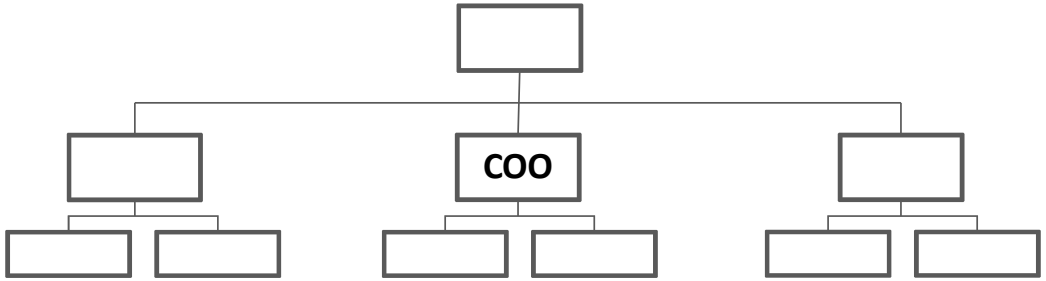
Top Business Questions from CMO



Chief Marketing Officer (CMO), Innovation, Sales and Promotion:-

- Which customers should we target?
- What has caused the change in my pipeline?
- Which are my most profitable campaigns/region?
- Did store sales spike when we advertised in the local paper or launched the campaign?
- What is the most profitable sales channel and how has that changed over time?

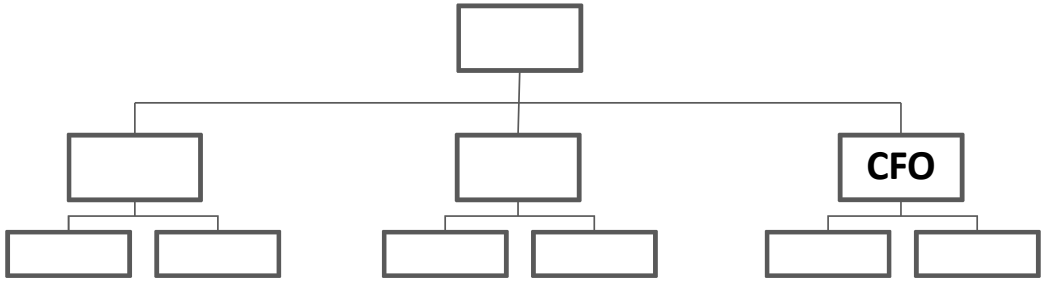
Top Business Questions from COO



Chief Operation Officer (COO):-

- Lead time and cost of production for each products
- Which order processing processes are most inefficient?
- Which vendors are best at delivering on time and on budget?–
- How many additional personnel do we need to add per branch?
- Percent of error or defect trend for each product

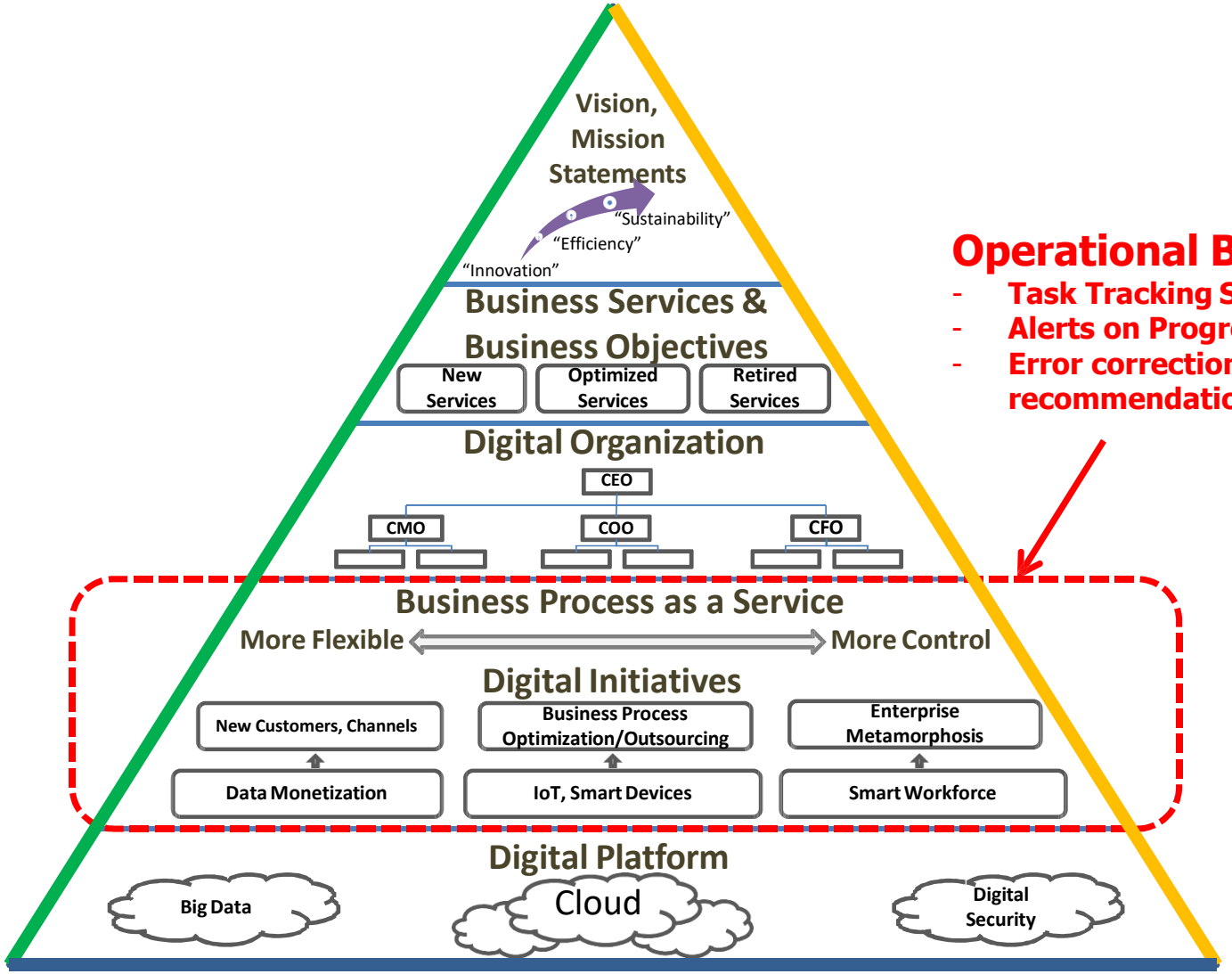
Top Business Questions from CFO



Chief Financial Officer (CFO):-

- What is the fully loaded cost of new products deployment?
- What are the current trends in cash flow, accounts payable and accounts receivable and how do they compare with plan?
- What is the expected annual profit/loss based on current marketing and sales forecasts?
- How are forecasts trending against the annual plan?

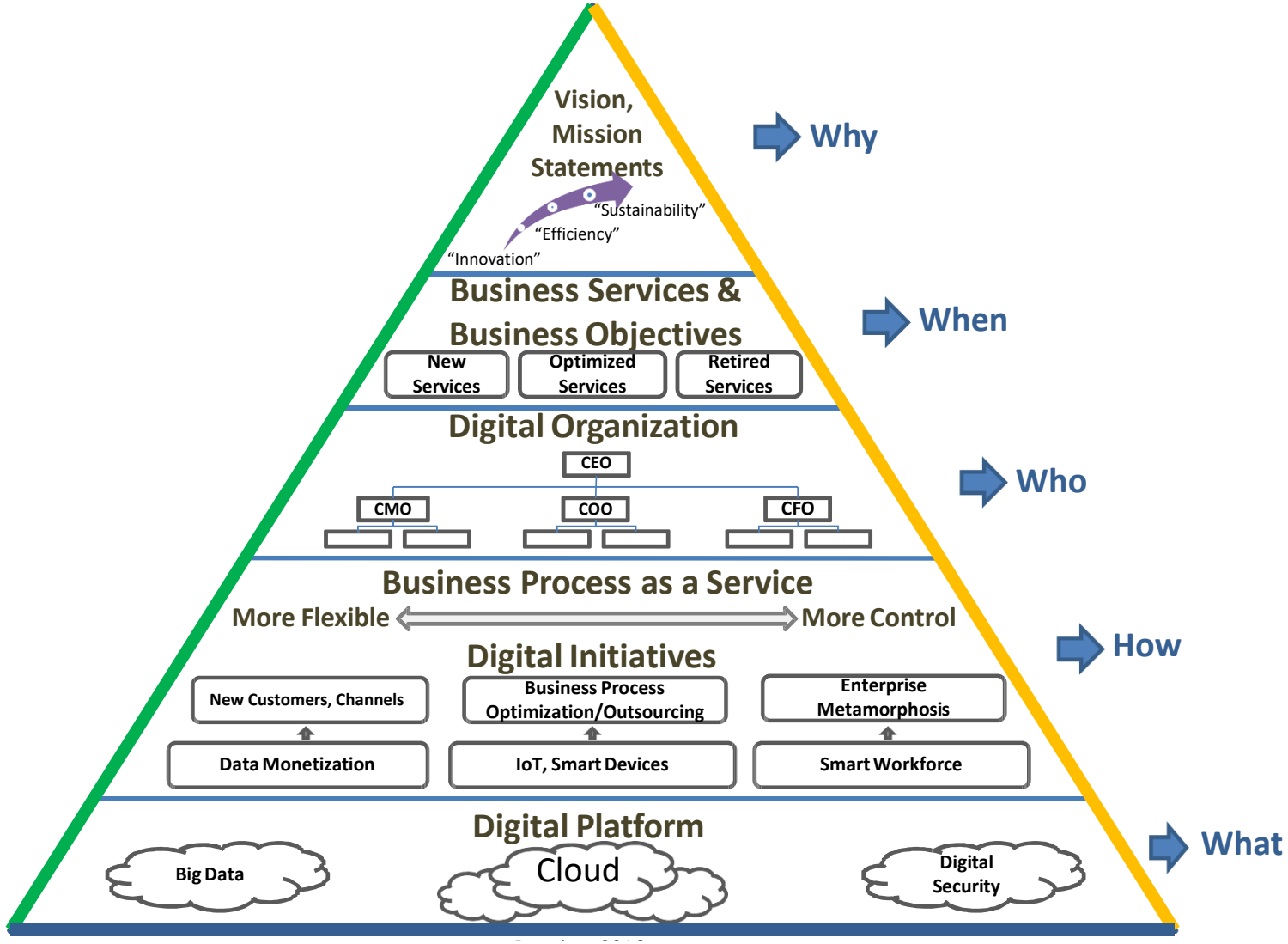
4. Identifying Operational BI



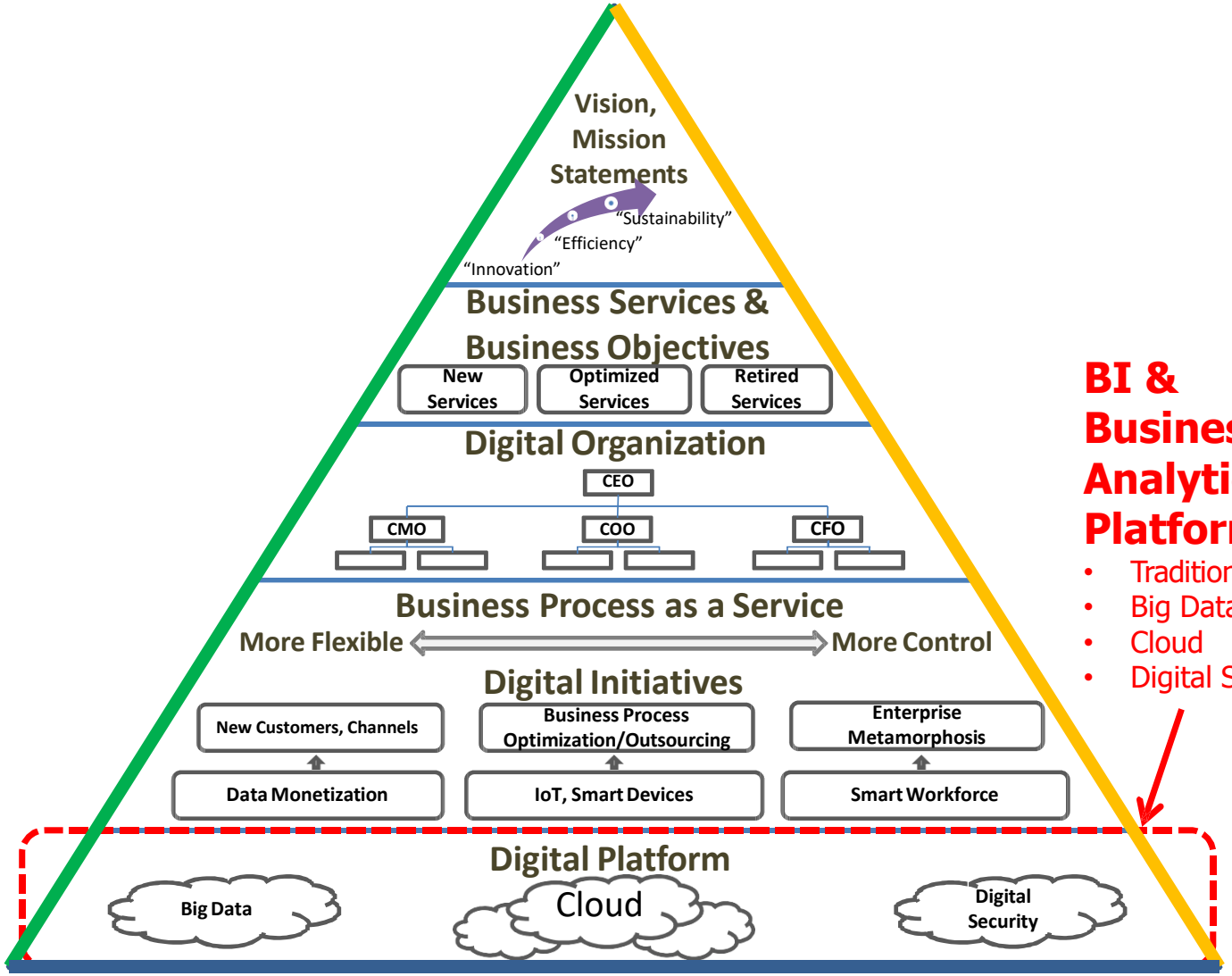
Operational BI:-

- **Task Tracking Status**
- **Alerts on Progression**
- **Error correction and recommendation**

Key Questions Type in each level of enterprise



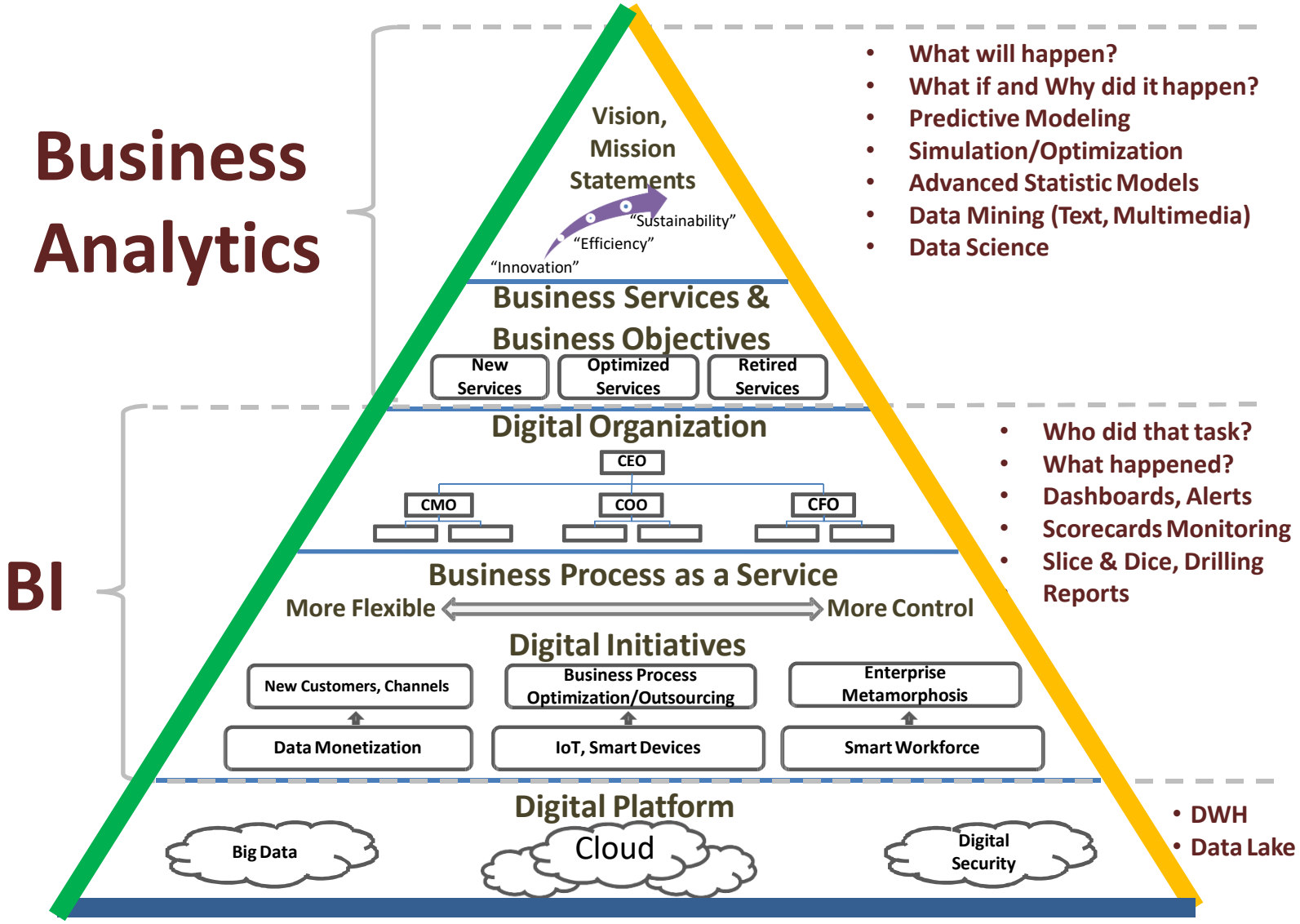
5. Identifying BI and BA Platform



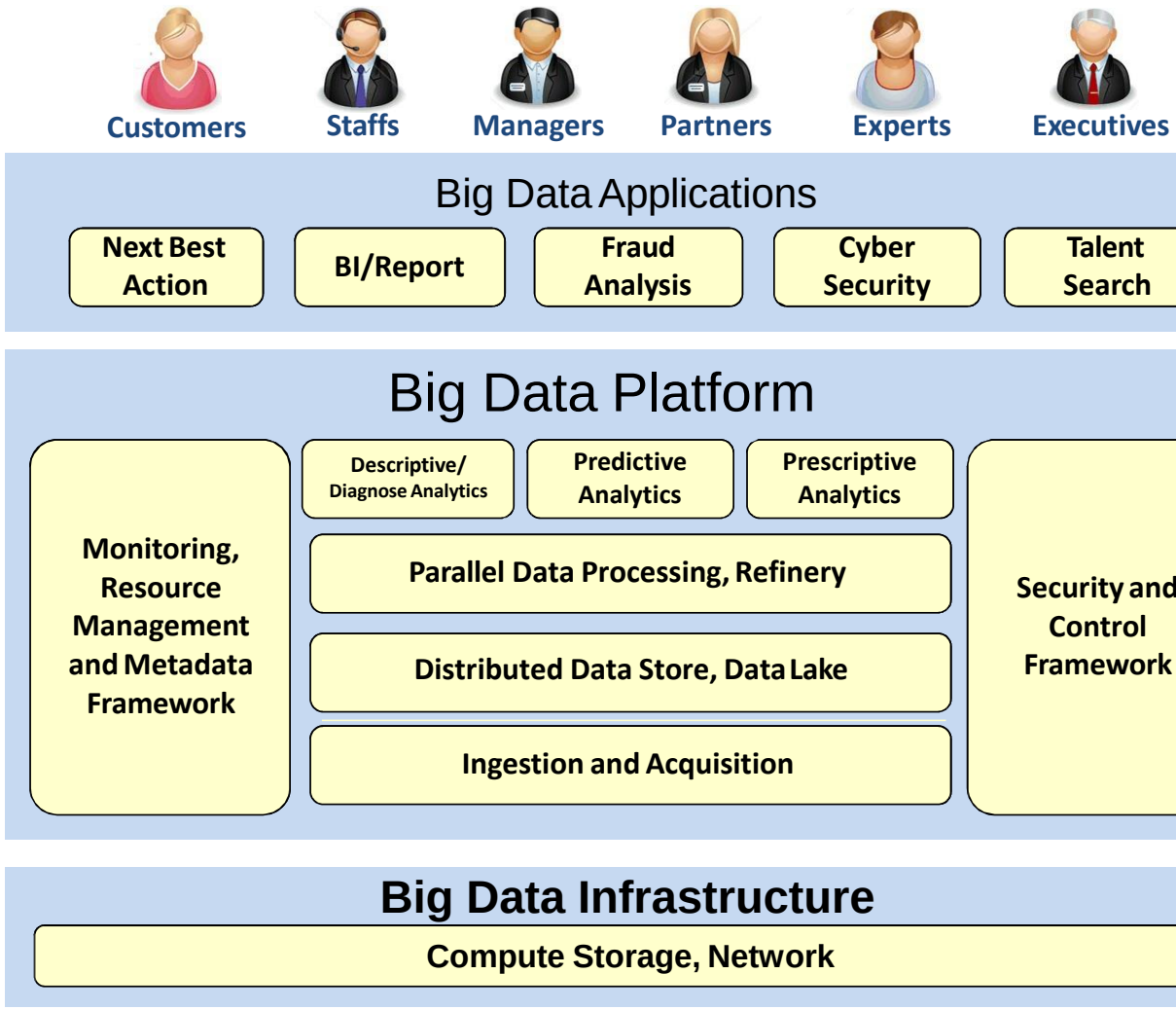
BI & Business Analytics Platforms :-

- Traditional BI
- Big Data
- Cloud
- Digital Security

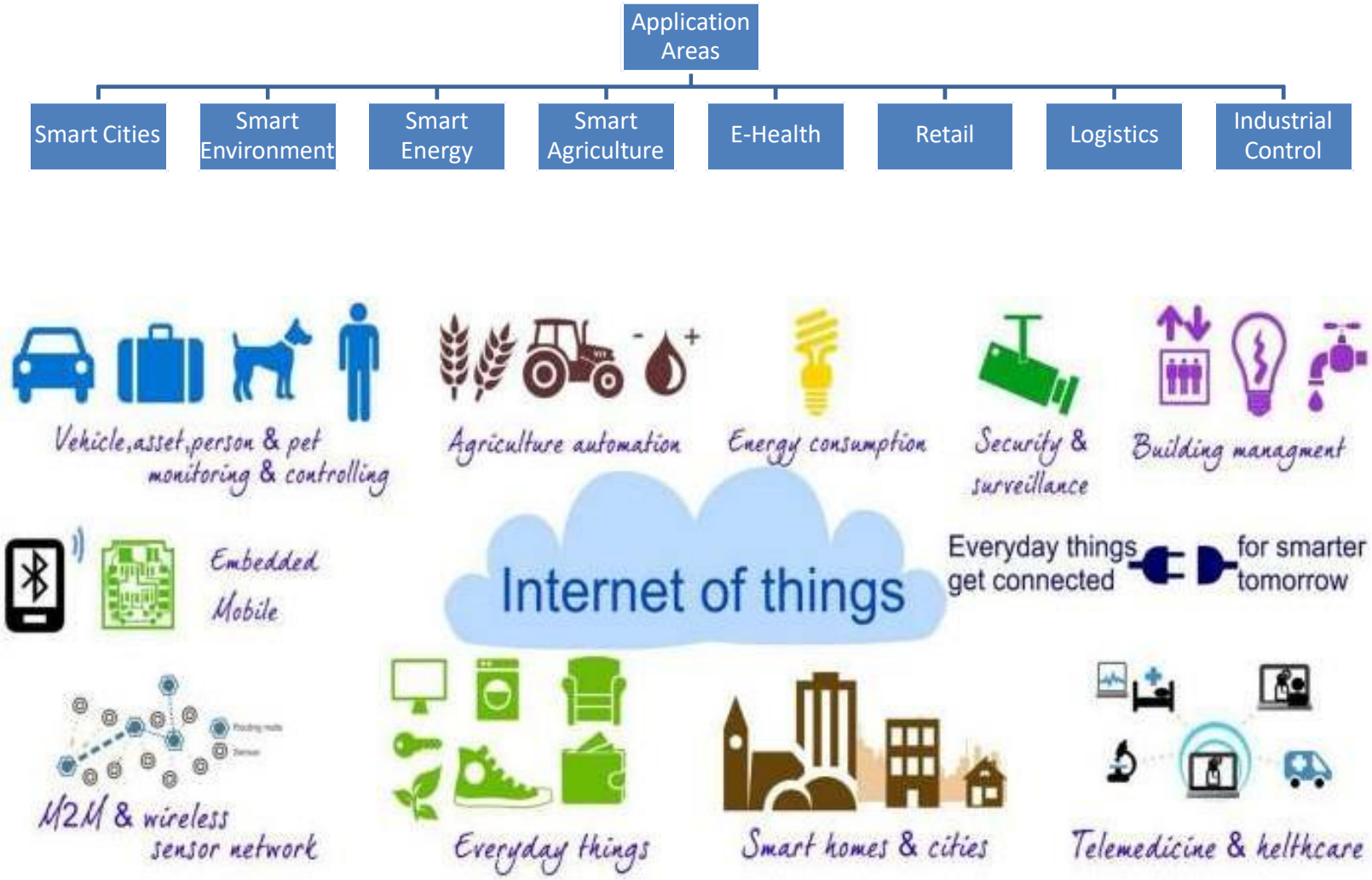
Summary


Business Analytics
BI
Vision, Mission Statements
"Sustainability"
"Efficiency"
"Innovation"
Business Services & Business Objectives
New Services
Optimized Services
Retired Services
Digital Organization
CEO
CMO
COO
CFO
Business Process as a Service
More Flexible
More Control
Digital Initiatives
New Customers, Channels
Business Process Optimization/Outsourcing
Enterprise Metamorphosis
Data Monetization
IoT, Smart Devices
Smart Workforce
Digital Platform
Big Data
Cloud
Digital Security
What will happen?
What if and Why did it happen?
Predictive Modeling
Simulation/Optimization
Advanced Statistic Models
Data Mining (Text, Multimedia)
Data Science
Who did that task?
What happened?
Dashboards, Alerts
Scorecards Monitoring
Slice & Dice, Drilling Reports
DWH
Data Lake

Big Data for Business Analytics Platform

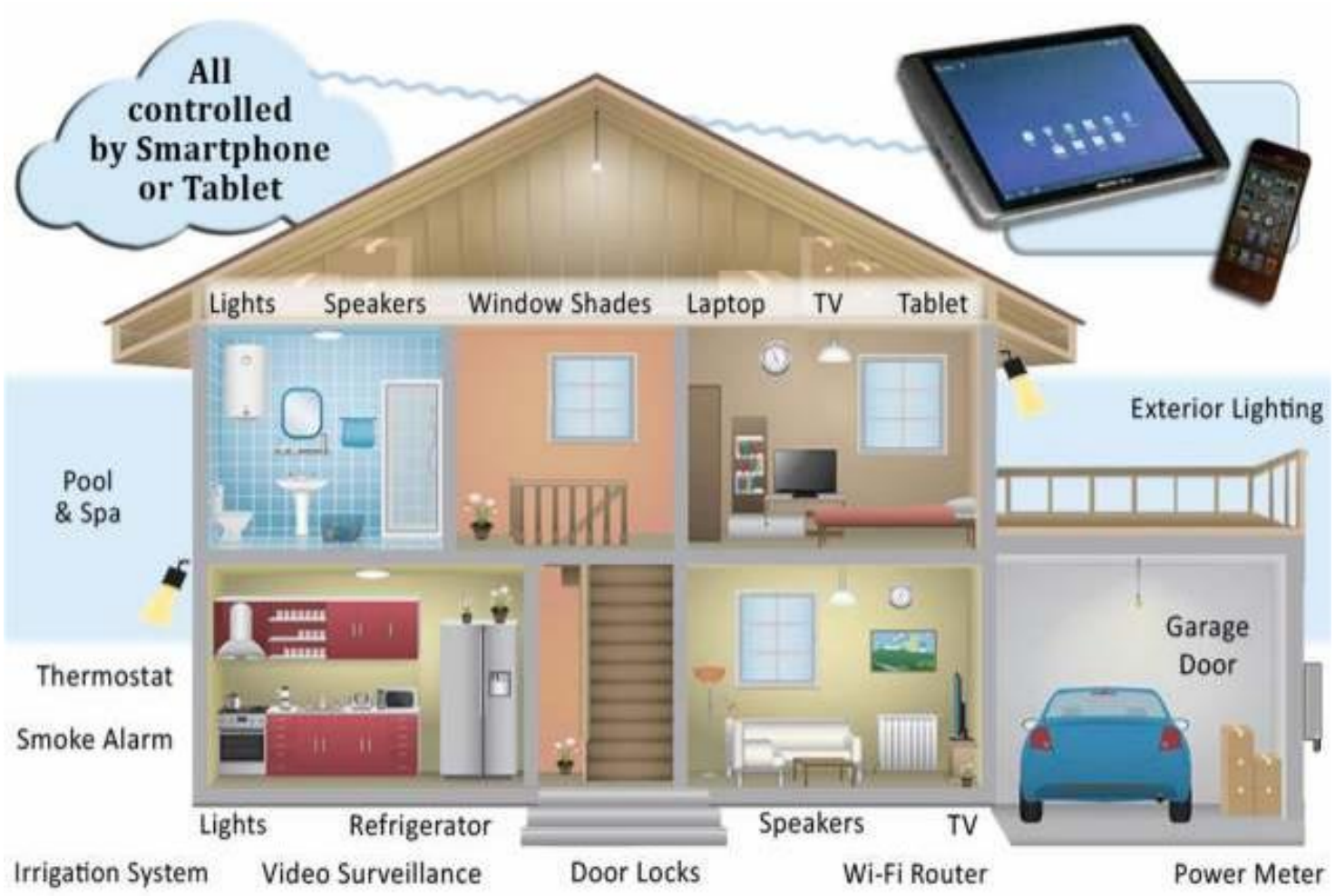


Internet of Things and Real-time Data Feeds



Monika, 2015

Smart Home



38

Monika, 2015

Based on these capabilities, we have seen multiple applications of artificial intelligence in business in the form of:

Chatbots, virtual assistants, and business intelligence bots

Targeted online advertising

Predictive analytics

Voice recognition

Pattern recognition

Sentiment Analysis

Sentiment Sentiment Analysis

A thought, view, or attitude, especially one based mainly on emotion instead of reasonSentiment Analysisaka opinion mininguse of natural language processing (NLP) and computational techniques to automate the extraction or classification of sentiment from typically unstructured text

What is Sentiment Analysis

- ✓ Sentiments are feelings, opinions, emotions, likes/dislikes, good/bad
- ✓ Sentiment Analysis is a *Natural Language Processing and Information Extraction task that* aims to obtain writer's feelings expressed in positive or negative comments, questions and requests, by analyzing a large numbers of documents.
- ✓ Sentiment Analysis is a study of human behavior in which we extract user opinion and emotion from plain text.
- ✓ Sentiment Analysis is also known as Opinion Mining.

Sentiment Analysis contd....

- ✓ It is a task of identifying whether the opinion expressed in a text is positive or negative.
- ✓ Automatically extracting opinions, emotions and sentiments in text.
- ✓ Language-independent technology that understand the meaning of the text.
- ✓ It identifies the opinion or attitude that a person has towards a topic or an object.

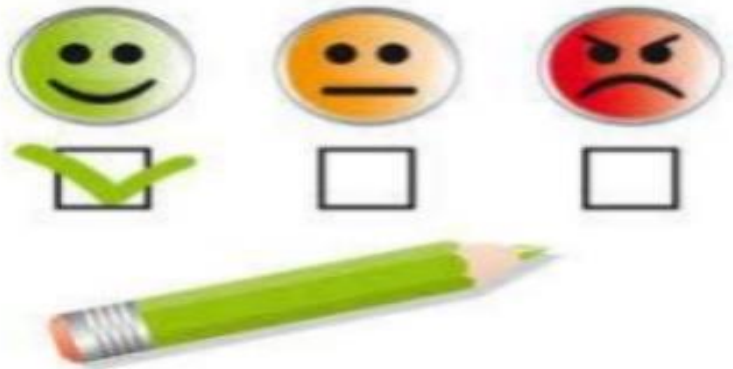
Example

- ✓ User's Opinions :

Sameer : It's a great movie (Positive statement)
Neha : Nah!! I didn't like it at all (Negative statement)

Mayur : The new iOS7 is awesome..!!!(Positive statement)

- ✓ Polarity :
  - Positive
  - Negative
  - Complex

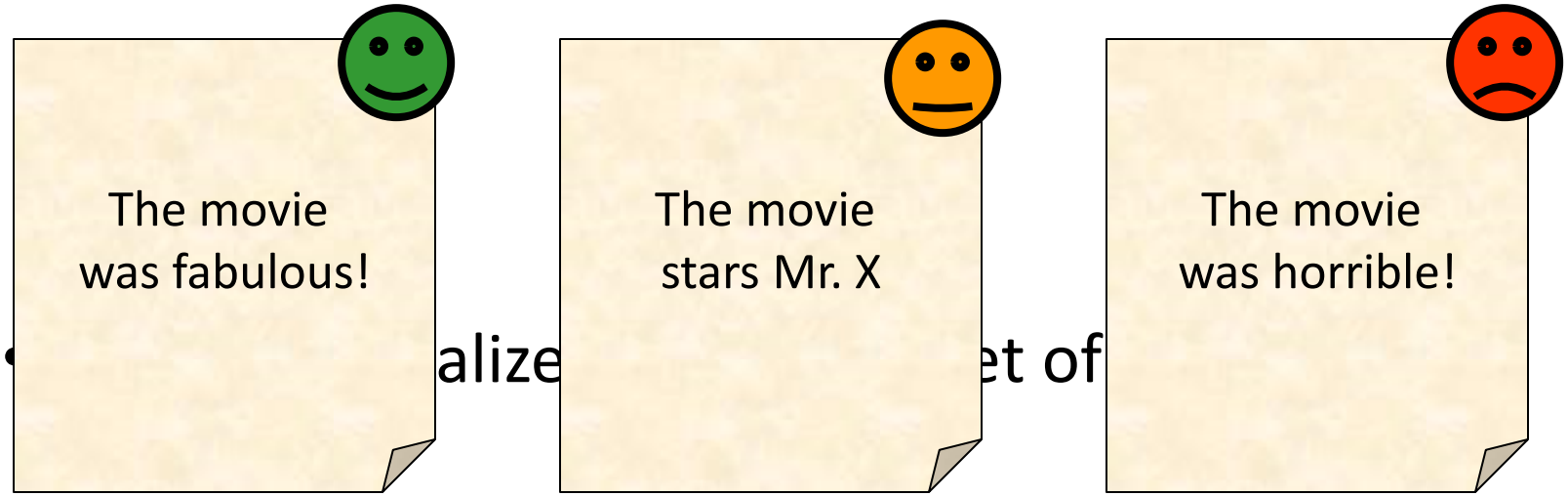


Sentiment analysis has many other names

- Opinion extraction
- Opinion mining
- Sentiment mining
- Subjectivity analysis

What is SA & OM?

- Identify the orientation of opinion in a piece of text



- alize et of

Positive or negative movie review?

- *unbelievably disappointing*

- Full of zany characters and *richly applied satire*, and some great plot twists

- this is the greatest *screwball comedy* ever filmed

- It was *pathetic*. The worst part about it was the boxing scenes.

Why sentiment analysis?

- *Movie*: is this review positive or negative?
- *Products*: what do people think about the new iPhone?
- *Public sentiment*: how is consumer confidence? Is despair increasing?
- *Politics*: what do people think about this candidate or issue?
- *Prediction*: predict election outcomes or market trends from sentiment

Why compute affective meaning?

- Detecting:
  - sentiment towards politicians, products, countries, ideas
  - frustration of callers to a help line
  - stress in drivers or pilots
  - depression and other medical conditions
  - confusion in students talking to e-tutors
  - emotions in novels (e.g., for studying groups that are feared over time)
- Could we generate:
  - emotions or moods for literacy tutors in the children's storybook domain
  - emotions or moods for computer games
  - personalities for dialogue systems to match the user

Scherer's typology of affective states

Emotion: relatively brief episode of synchronized response of all or most organismic subsystems in response to the evaluation of an event as being of major significance

angry, sad, joyful, fearful, ashamed, proud, desperate

Mood: diffuse affect state …change in subjective feeling, of low intensity but relatively long duration, ojen without apparent cause

cheerful, gloomy, irritable, listless, depressed, buoyant

Interpersonal stance: affective stance taken toward another person in a specific interaction, coloring the interpersonal exchange

distant, cold, warm, supportive, contemptuous

A]tudes: relatively enduring, affectively colored beliefs, preferences predispositions towards objects or persons

liking, loving, hating, valuing, desiring

Personality traits: emotionally laden, stable personality dispositions and behavior tendencies, typical for a person

nervous, anxious, reckless, morose, hostile, envious, jealous

Google Product Search

HP Officejet 6500A Plus e-All-in-One Color Ink-jet - Fax / copier / printer / scanner

$89 online, **$100** nearby ★★★★☆ 377 reviews

September 2010 - Printer - HP - Inkjet - Office - Copier - Color - Scanner - Fax - 250 sh

Reviews

Summary - Based on 377 reviews

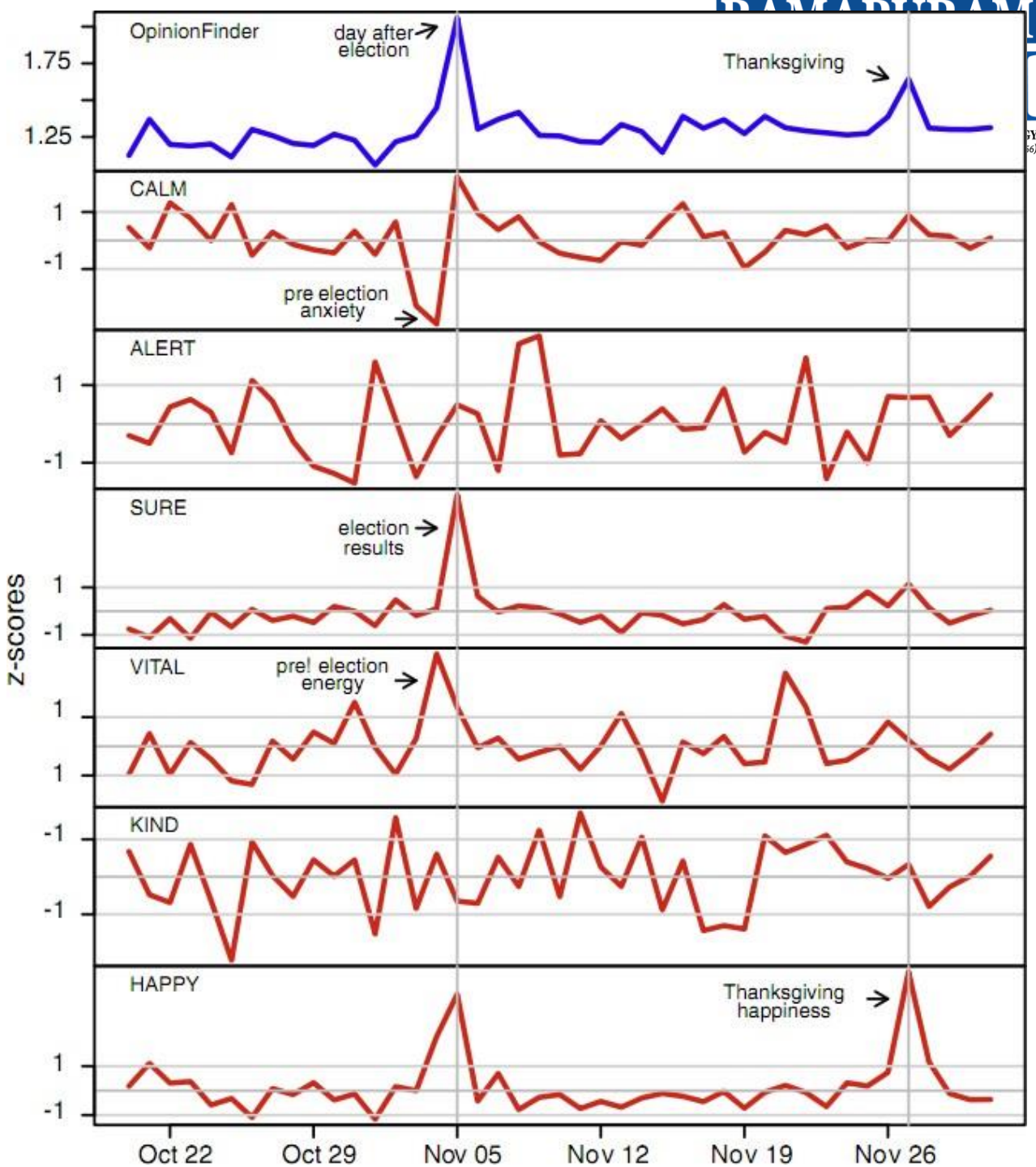
1 star	2	3	4 stars	5 stars

What people are saying

ease of use	"This was very easy to setup to four computers."
value	"Appreciate good quality at a fair price."
setup	"Overall pretty easy setup."
customer service	"I DO like honest tech support people."
size	"Pretty Paper weight."
mode	"Photos were fair on the high quality mode."
colors	"Full color prints came out with great quality."

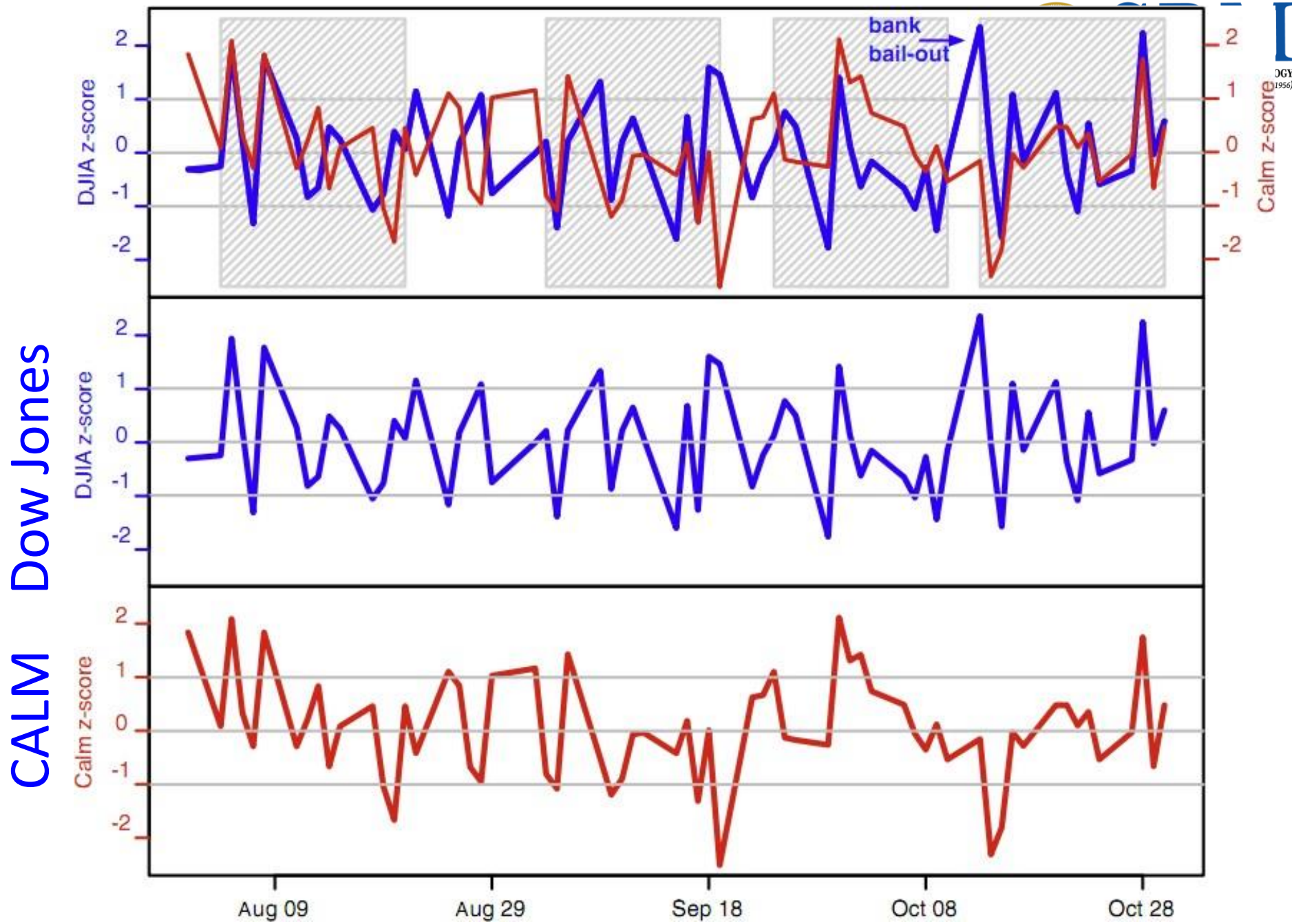
TwiGer sentiment:

Johan Bollen, Huina Mao, Xiaojun Zeng. 2011.
Twitter mood predicts the stock market,
Journal of Computational Science 2:1, 1-8
10.1016/j.jocs.2010.12.007.



Bollen et al. (2011)

- CALM predicts DJIA 3 days later
- At least one current hedge fund uses this algorithm



Applications

✓ Businesses and Organizations :

- Brand analysis
- New product perception
- Product and Service benchmarking
- Business spends a huge amount of money to find consumer sentiments and opinions.

✓ Individuals : Interested in other's opinions when...

- Purchasing a product or using a service
- Finding opinions on political topics ,movies,etc.

Applications

Clip slide

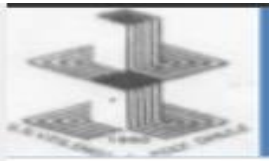
- ✓ **Social Media :**

  Finding general opinion about recent hot topics in town

- ✓ **Ads Placements :**

  Placing ads in the user-generated content

  - ❑ Place an ad when one praises a product.
  - ❑ Place an ad from a competitor if one criticizes a product.



Approach

Clip slide

- ✓ **NLP**
  - Use semantics to understand the language.
  - Uses SentiWordNet
- ✓ **Machine Learning**
  - Don't have to understand the meaning
  - Uses classifiers such as Naïve Byes, SVM, etc.

RAMAPURAM
SRM
INSTITUTE OF SCIENCE & TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)

DEEP LEARNING

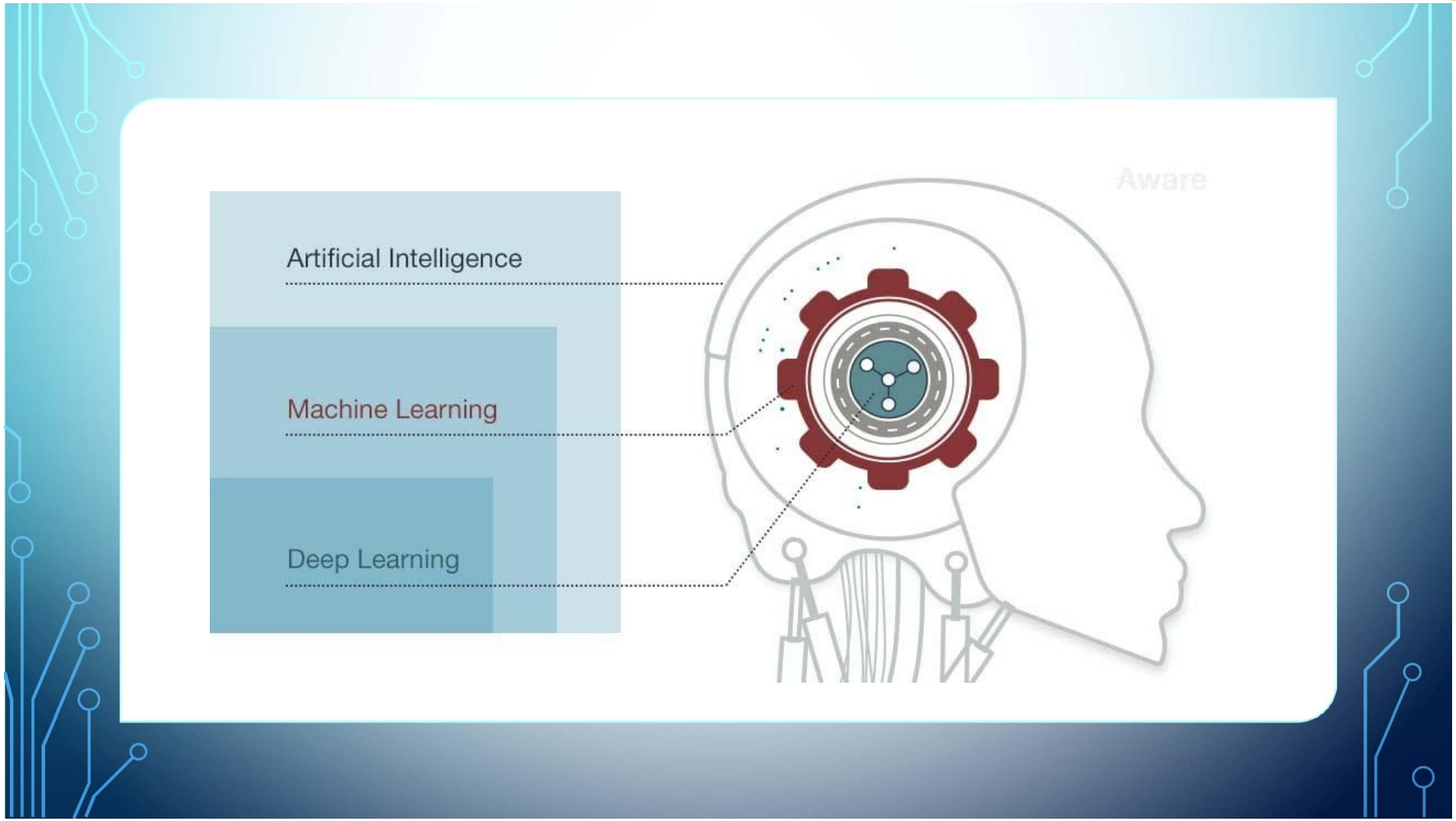
History of Deep Learning Ideas and Milestones*

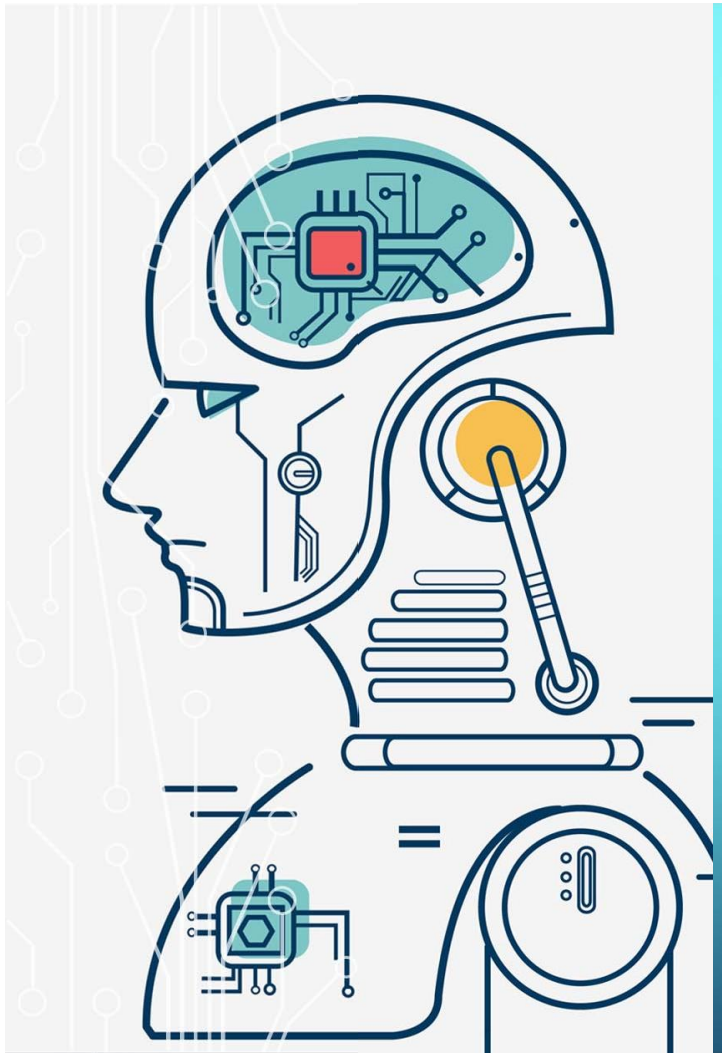
Perspective:

- Universe created
  13.8 billion years ago
- Earth created
  4.54 billion years ago
- Modern humans
  300,000 years ago
- Civilization
  12,000 years ago
- Written record
  5,000 years ago

- 1943: Neural networks
- 1957-62: Perceptron
- 1970-86: Backpropagation, RBM, RNN
- 1979-98: CNN, MNIST, LSTM, Bidirectional RNN
- 2006: "Deep Learning", DBN
- 2009: ImageNet + AlexNet
- 2014: GANs
- 2016-17: AlphaGo, AlphaZero
- 2017: 2017-19: Transformers

* Dates are for perspective and not as definitive historical record of invention or credit

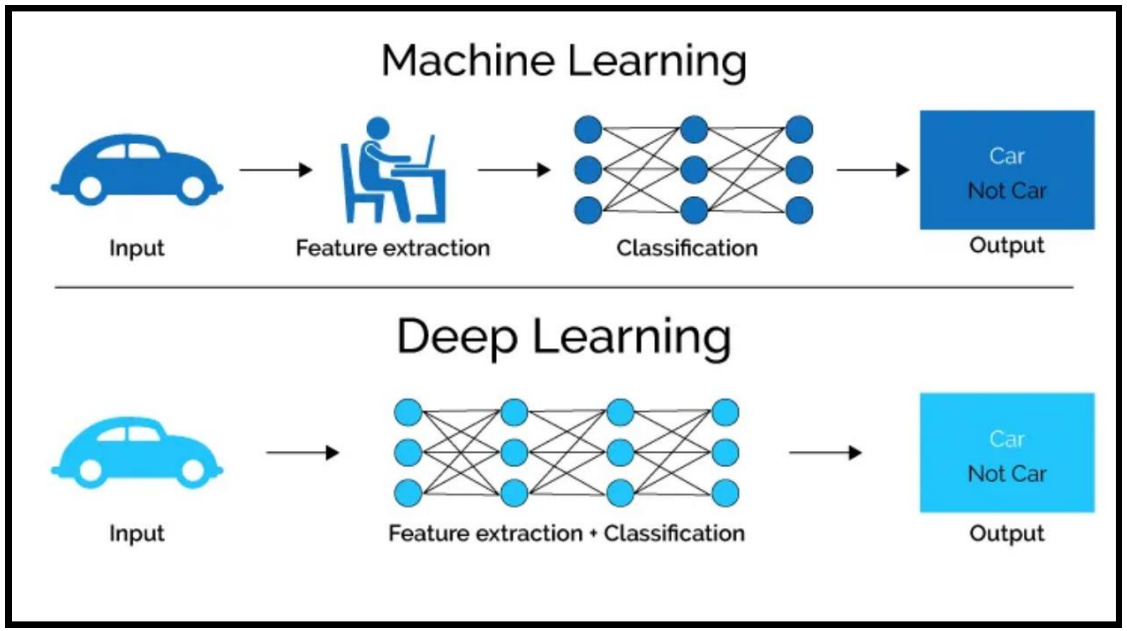

Artificial Intelligence
Machine Learning
Deep Learning



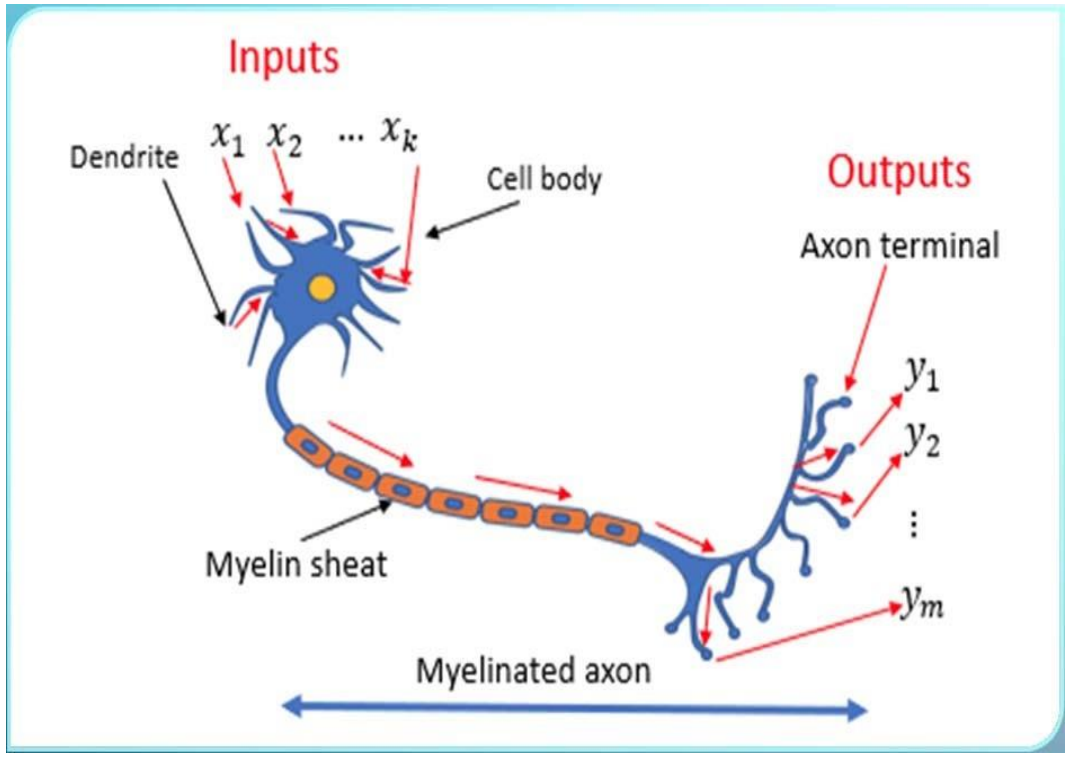
MACHINE LEARNING

- Train machine by ourselves
- Extract feature and feed to the machine then apply algorithm to train it

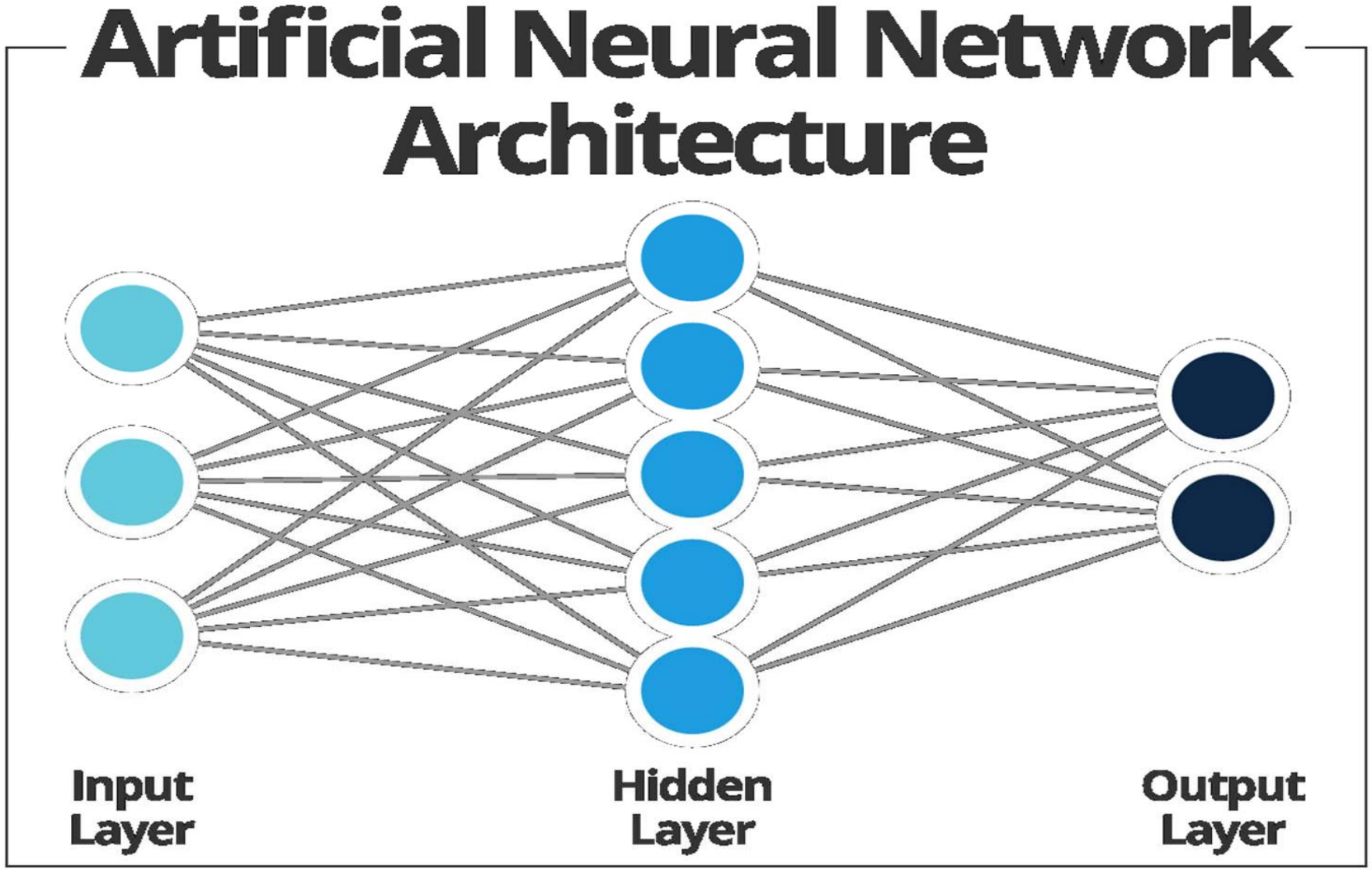
ML VS DL


Machine Learning
Input
Feature extraction
Classification
Car
Not Car
Output
Deep Learning
Input
Feature extraction + Classification
Car
Not Car
Output

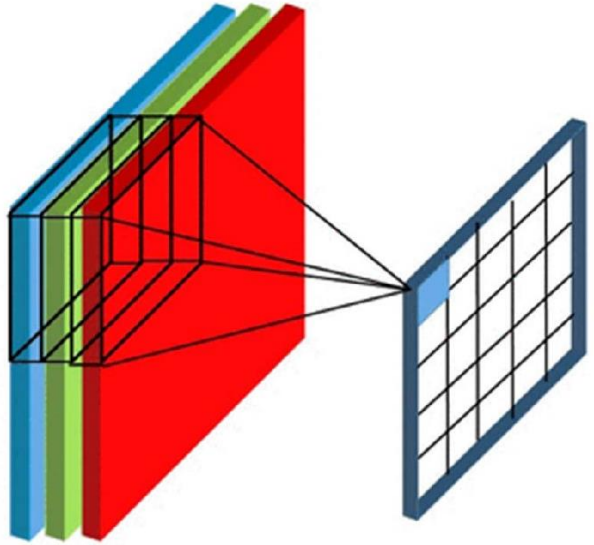
DEEP LEARNING



- It's a type of ML inspired by human brain.
- In DL, the structure is called **artificial neural network**.
- In DL, machine learns itself using **artificial neural network** that mimics biological neural network.

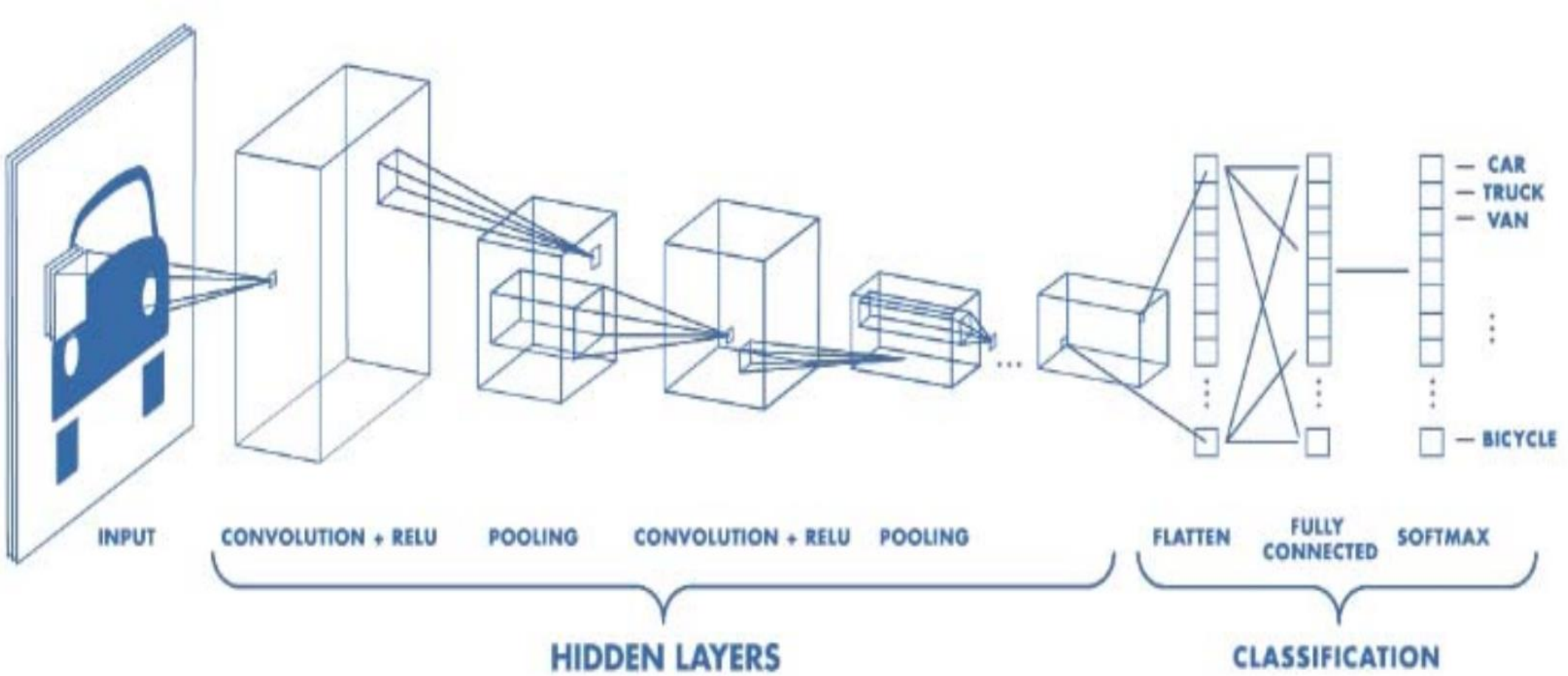

Artificial Neural Network Architecture
Input Layer
Hidden Layer
Output Layer

CONVOLUTIONAL NEURAL NETWORK



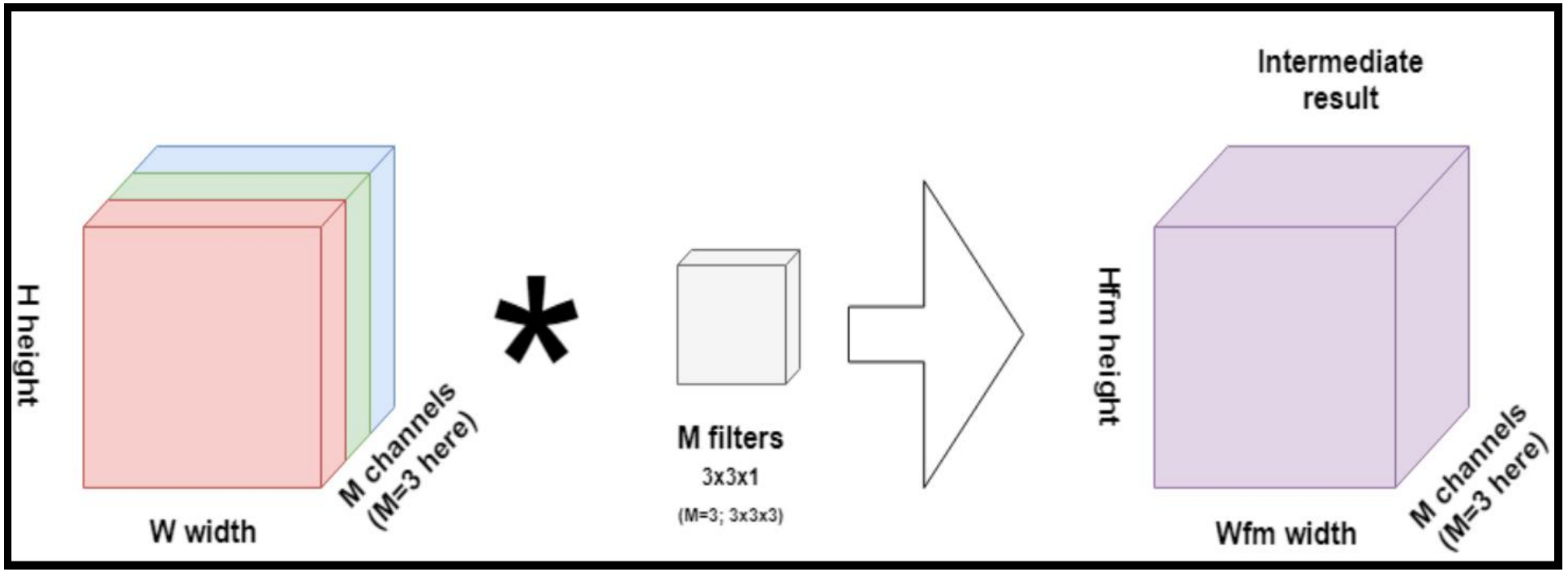
CONVOLUTIONAL NEURAL NETWORK (CNN)

- Image recognition
- Image classification
- Object detection


CAR
TRUCK
VAN
BICYCLE
INPUT
CONVOLUTION + RELU
POOLING
CONVOLUTION + RELU
POOLING
FLATTEN
FULLY CONNECTED
SOFTMAX
HIDDEN LAYERS
CLASSIFICATION

CONVOLUTION LAYER

- An image matrix (volume) of dimension (h x w x d)
- A filter (f_h x f_w x f_d)
- Out put a volume dimension



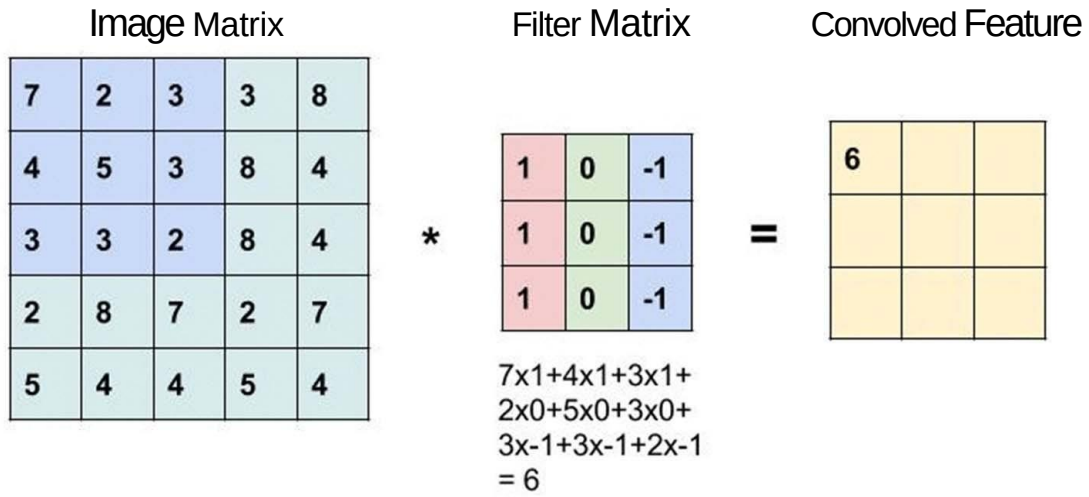
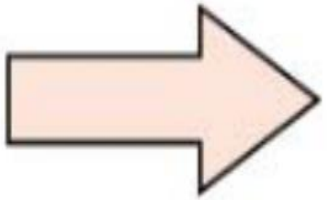


Image Matrix
Filter Matrix
Convolved Feature
7 2 3 3 8
4 5 3 8 4
3 3 2 8 4
2 8 7 2 7
5 4 4 5 4
*
1 0 -1
1 0 -1
1 0 -1
=
6
7x1+4x1+3x1+
2x0+5x0+3x0+
3x-1+3x-1+2x-1
= 6

STRIDE

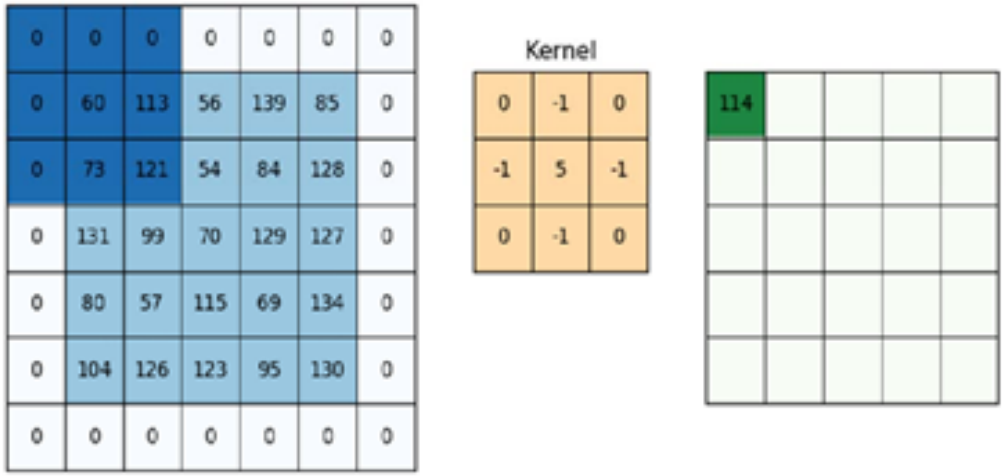
1	2	3	4	5	6	7
11	12	13	14	15	16	17
21	22	23	24	25	26	27
31	32	33	34	35	36	37
41	42	43	44	45	46	47
51	52	53	54	55	56	57
61	62	63	64	65	66	67
71	72	73	74	75	76	77

Convolve with 3x3 filters filled with ones



108	126	
288	306	

PADDING



- Pad the picture with zero-padding so that it fits.
- Drop part where image did not if. This called **valid padding** which keep only valid part of image.

SEPARABLE CONVOLUTION LAYER

- The spatial separable convolution is so named because it deals primarily with the **spatial dimensions** of an image and kernel: the width and the height. (The other dimension, the "depth" dimension, is the number of channels of each image).