Artificial Intelligence Techniques Using Structured Health Records

John Chelsom
john.chelsom@seveninformatics.com

Artificial Intelligence

We are currently at the start of the third wave of the discipline of Artificial Intelligence.

The *first wave*, and the birth of the discipline, started with the Dartmouth Conference in 1956, organised by John McCarthy and Claude Shannon, and attended by early pioneers such as Marvin Minsky, Allen Newell and Herb Simon.

The *second wave*, in the mid 1980's, saw the emergence of 'expert systems' and 'knowledge-based systems' as a paradigm for reasoning with declarative, rather than procedural knowledge.

Although the technology for pattern recognition and machine learning continued to develop throughout the two decades around the turn of the century, the widespread deployment of knowledge-based systems was hampered by a lack of robust, large-scale declarative knowledge bases and structured observational data.

The *third wave* of AI is driven by advances in speech recognition, natural language processing, speech synthesis, pattern matching and machine learning for mining unstructured or semi-structured Big Data sets.

Artificial Intelligence Techniques Using Structured Health Records

In this seminar we are going to buck the trend of the Third Wave and look at AI techniques using structured health records and knowledge bases.

- Structured Health Records
- Knowledge and Reasoning
- Artificial Intelligence

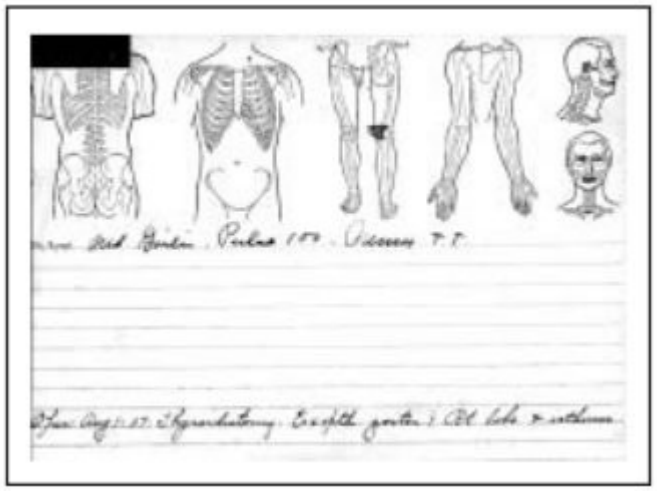
Artificial Intelligence Techniques Using Structured Health Records

- Linking Health Records with Knowledge Sources
- Rules-based Reasoning
- Bayesian Inference

Structured Health Records

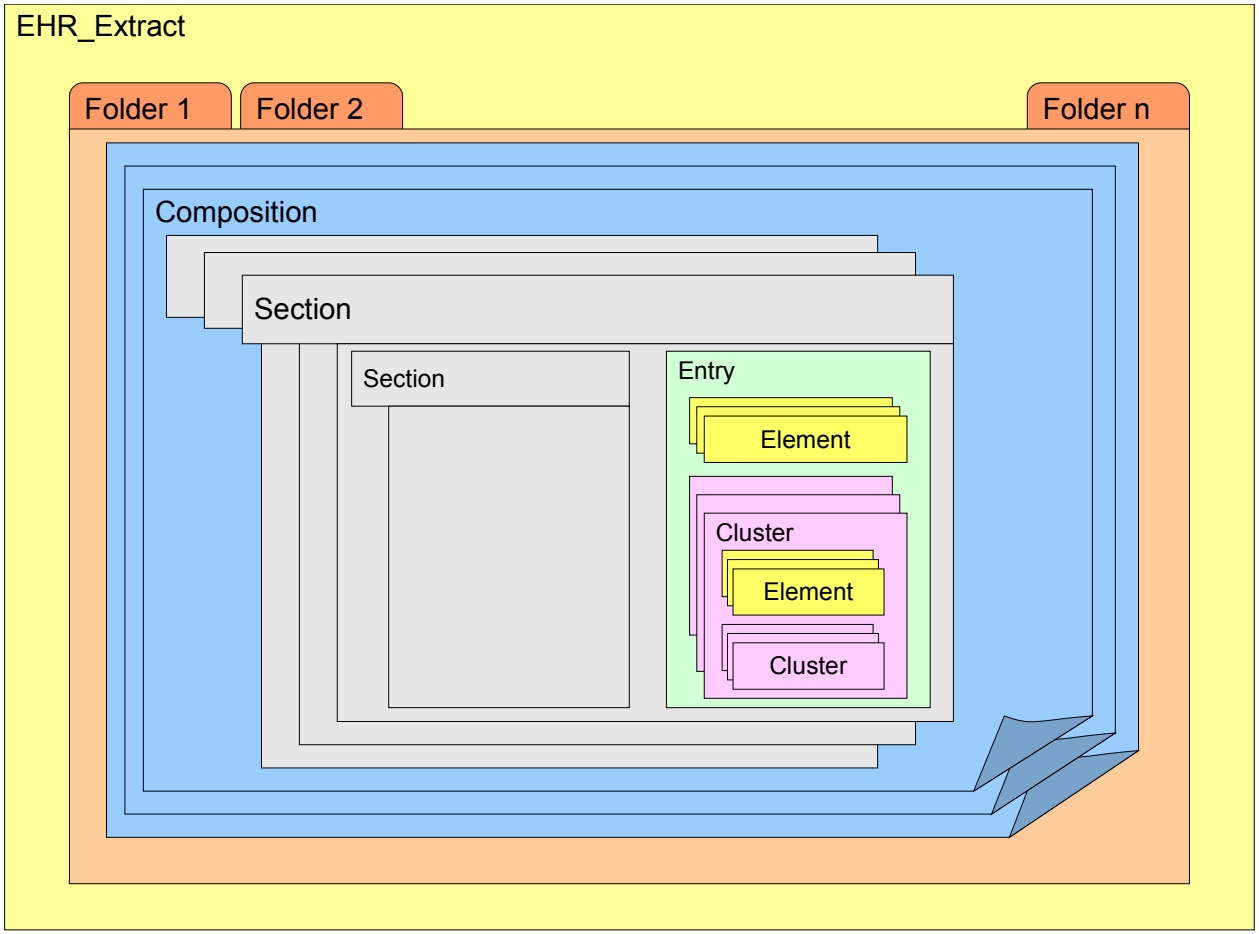
Structure of a Health Record

- ISO-13606 standard
- Started as a European standards activity, adopted by ISO
- Defines the general structure of a health record (amongst other things)

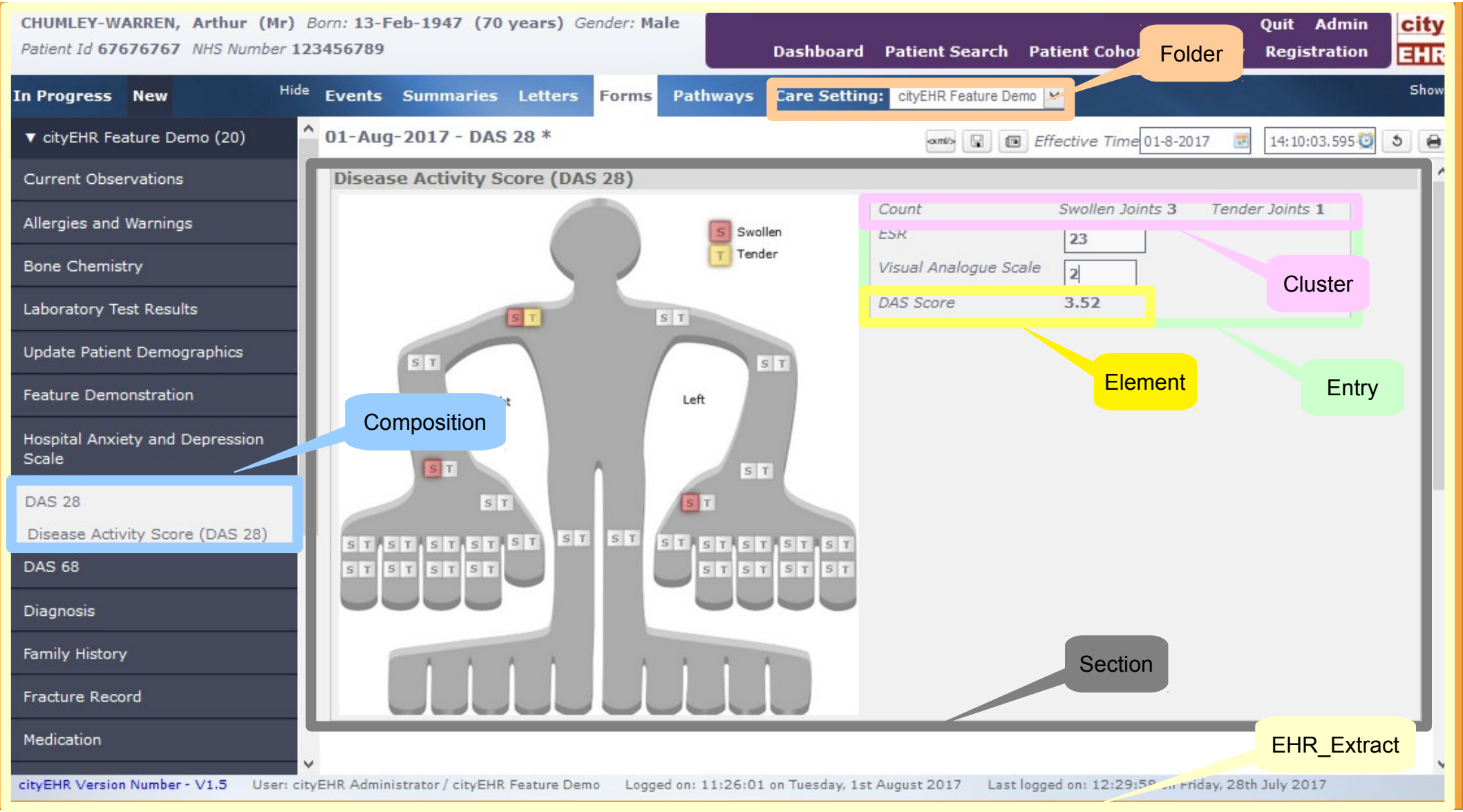


The organisation of patient records as paper 'dossiers' was introduced by Henry S Plummer at the Mayo Clinic in 1907.

Camp et al, 2008.



Structure of a Health Record

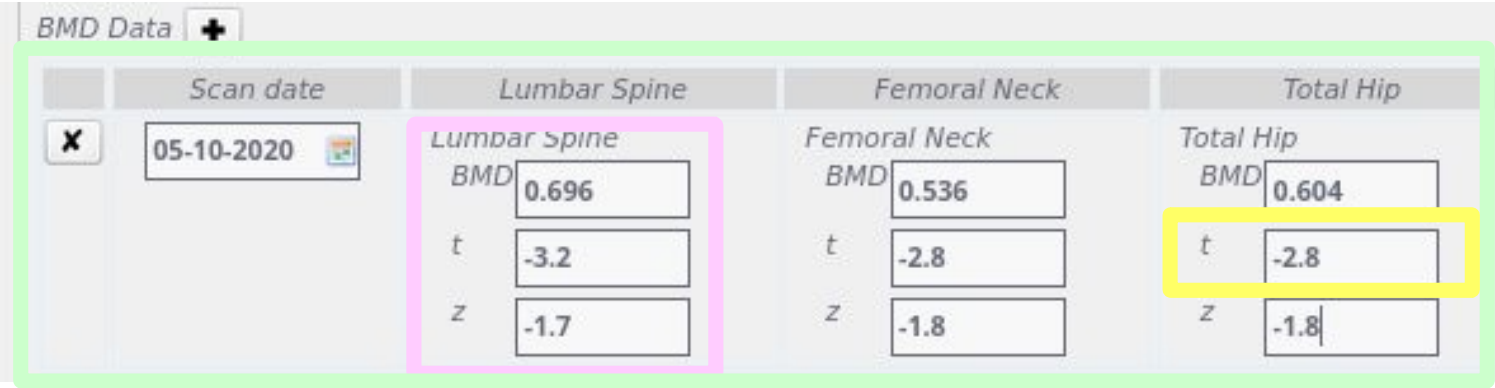


Clinical Statement

0.604 **05-10-2020** **-2.8**

These data carry no meaning.

In ISO-13606, data items are modelled as Element, and grouped as Cluster



The Clinical Statement is the lowest level of modelling that carries useful clinical context.

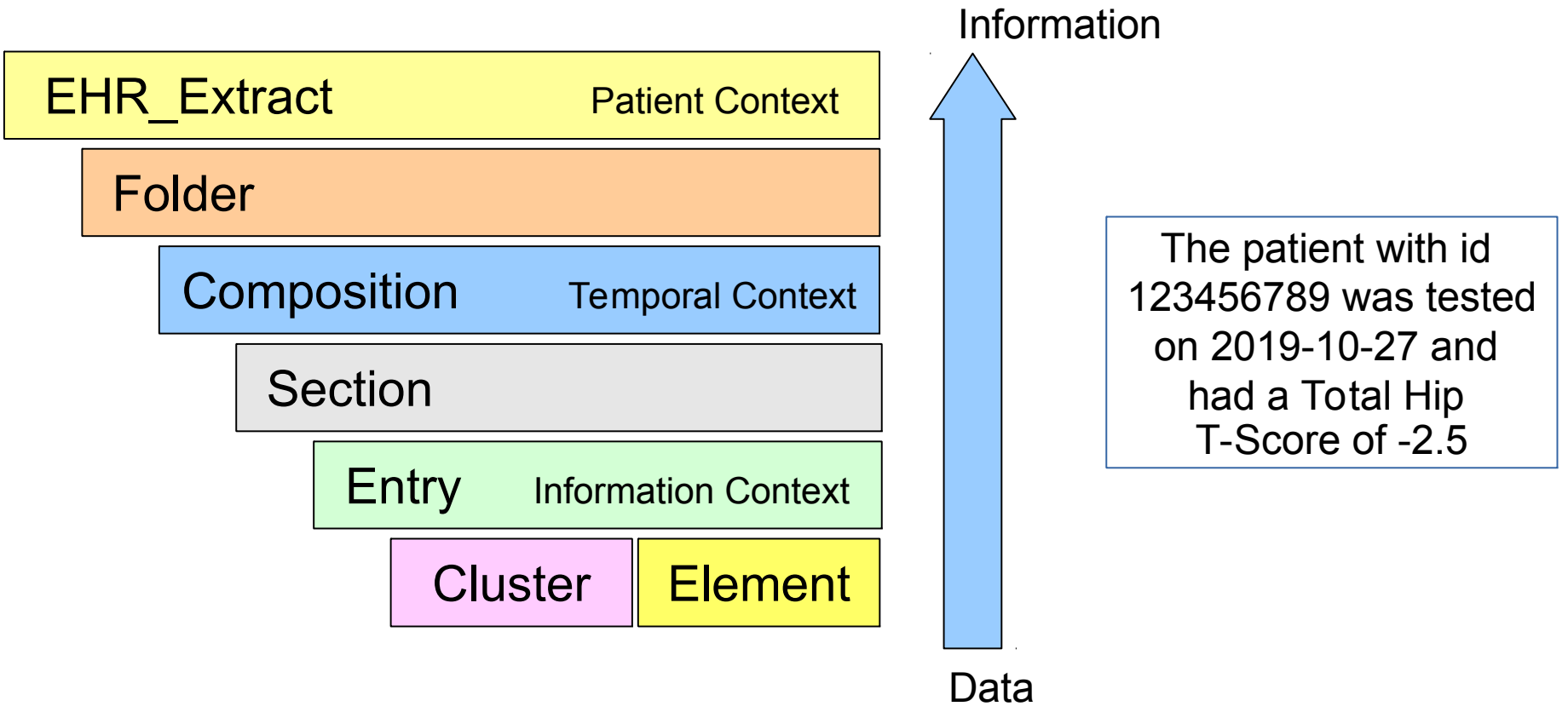
It is the point in the model at which data becomes information.

In ISO-13606 the Clinical Statement is modelled as Entry.

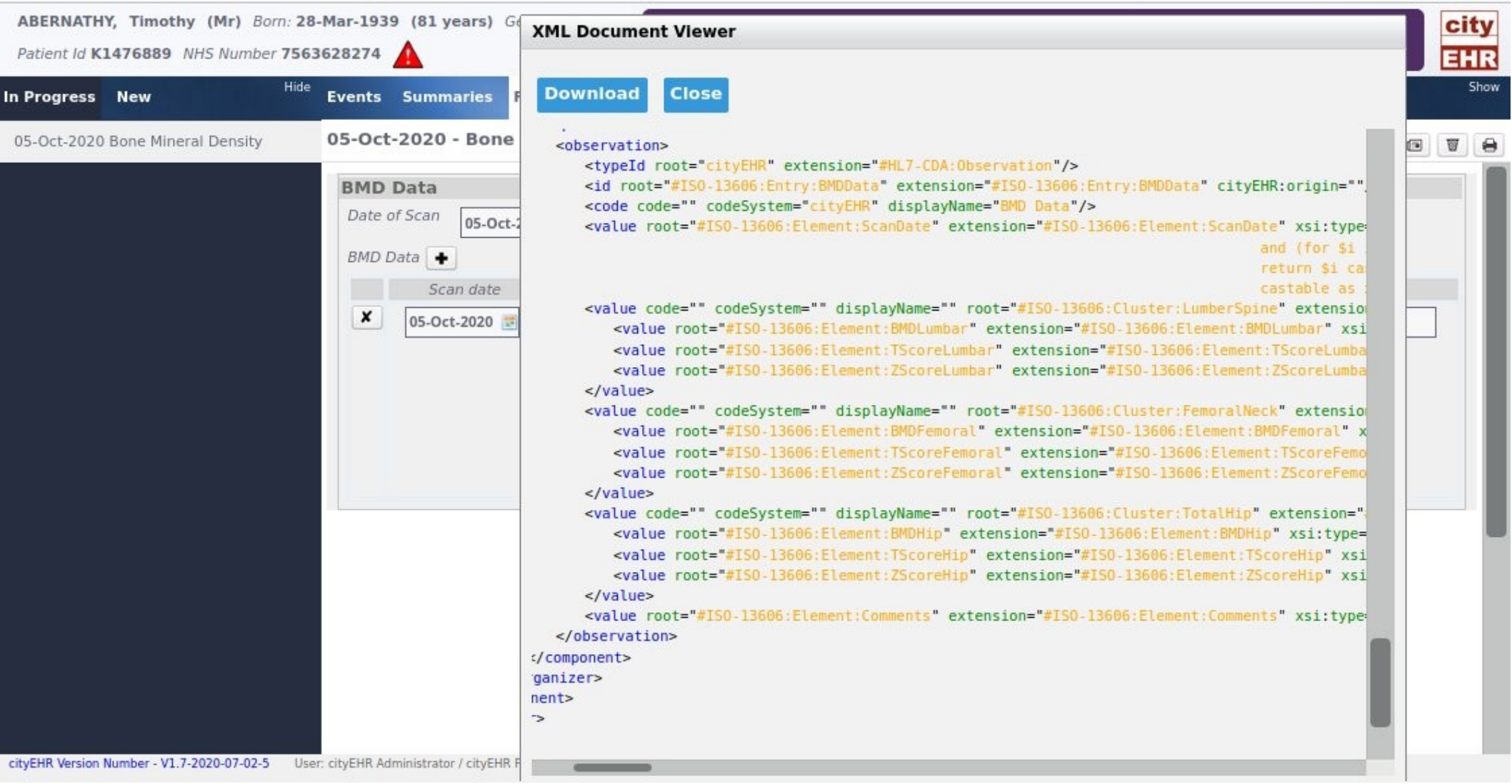
Clinical Context

Different levels of the ISO-13606 model carry different aspects of clinical context.

Other levels in the model help clinical users to interact with and understand the information, but are of no clinical significance (i.e. do not contribute to the clinical context)



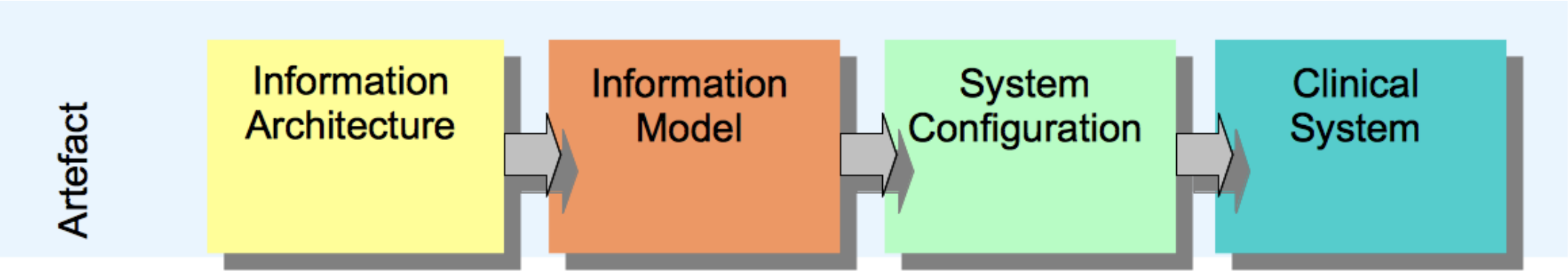
HL7 Clinical Document Architecture



Dolin, R.H., Alschuler, L., Boyer, S., Beebe, C., Behlen, F.M., Biron, P.V. and Shabo, A., 2006. HL7 clinical document architecture, release 2. Journal of the American Medical Informatics Association, 13(1), pp.30-39.

Model-Driven EHR

- A base Information Architecture provides the building blocks
- Clinicians create clinical Information Models for their own systems
- The System Configuration is generated from the Information Model
- The runtime EHR system operates using the System Configuration
  - views of the patient record
  - search criteria
  - data collection forms
  - clinical messages
  - (pathways, orders, prescriptions)



Based on the concept of Archetypes introduced by Beale (2002) and included in ISO 13606

EHR Ontology Model

```
<Declaration>
   <Class IRI="#ISO-13606:Entry"/>
</Declaration>
<Declaration>
   <NamedIndividual IRI="#ISO-13606:Entry:BMDData"/>
</Declaration>
<ClassAssertion>
   <Class IRI="#ISO-13606:Entry"/>
   <NamedIndividual IRI="#ISO-13606:Entry:BMDData"/>
</ClassAssertion>
<DataPropertyAssertion>
   <DataProperty IRI="#hasDisplayName"/>
   <NamedIndividual IRI="#ISO-13606:Entry:BMDData"/>
   <Literal xml:lang="en-gb">DXA Scan Results</Literal>
</DataPropertyAssertion>
```

```
<ObjectPropertyAssertion>
   <ObjectProperty IRI="#hasContent"/>
   <NamedIndividual IRI="#ISO-13606:Entry:BMDData"/>
   <NamedIndividual IRI="#ISO-3606:Element:BMDMeasurement"/>
</ObjectPropertyAssertion>
<ObjectPropertyAssertion>
   <ObjectProperty IRI="#hasContent"/>
   <NamedIndividual IRI="#ISO-13606:Entry:BMDData"/>
   <NamedIndividual IRI="#ISO-13606:Element:TScore"/>
</ObjectPropertyAssertion>
<ObjectPropertyAssertion>
   <ObjectProperty IRI="#hasContent"/>
   <NamedIndividual IRI="#ISO-13606:Entry:BMDData"/>
   <NamedIndividual IRI="#ISO-13606:Element:ZScore"/>
</ObjectPropertyAssertion>
```

Information model specified by assertions using the Web Ontology Language

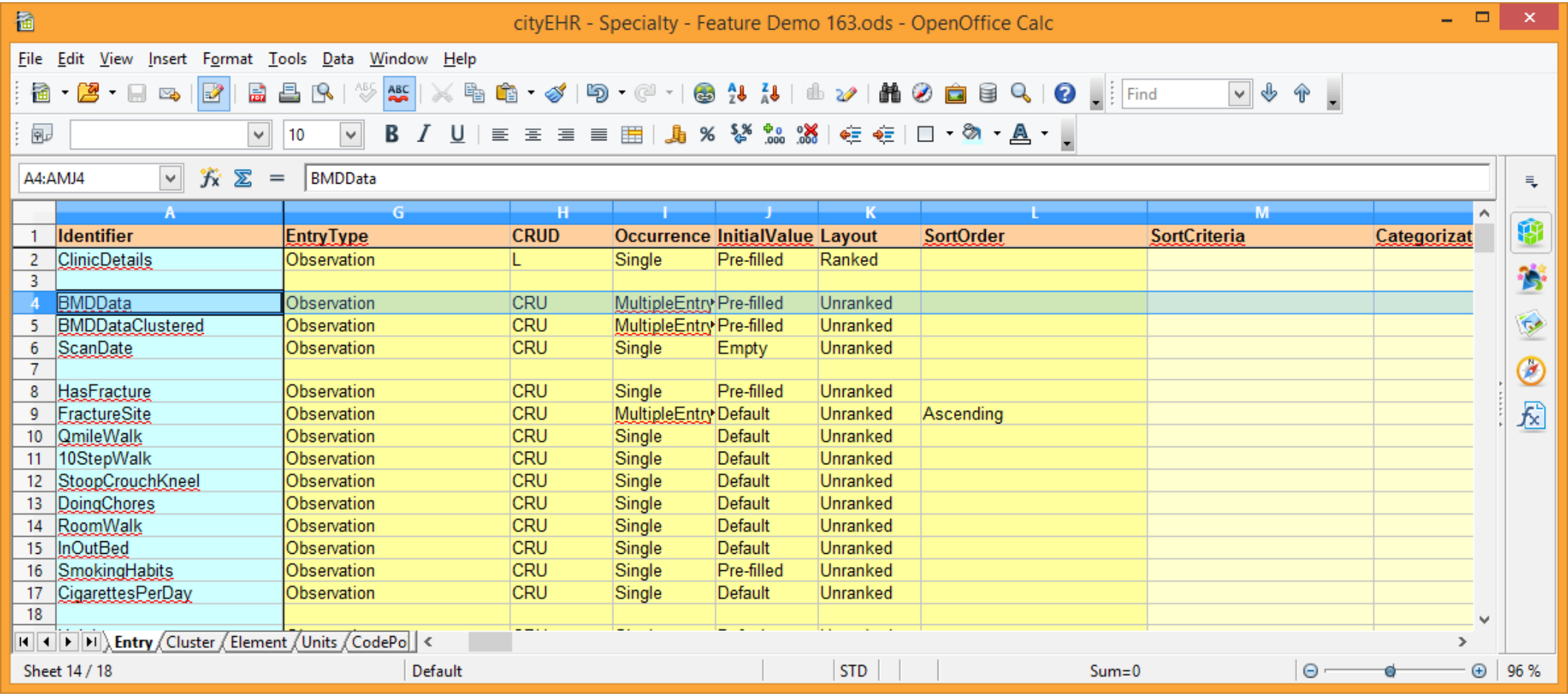
The OWL/XML format is one of several formats included in the standard

W3C Owl Working Group. 2012. OWL 2 Web Ontology Language. Document Overview (Second Edition) W3C Recommendation 11 December 2012.

Information Model – Spreadsheet Input

Clinicians can create their own information models using a spreadsheet

Converted to an OWL/XML ontology when imported to cityEHR



	A	G	H	I	J	K	L	M	
1	Identifier	EntryType	CRUD	Occurrence	InitialValue	Layout	SortOrder	SortCriteria	Categorizat
2	ClinicDetails	Observation	L	Single	Pre-filled	Ranked			
3									
4	BMDData	Observation	CRU	MultipleEntry	Pre-filled	Unranked			
5	BMDDataClustered	Observation	CRU	MultipleEntry	Pre-filled	Unranked			
6	ScanDate	Observation	CRU	Single	Empty	Unranked			
7									
8	HasFracture	Observation	CRU	Single	Pre-filled	Unranked			
9	FractureSite	Observation	CRU	MultipleEntry	Default	Unranked	Ascending		
10	QmileWalk	Observation	CRU	Single	Default	Unranked			
11	10StepWalk	Observation	CRU	Single	Default	Unranked			
12	StoopCrouchKneel	Observation	CRU	Single	Default	Unranked			
13	DoingChores	Observation	CRU	Single	Default	Unranked			
14	RoomWalk	Observation	CRU	Single	Default	Unranked			
15	InOutBed	Observation	CRU	Single	Default	Unranked			
16	SmokingHabits	Observation	CRU	Single	Pre-filled	Unranked			
17	CigarettesPerDay	Observation	CRU	Single	Default	Unranked			
18									

Catching Up With the Mayo Clinic

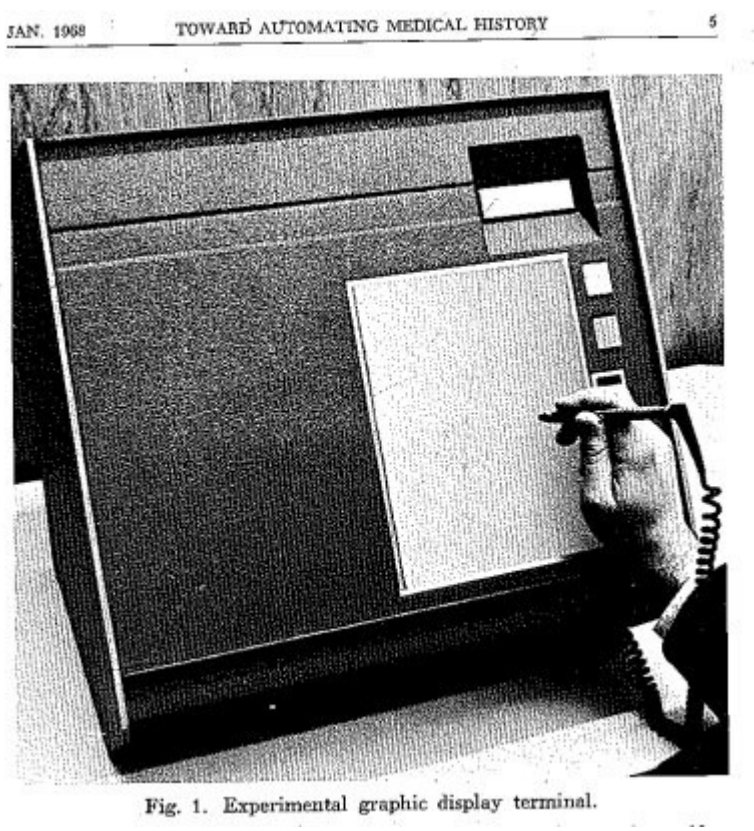


Fig. 1. Experimental graphic display terminal.

Mayne, J.G., Weksel, W. and Sholtz, P.N., 1968, January. Toward automating the medical history. In Mayo Clinic Proceedings (Vol. 43, No. 1, p. 1).

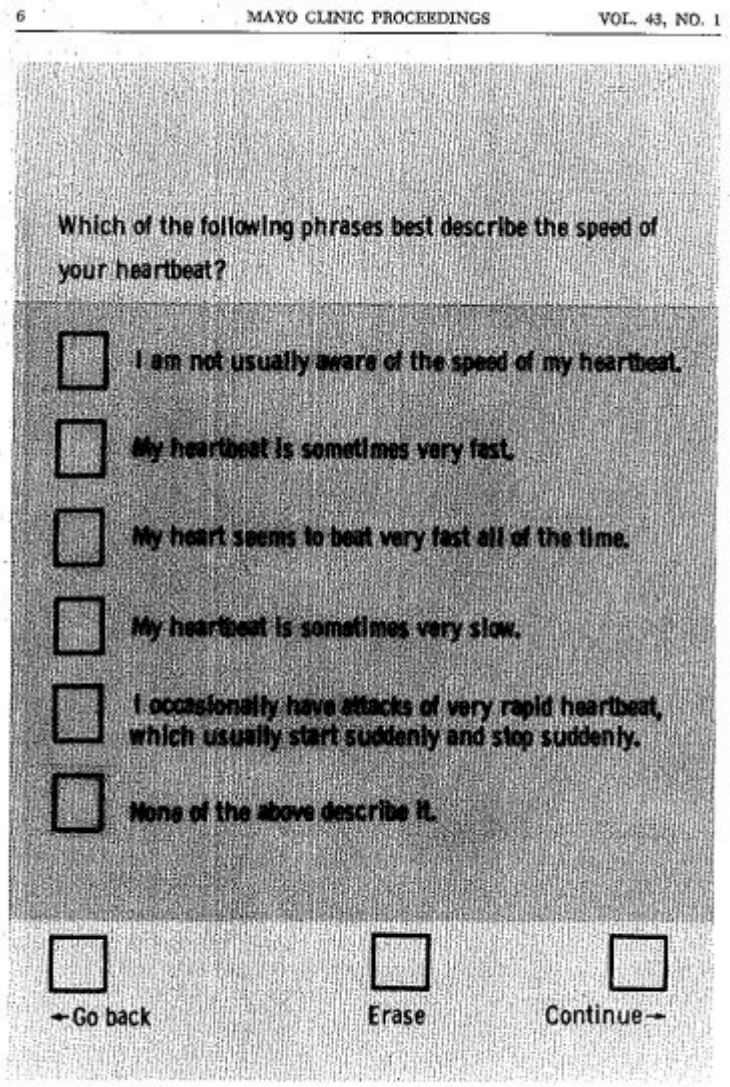
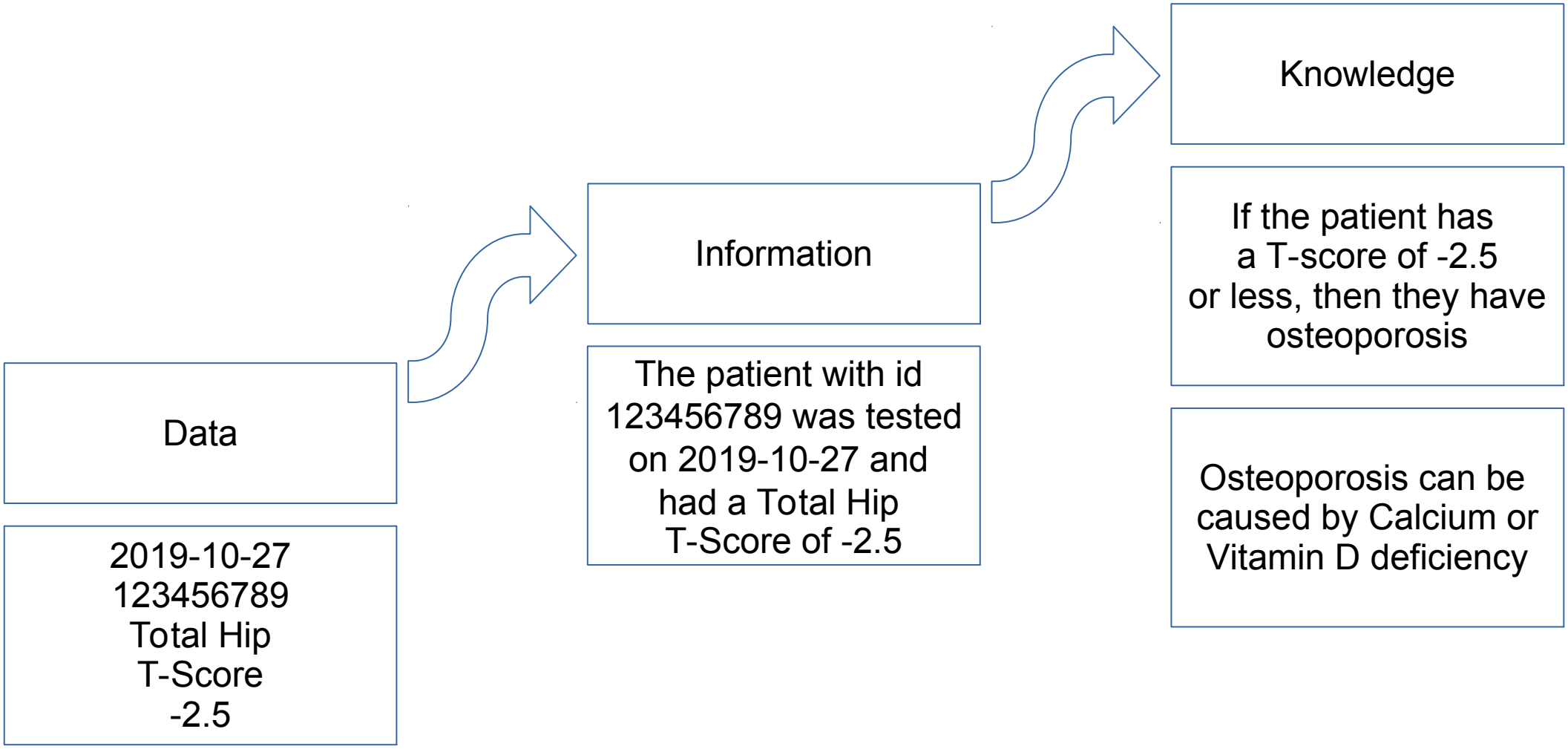


Fig. 2. Sample image, showing the three functional areas of each frame: question *(top)*; response alternatives *(middle)*; and instructions *(bottom)*. In the actual projected image, each functional area is a different color.

Knowledge and Reasoning

Data – Information - Knowledge

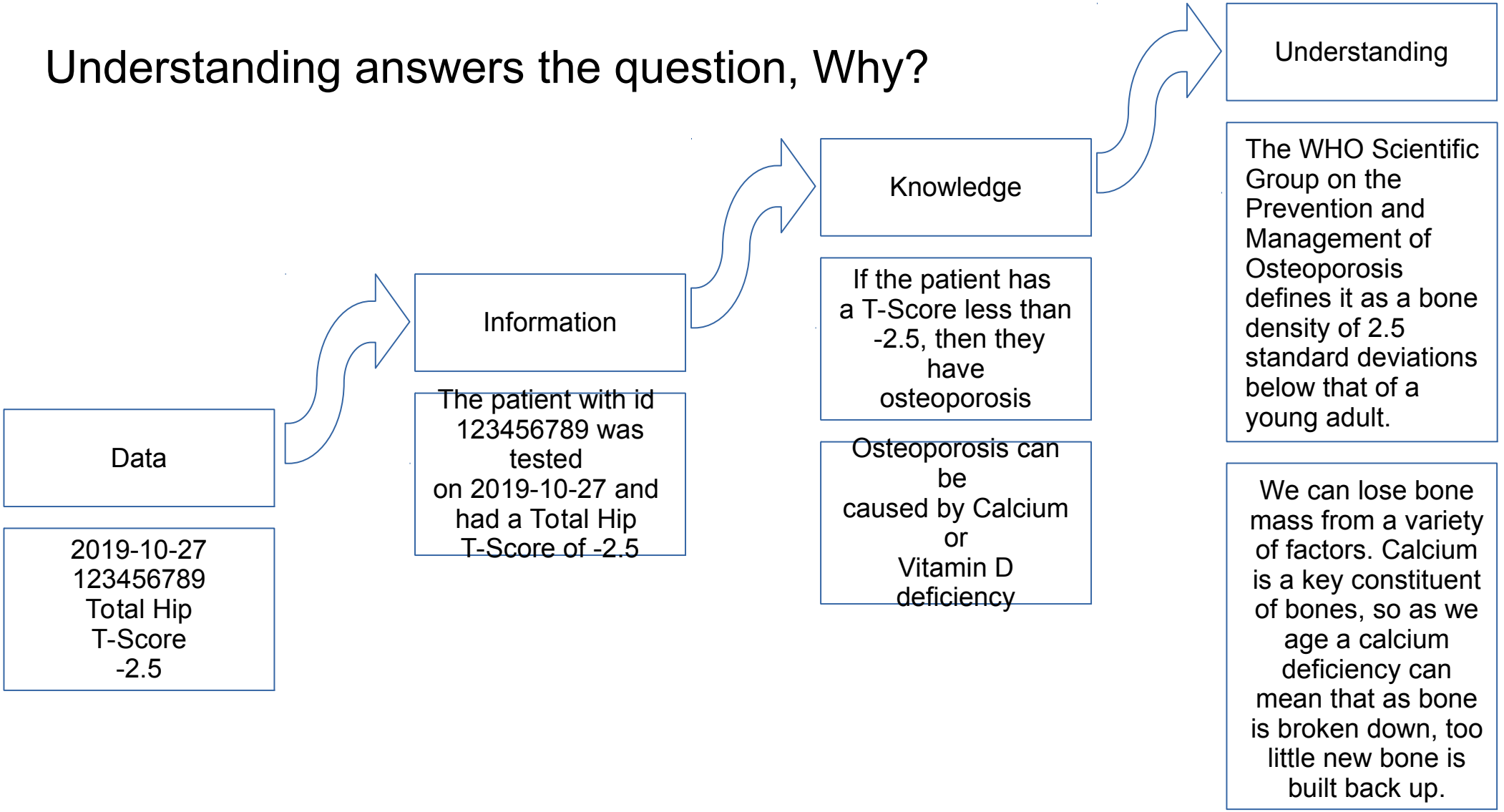
The progression had been well described by Russell Ackoff (1989).



Understanding

Ackoff extends beyond Knowledge to Understanding and Wisdom

Understanding answers the question, Why?



What is Ontology?

The concept of Ontology has been hijacked by Computer Scientists from the discipline of Philosophy, where it can be defined as

> The branch of metaphysics dealing with the nature of being.
>
> Oxford English Dictionary

Social Scientist, Anthony Giddens introduced the concept of Ontological Security

> We begin from the premise that to be a human being is to know, virtually all of the time, in terms of some description or another, both what one is doing and why one is doing it […] All individuals develop a framework of ontological security of some sort, based on routines of various forms.
>
> Giddens (1991) Modernity and Self-Identity: Self and Society in the Late Modern Age

> An ontologically insecure person does not accept at a fundamental level the reality or existence of things, themselves, and others.
>
> Jackson II, R.L. and Hogg, M.A. eds., 2010. Encyclopedia of Identity

Or in Computer Science, we might say that

> Ontological insecurity refers to a person who uses the word 'ontology' without really knowing what it means. Their insecurity arises from the fear that one day someone might ask them to define 'ontology' and they will be found lacking.

Ontology

A body of formally represented knowledge is based on a conceptualization: the objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them.

A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose. Every knowledge base, knowledge-based system, or knowledge-level agent is committed to some conceptualization, explicitly or implicitly.

An ontology is an explicit specification of a conceptualization.

T. R. Gruber. A translation approach to portable ontologies. Knowledge Acquisition, 5(2):199-220, 1993

Tom Gruber, Stanford University, Knowledge Systems Lab

Explicit Specification of a Conceptualization

Yellow walls
Grey carpet
Can't understand lecture

White walls
No windows
Boring lecture

Cream walls
Windows with blinds
Interesting lecture

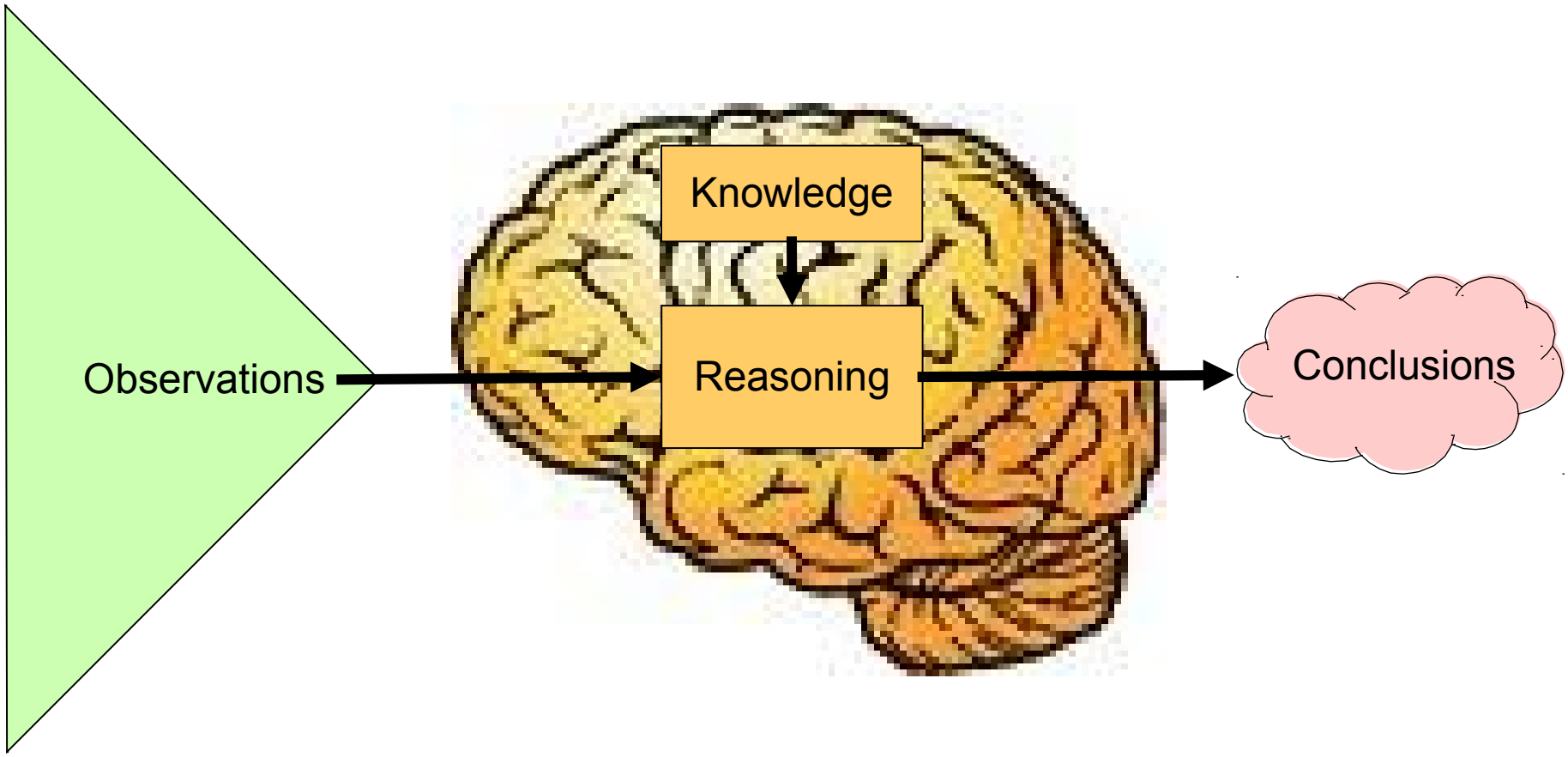
Knowledge

Knowledge is familiarity, awareness, or understanding gained through experience or study.

https://en.wikipedia.org/wiki/Knowledge

- Western philosophers for centuries have distinguished between two kinds of knowledge
- A priori knowledge is gained or justified by reason alone, without the direct or indirect influence of any particular experience (e.g. the truths of logic and mathematics)
- A posteriori knowledge is attained or justified by reference to experience (i.e. observation or measurement)
  This is also called empirical knowledge.

Knowledge Fuels Reasoning



- Observations provide information on the current (problem) context
- Knowledge base and observations are inputs to a reasoning engine
- Reasoning generates conclusions (solutions) relevant to the current context, based on observations and knowledge

Reasoning

Reasoning**.** The act of using reason, to derive a conclusion from certain premises, using a given methodology.

Reason is the ability to form and operate on concepts in abstraction, using rationality and logic

https://en.wikipedia.org/wiki/Reason

- logic deals principally with provable facts and demonstrably valid relations between them

Mr Spock

- rationality is a much broader term than logic, as it includes "uncertain but sensible" arguments based on probability, expectation and personal experience

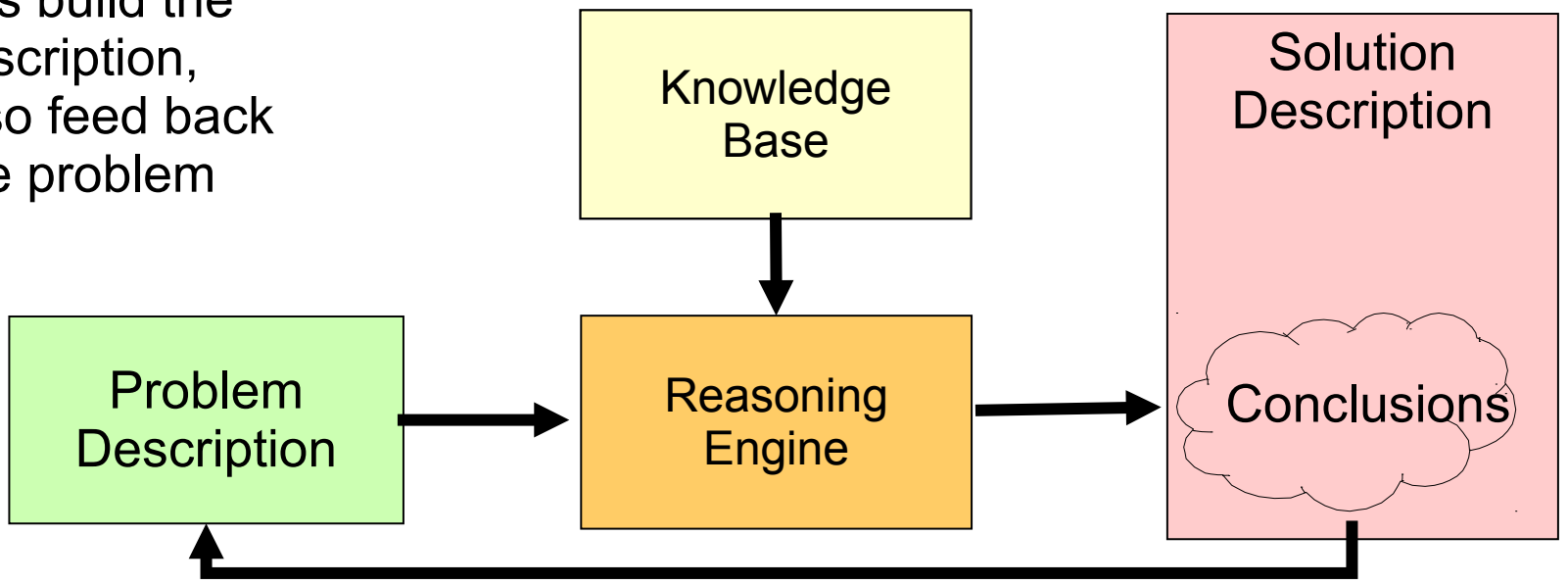
Captain Kirk

Reasoning Engine

A reasoning engine is a computer program that tries to derive answers from a knowledge base.

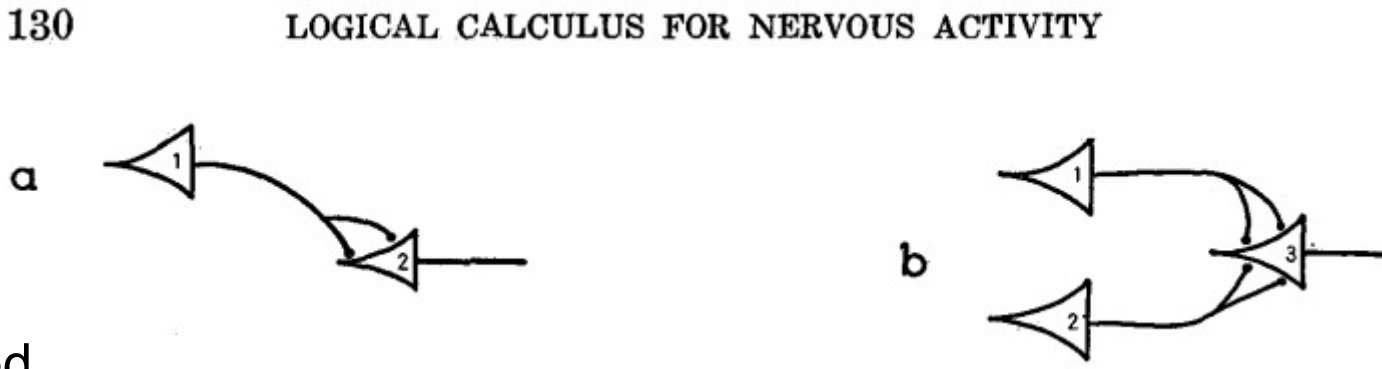
It is the “brain” that knowledge based systems use to reason about the information in the knowledge base, for the ultimate purpose of formulating new conclusions.

Conclusions build the solution description, but may also feed back to refine the problem description



Models of Human Thought Processes

In 1943, McCullogh and Pitts proposed a mathematical model of nervous activity.



In 1958, Rosenblatt introduced the theory of the perceptron.

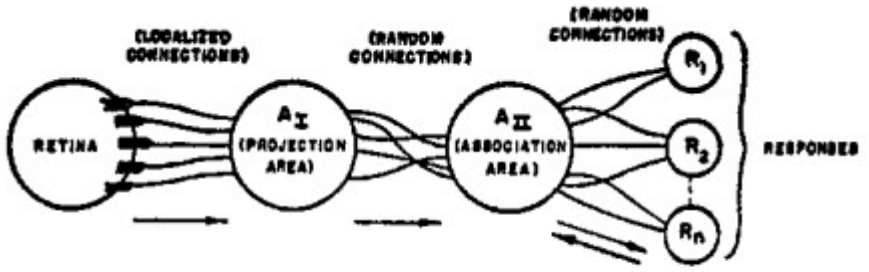


FIG. 1. Organization of a perceptron.

The theory has been developed for a hypothetical nervous system, or machine, called a perceptron. The perceptron is designed to illustrate some of the fundamental properties of intelligent systems in general, without becoming too deeply enmeshed in the special, and frequently unknown, conditions which hold for particular biological organisms. The analogy between the perceptron and biological systems should be readily apparent to the reader.

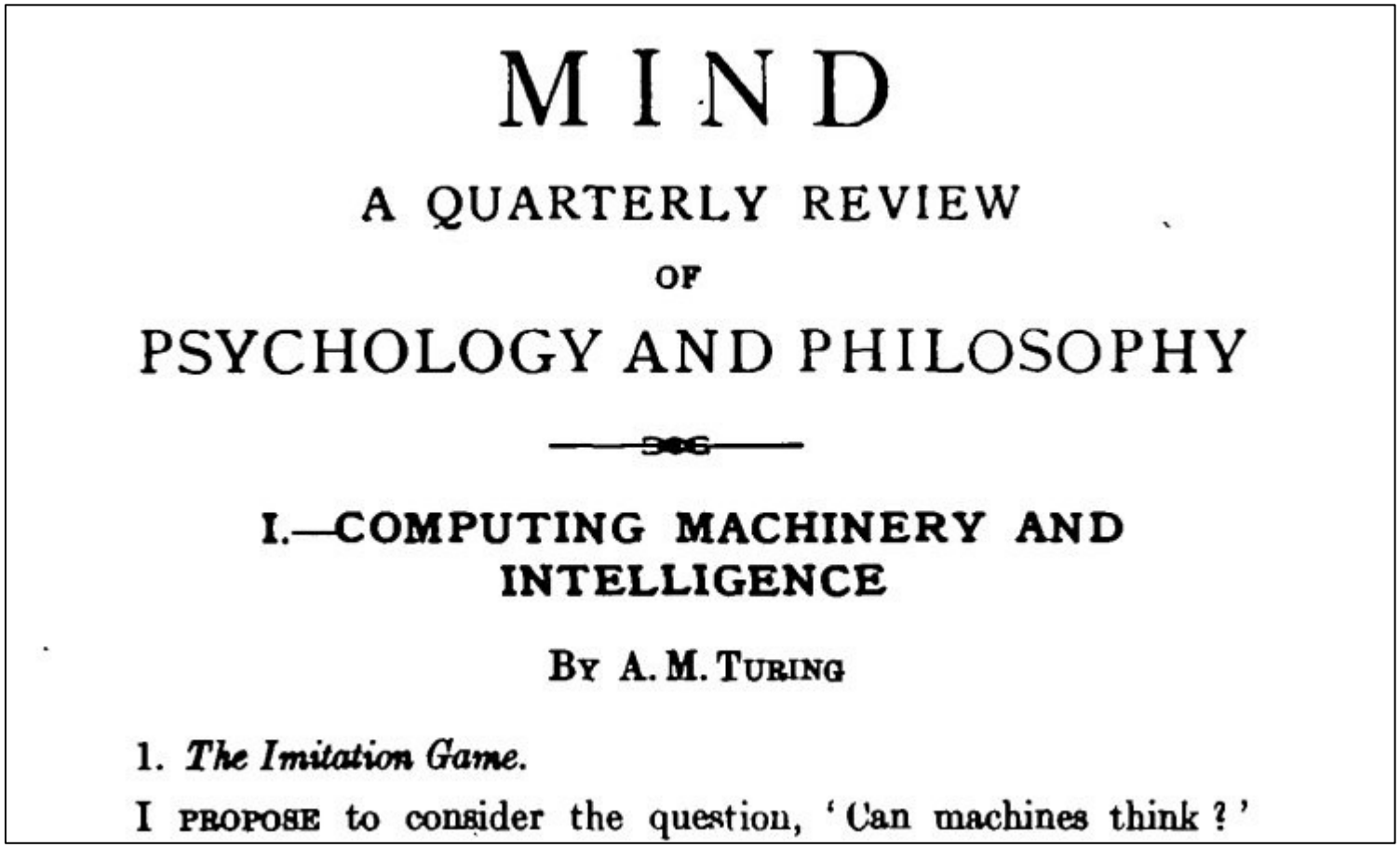
These two papers laid the foundations for the simulation of human thought processes using computers, leading to the field of Artificial Neural Networks (ANNs).

Artificial Intelligence

Artificial Intelligence – Beginnings

The academic discipline of Artificial Intelligence can be said to have begun in 1950 when Alan Turing posed the question "Can Machines Think?"

Turing answered the question by introducing "The Imitation Game" – which has since become known as "The Turing Test".



MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?'

The Imitation Game

The essence of the Imitation Game is that a human and a computer are interrogated by a second human, who must decide which of the entities under interrogation is the computer and which is the human, based solely on their answers to the questions posed; if the interrogator is unable to say, definitively, which is the human and which is the computer, then the computer is deemed to be a 'thinking machine'.

Some important questions follow:

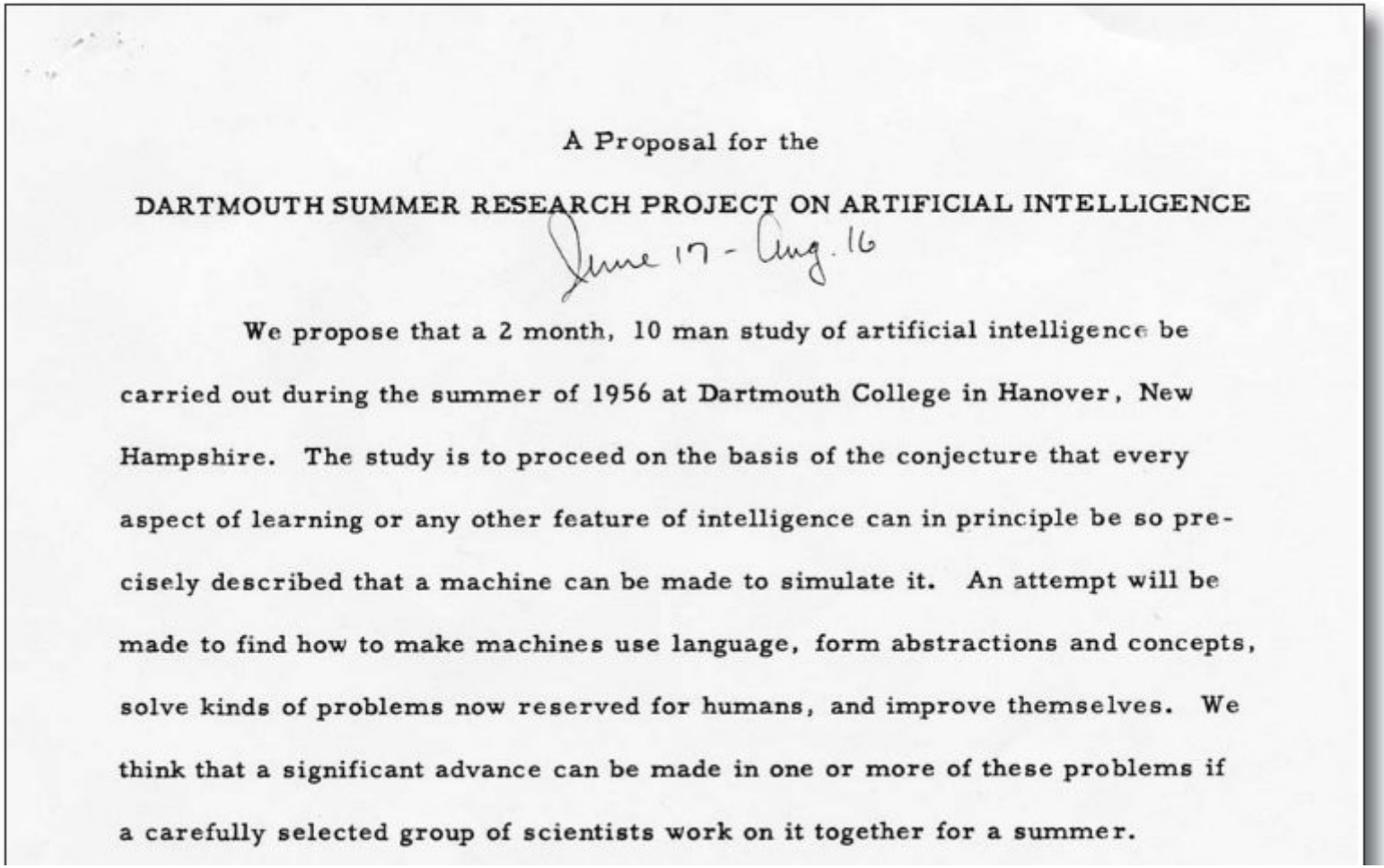
Does the 'thinking machine' display artificial intelligence or artificial stupidity?

If the machine gets the same questions 'wrong' as the human, is this useful to us?

Should we try to make machines which think like humans (i.e. use similar processes) or just appear to think like humans?

Artificial Intelligence – Dartmouth Summer Project

The term "Artificial Intelligence" appeared in John McCarthy's proposal for a summer research project at Dartmouth College in 1956, which brought together many of the early pioneers and really kicked off the study of AI.



A Proposal for the

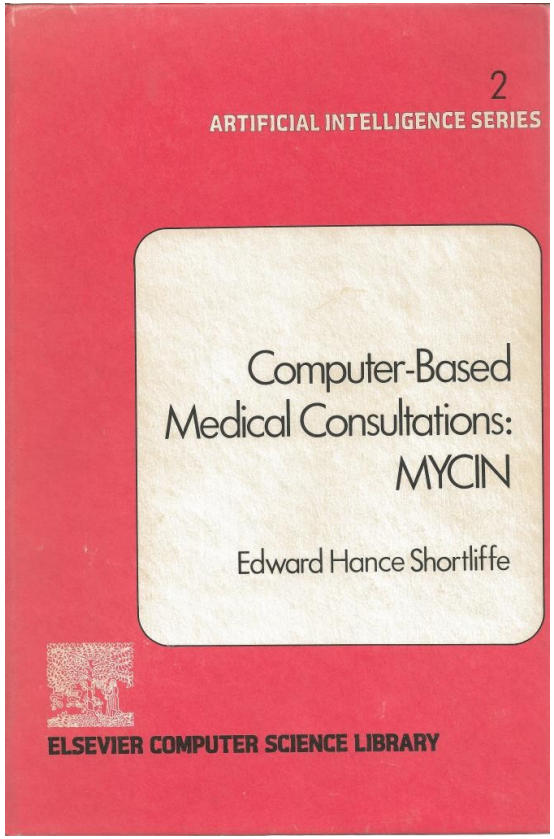
DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

June 17 - Aug. 16

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

Artificial Intelligence in Medicine

The MYCIN rule-based expert system, for the identification and treatment of bacterial infections, became the most widely known AI system in the medical arena.



RULE037

IF: 1) THE IDENTITY OF THE ORGANISM IS NOT KNOWN WITH CERTAINTY, AND
2) THE STAIN OF THE ORGANISM IS GRAMNEG, AND
3) THE MORPHOLOGY OF THE ORGANISM IS ROD, AND
4) THE AEROBICITY OF THE ORGANISM IS AEROBIC
THEN: THERE IS STRONGLY SUGGESTIVE EVIDENCE (.8) THAT THE CLASS OF THE ORGANISM IS ENTEROBACTERIACEAE

RULE145

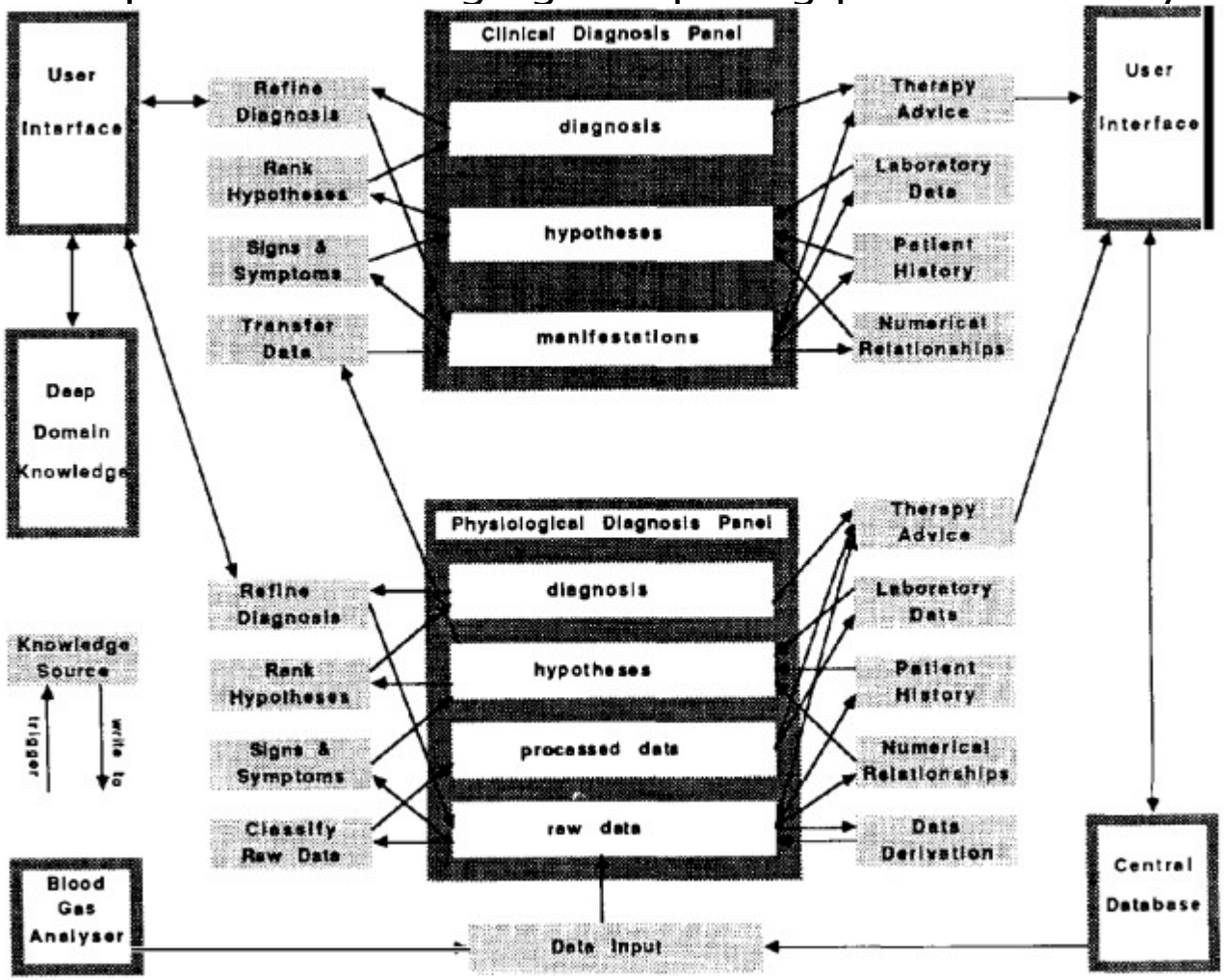
IF: 1) THE THERAPY UNDER CONSIDERATION IS ONE OF: CEPHALOTHIN CLINDAMYCIN ERYTHROMYCIN LINCOMYCIN VANCOMYCIN, AND
2) MENINGITIS IS AN INFECTIOUS DISEASE DIAGNOSIS FOR THE PATIENT
THEN: IT IS DEFINITE (1) THE THE THERAPY UNDER CONSIDERATION IS NOT A POTENTIAL THERAPY FOR USE AGAINST THE ORGANISM

Shortliffe, E.H., 1974. MYCIN: a rule-based computer program for advising physicians regarding antimicrobial therapy selection. STANFORD UNIV CALIF DEPT OF COMPUTER SCIENCE.

Available at: http://people.dbmi.columbia.edu/~ehs7001/Shortliffe-1976/MYCIN%20thesis%20Book.htm

Knowledge-Based Expert Systems

The second wave, in the mid-1980s, arrived just when the (IBM) Personal Computer was bringing computing power to every desktop.



Blackboard paradigm for developing hypotheses.

Physiological diagnosis based on data from critical care unit

Clinical diagnosis of acid-base Disorders

Used statistical models of uncertainty, including a Bayesian Belief Network

Chelsom, J.J. et al 1988. Blood gas analysis: a knowledge-based adviser for the interpretation of results.

Pearl, J., 1986. Fusion, propagation, and structuring in belief networks.

Artificial Neural Networks

The third wave, from 2010, when cloud computing, big data and personal assistants caught up with various theories that started in the earlier waves.

Artificial Neural Networks (ANN) are a network of nodes where the signal pattern presented at the inputs is propagated through the connected (and weighted) nodes to produce an output.

The network is 'trained' by presenting a series of known test patterns, and the weights (and node architecture) are adjusted until the desired outputs are obtained.

One key training algorithm is that of 'back propagation', first described by Paul Werbos in 1974, which started to gain attention in the 1980s.

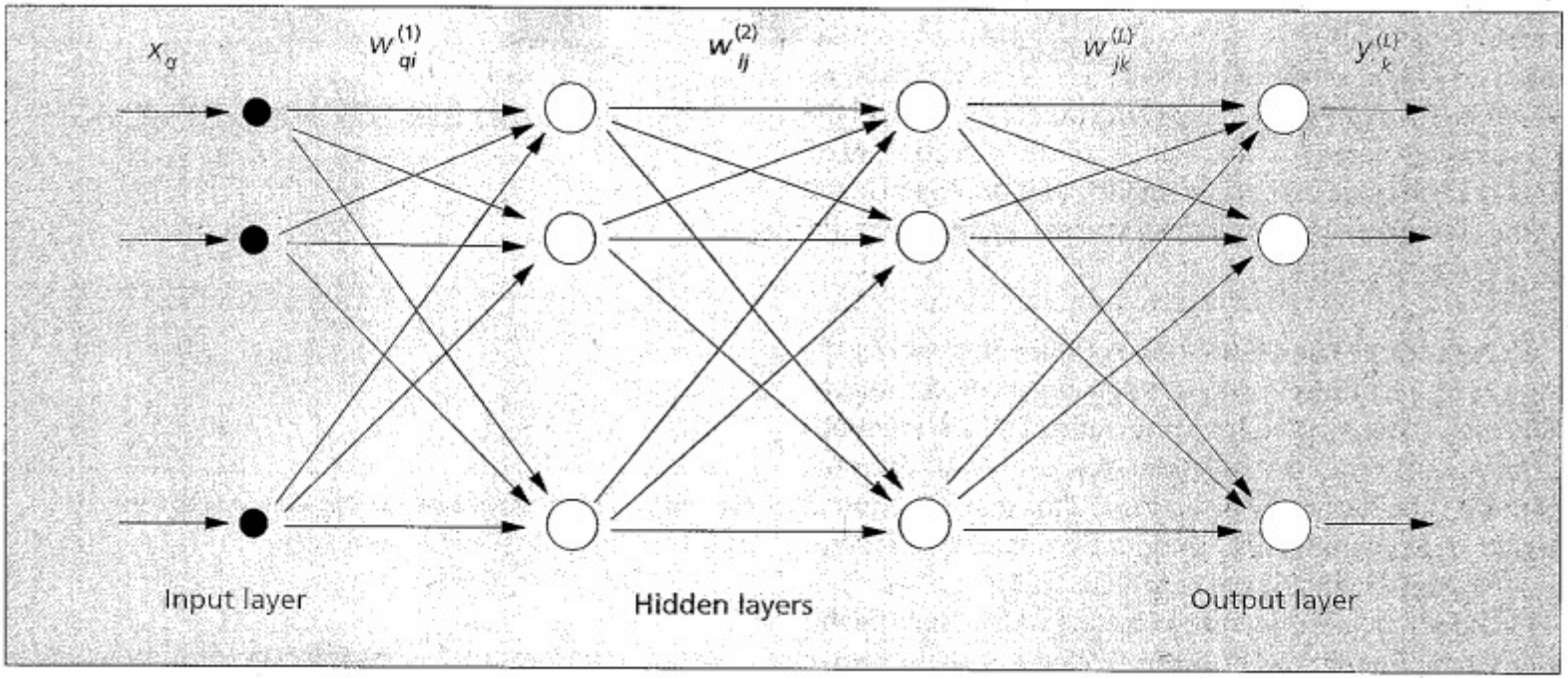


Figure 7. A typical three-layer feed-forward network architecture.

Jain, A.K., Mao, J. and Mohiuddin, K.M., 1996. Artificial neural networks: A tutorial. Computer, 29(3), pp.31-44.

Artificial Intelligence Techniques Using Structured Health Records

Linking Health Records with Knowledge Sources

Rules-based Reasoning

Bayesian Inference

Clinical Coding

SNOMED CT consists of over a million medical Concepts.

For example 64859006 means Osteoporosis.

http://www.snomedbrowser.com/

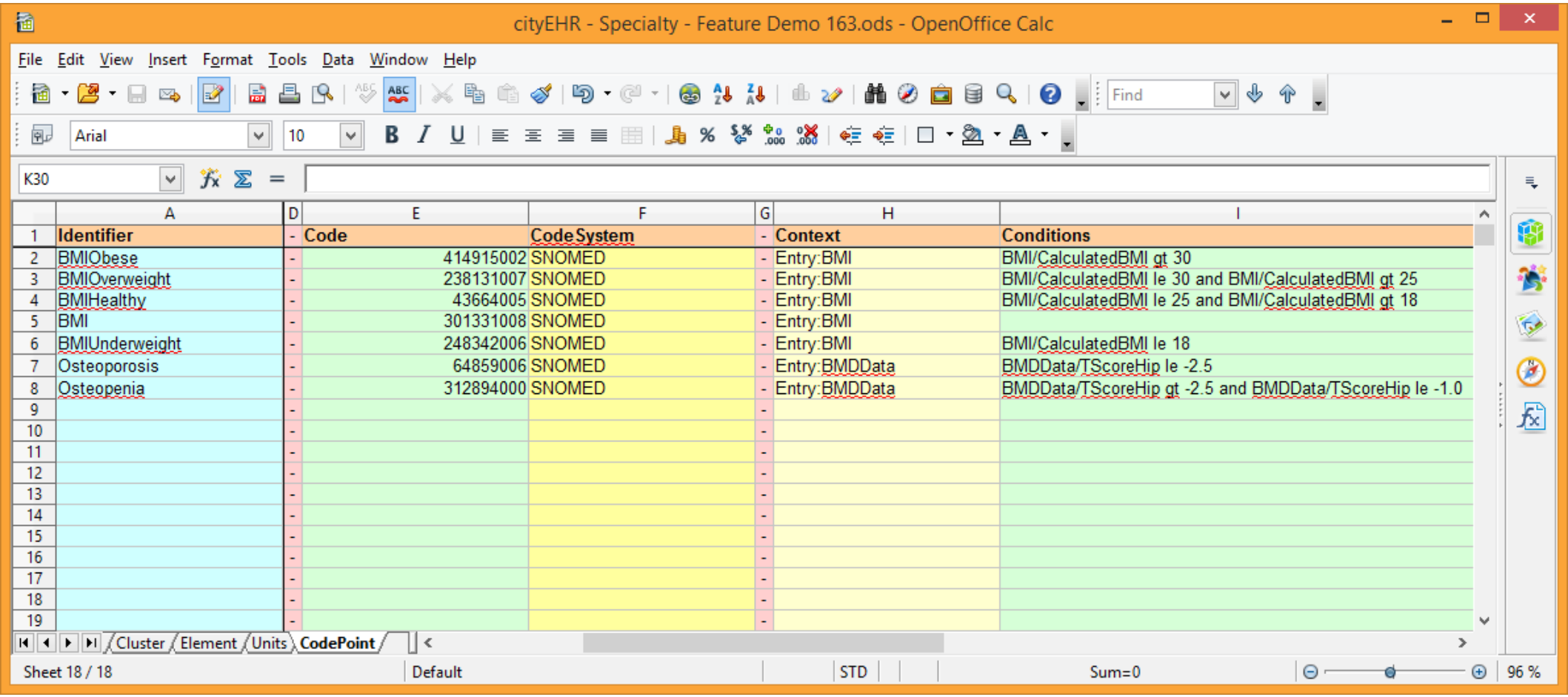
Ontology Model of Clinical Coding

```
<Declaration>
   <Class IRI="#CodingSystem:SNOMED"/>
</Declaration>
<Declaration>
   <NamedIndividual IRI="#Code:SNOMED:64859006"/>
</Declaration>
<ClassAssertion>
   <Class IRI="#CodingSystem:SNOMED"/>
   <NamedIndividual IRI="#Code:SNOMED:64859006"/>
</ClassAssertion>
<DataPropertyAssertion>
   <DataProperty IRI="#hasDisplayName"/>
   <NamedIndividual IRI="#Code:SNOMED:64859006"/>
   <Literal xml:lang="en-gb">Osteoporosis</Literal>
</DataPropertyAssertion>
<DataPropertyAssertion>
   <DataProperty IRI="#hasCode"/>
   <NamedIndividual IRI="#Code:SNOMED:64859006"/>
   <Literal>64859006</Literal>
</DataPropertyAssertion>
```

```
<Declaration>
   <Class IRI="#CodePoint"/>
</Declaration>
<Declaration>
   <NamedIndividual IRI="#CodePoint:BDMData:64859006"/>
</Declaration>
<ClassAssertion>
   <Class IRI="#CodePoint"/>
   <NamedIndividual IRI="#CodePoint:BDMData:64859006"/>
</ClassAssertion>
<DataPropertyAssertion>
   <DataProperty IRI="#hasCondition"/>
   <NamedIndividual IRI="#CodePoint:BDMData:64859006"/>
   <Literal>BMDData/TScore lt -2.5</Literal>
</DataPropertyAssertion>
<ObjectPropertyAssertion>
   <ObjectProperty IRI="#hasCode"/>
   <NamedIndividual IRI="#CodePoint:BDMData:64859006"/>
   <NamedIndividual IRI="#Code:SNOMED:64859006"/>
</ObjectPropertyAssertion>

<ObjectPropertyAssertion>
   <ObjectProperty IRI="#hasCodePoint"/>
   <NamedIndividual IRI="#ISO-13606:Entry:BMDData""/>
   <NamedIndividual IRI="#CodePoint:BDMData:64859006"/>
</ObjectPropertyAssertion>
```

Code Points – Spreadsheet Entry

Code Points can be created in the Information Model (using the spreadsheet)



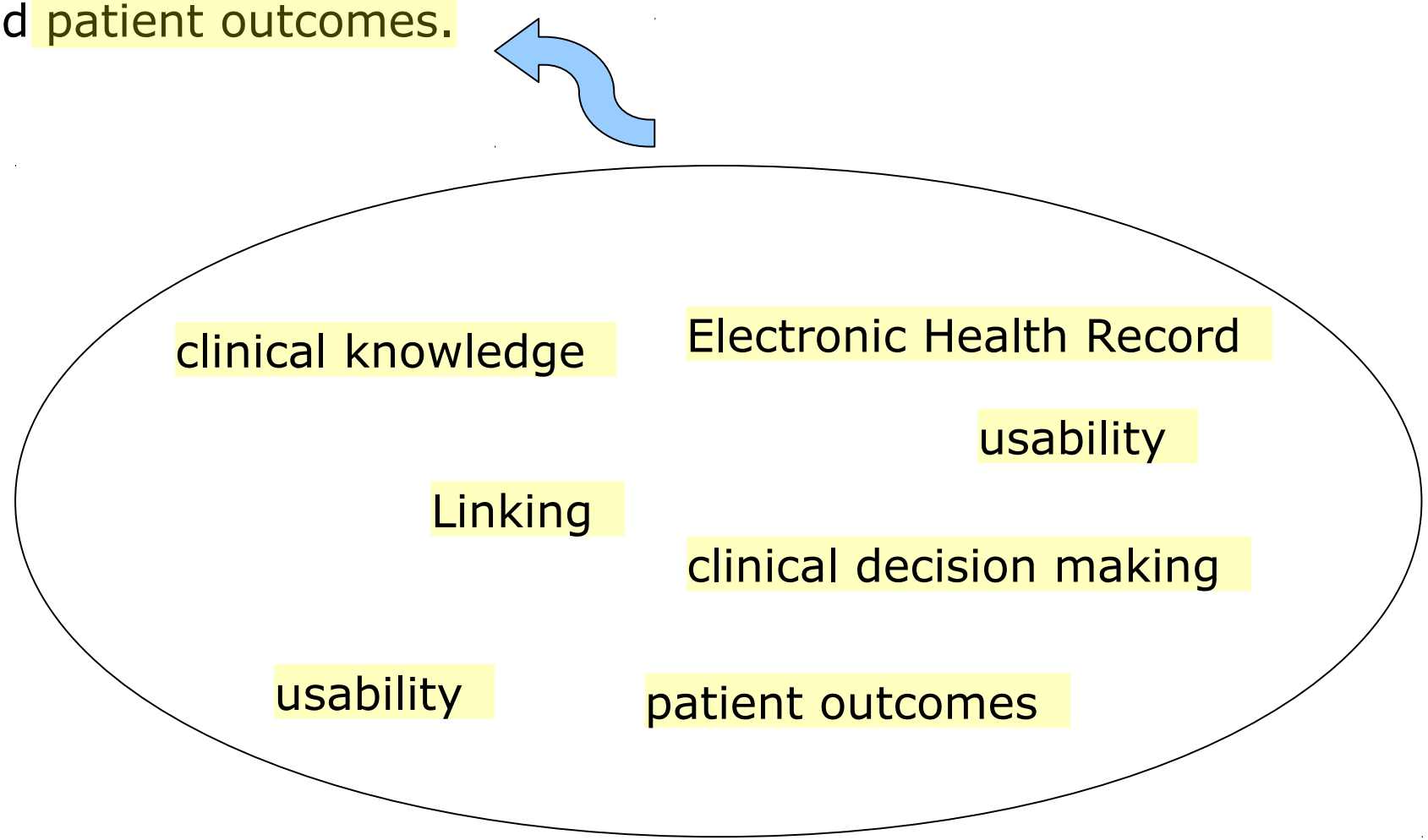
	A	D	E	F	G	H	I
1	Identifier	-	Code	CodeSystem	-	Context	Conditions
2	BMIObese	-	414915002	SNOMED	-	Entry:BMI	BMI/CalculatedBMI gt 30
3	BMIOverweight	-	238131007	SNOMED	-	Entry:BMI	BMI/CalculatedBMI le 30 and BMI/CalculatedBMI gt 25
4	BMIHealthy	-	43664005	SNOMED	-	Entry:BMI	BMI/CalculatedBMI le 25 and BMI/CalculatedBMI gt 18
5	BMI	-	301331008	SNOMED	-	Entry:BMI	
6	BMIUnderweight	-	248342006	SNOMED	-	Entry:BMI	BMI/CalculatedBMI le 18
7	Osteoporosis	-	64859006	SNOMED	-	Entry:BMDData	BMDData/TScoreHip le -2.5
8	Osteopenia	-	312894000	SNOMED	-	Entry:BMDData	BMDData/TScoreHip gt -2.5 and BMDData/TScoreHip le -1.0
9		-			-		
10		-			-		
11		-			-		
12		-			-		
13		-			-		
14		-			-		
15		-			-		
16		-			-		
17		-			-		
18		-			-		
19		-			-		

Linking Health Records with Knowledge Sources

Linking sources of clinical knowledge with the Electronic Health Record, is a relatively 'quick win' for improving the usability of the EHR and supporting clinical decision making, with associated benefits for clinical safety and patient outcomes.

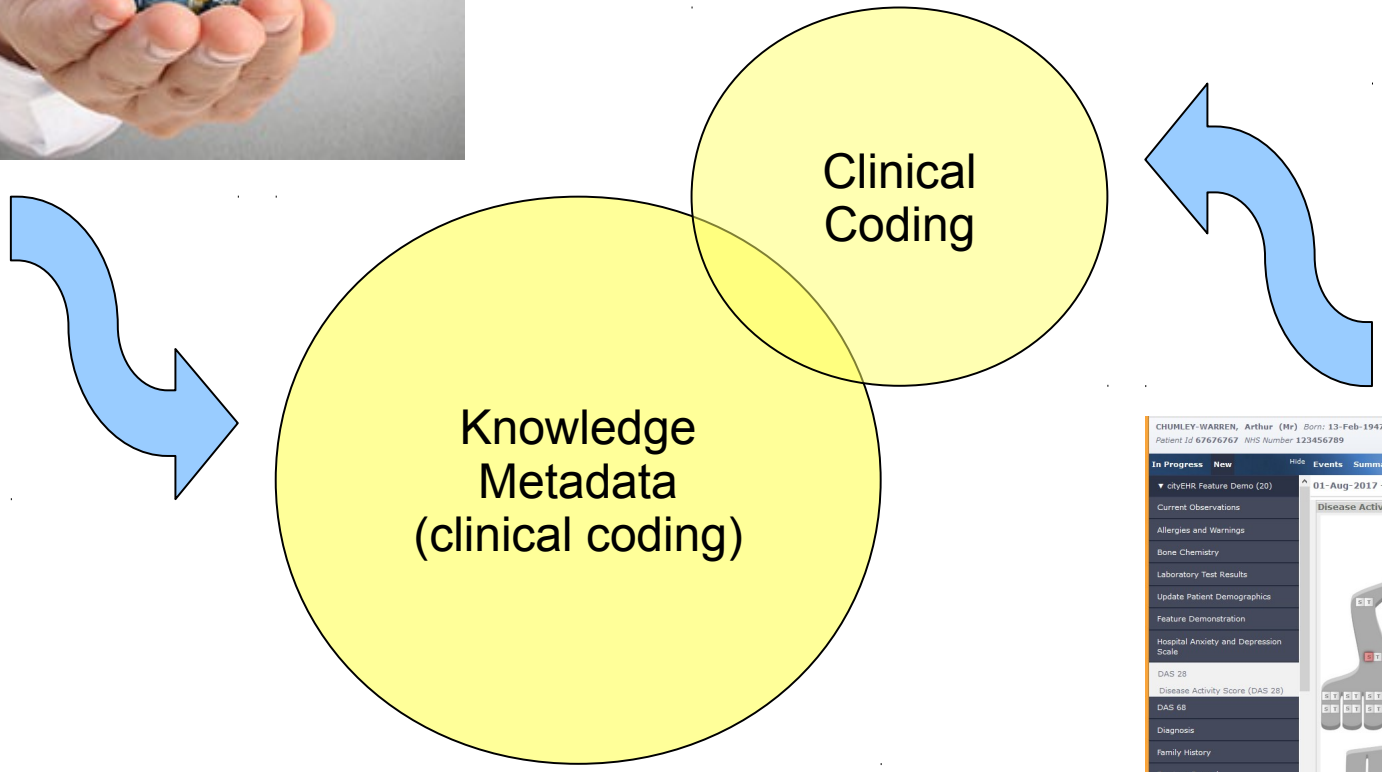
Linking Health Records with Knowledge Sources

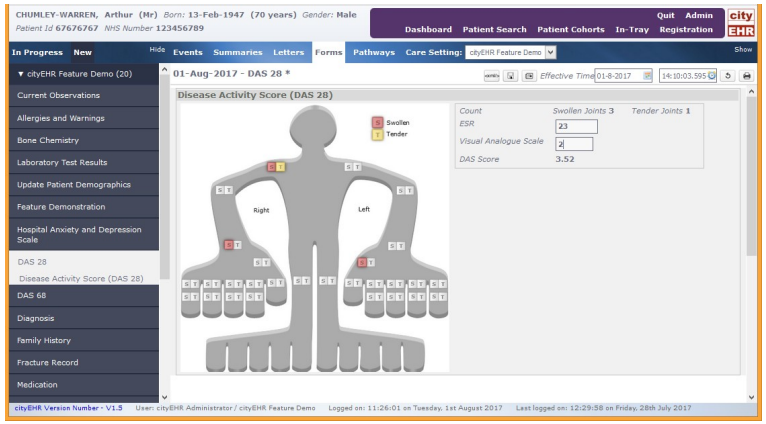
Linking sources of clinical knowledge with the Electronic Health Record, is a relatively 'quick win' for improving the usability of the EHR and supporting clinical decision making, with associated benefits for clinical safety and patient outcomes.



World Of Knowledge

Knowledge available on the Internet – the Semantic Web.





Clinical Information in a structured health record.

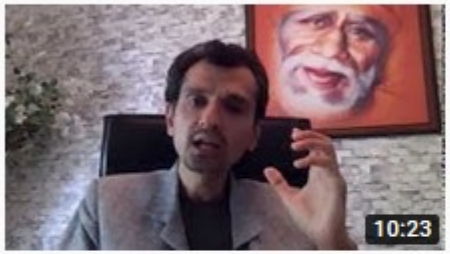
A World Of Knowledge

Here's some of the knowledge about osteoporosis, available on youtube

3011 - Osteoporosis / From Sickness to Health - Barbara O'Neill

Amazing Discoveries • 100K views • 1 year ago

In this lecture with Barbara O'Neill we will look at what causes the deterioration of our bones and learn what we can do to ...



HOW TO REVERSE OSTEOPOROSIS IN 6 MONTHS - Increase bone density by Amitabh Pandit

Amitabh Pandit • 198K views • 3 years ago

This has been a long awaited video I personally wanted to upload. First I wanted to see the results myself whether these foods did ...

5 Signs of Osteoporosis

Cure Diseases Naturally • 3.5K views • 7 months ago

Hey Everyone, Aqila here.. In this video I will talk about "5 Signs of **Osteoporosis**" Knowing the early 5 Common Signs of ...

Subtitles

Osteoporosis: Prevention and Treatment

Stanford Health Care • 10K views • 2 years ago

Over 50% of women and 25% of men over the age of 50 will sustain a fragility fracture due to **osteoporosis**. This talk addresses ...

Resource Description Framework (RDF)

Resource Description Framework (RDF)

Part of the Semantic Web activity from the World Wide Web Consortium

Suppose that I find a video on youtube about the disorder Osteoporosis

Meta data about the video (i.e. a resource on the web) can be represented as 'triples' of Subject – Predicate – Object

Video disorder is Osteoporosis

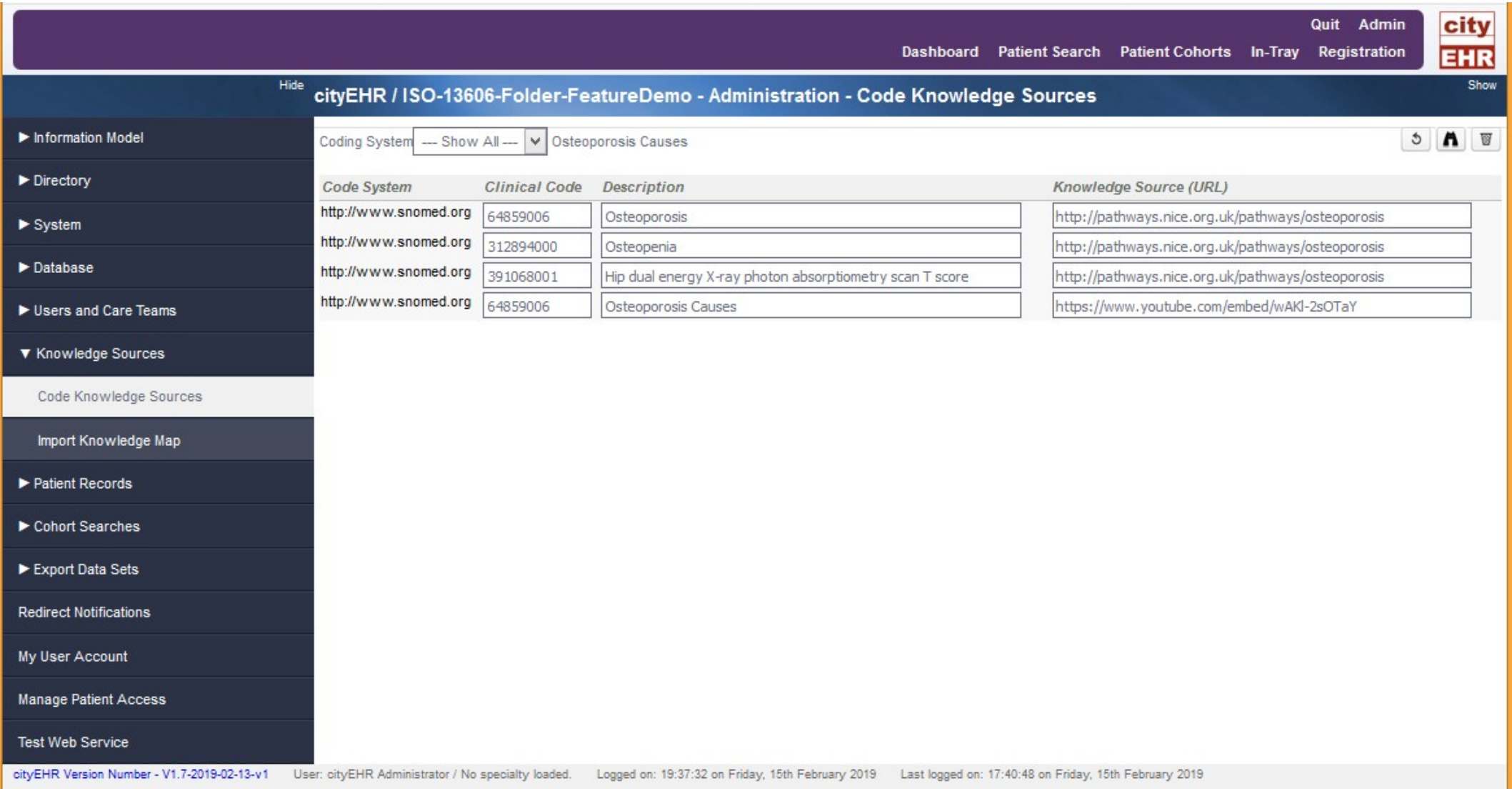
Disorder in the Video is Osteoporosis

```
<rdf:RDF  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:medicalTerms="http://medicalTerms.org/disorders"/>
  <rdf:Description rdf:about="https://www.youtube.com/embed/wAKl-2sOTaY">
        – <medicalTerms:Disorder>Osteoporosis</medicalTerms:Disorder>
  </rdf:Description>
</rdf:RDF>
```

RDF Knowledge Map

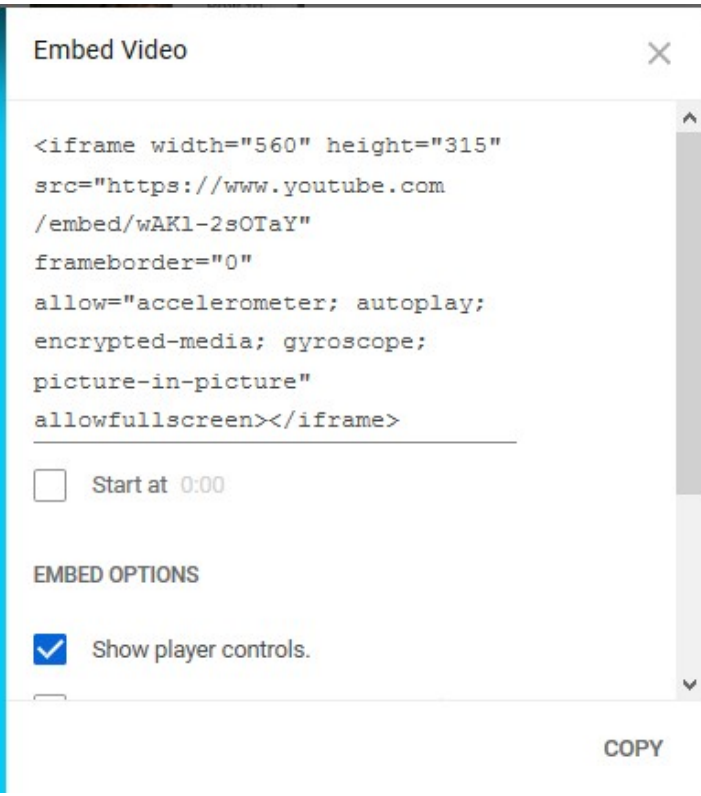
```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:dc="http://purl.org/dc/elements/1.1/" dc:title="Knowledge Map">

    <rdf:Description rdf:about="http://pathways.nice.org.uk/pathways/osteoporosis">
        <dc:subject rdf:datatype="http://www.snomed.org">391068001</dc:subject>
        <dc:description rdf:datatype="http://www.snomed.org">
           Hip dual energy X-ray photon absorptiometry scan T score
        </dc:description>
    </rdf:Description>

    <rdf:Description rdf:about="http://pathways.nice.org.uk/pathways/osteoporosis">
        <dc:subject rdf:datatype="http://www.snomed.org">64859006</dc:subject>
        <dc:description rdf:datatype="http://www.snomed.org">Osteoporosis</dc:description>
    </rdf:Description>

    <rdf:Description rdf:about="http://pathways.nice.org.uk/pathways/osteoporosis">
        <dc:subject rdf:datatype="http://www.snomed.org">312894000</dc:subject>
        <dc:description rdf:datatype="http://www.snomed.org">Osteopenia</dc:description>
    </rdf:Description>

</rdf:RDF>
```

Coding Knowledge Sources

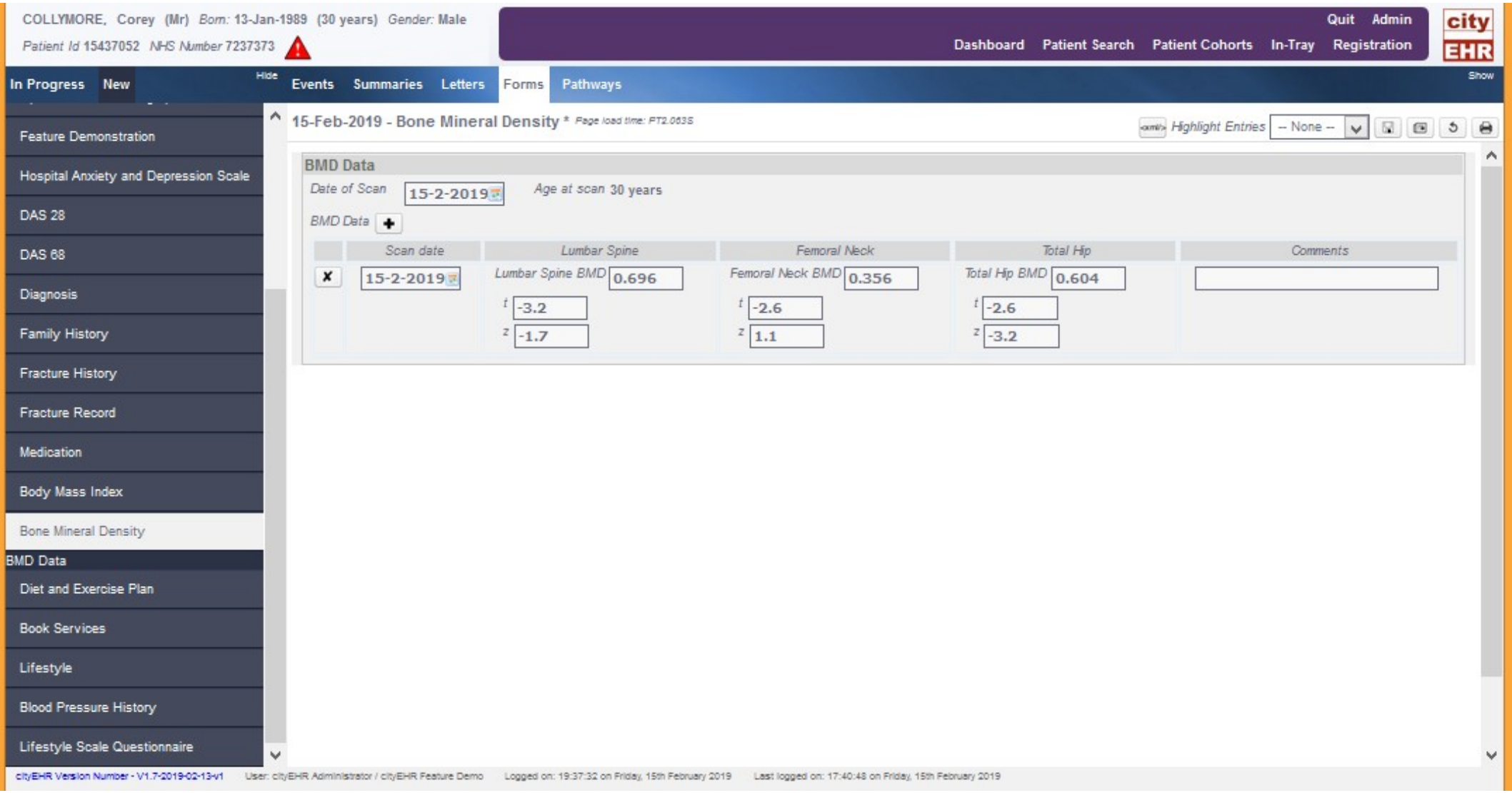


Addressable Knowledge can be Coded

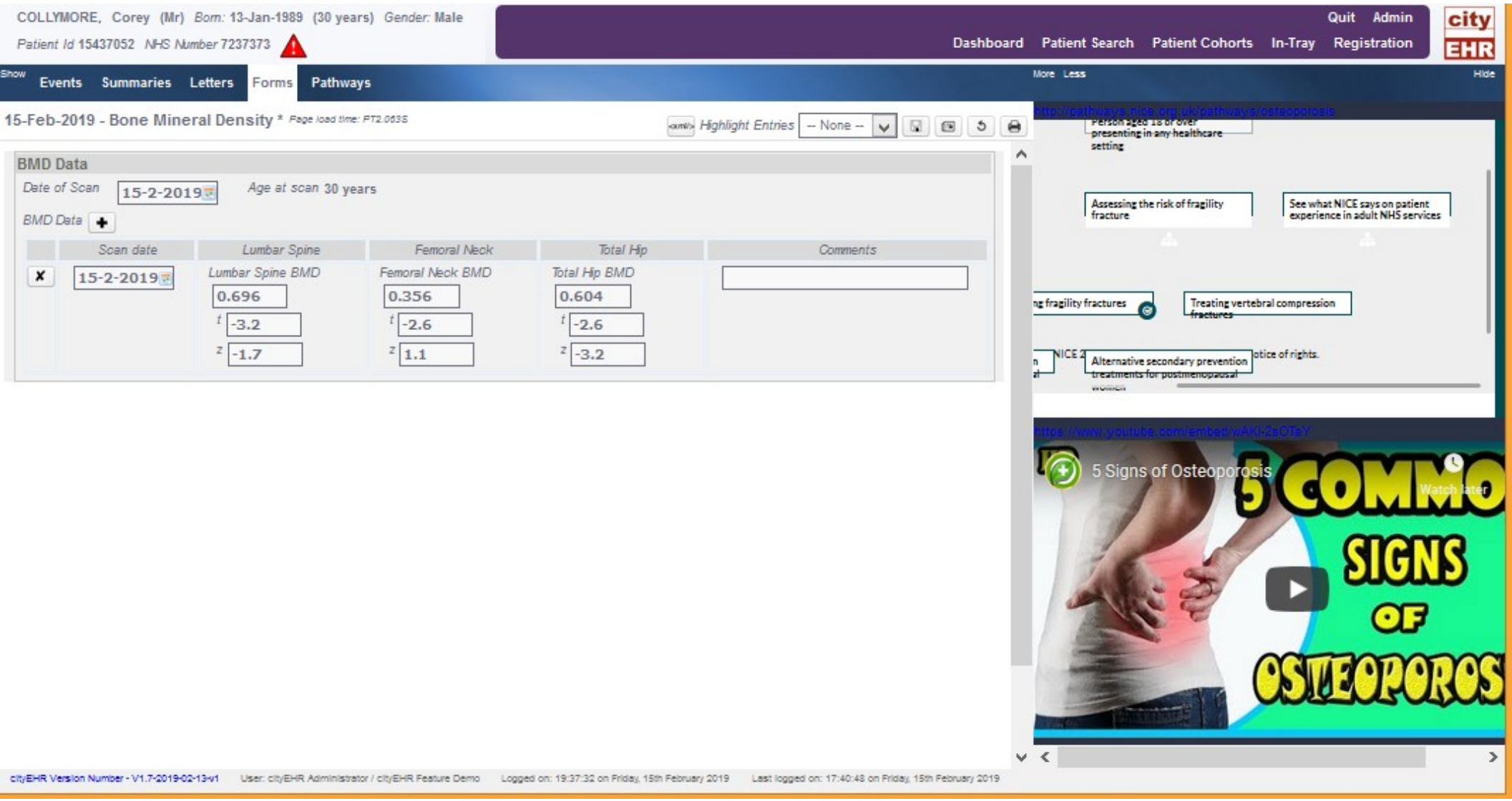
https://www.youtube.com/embed/wAKl-2sOTaY



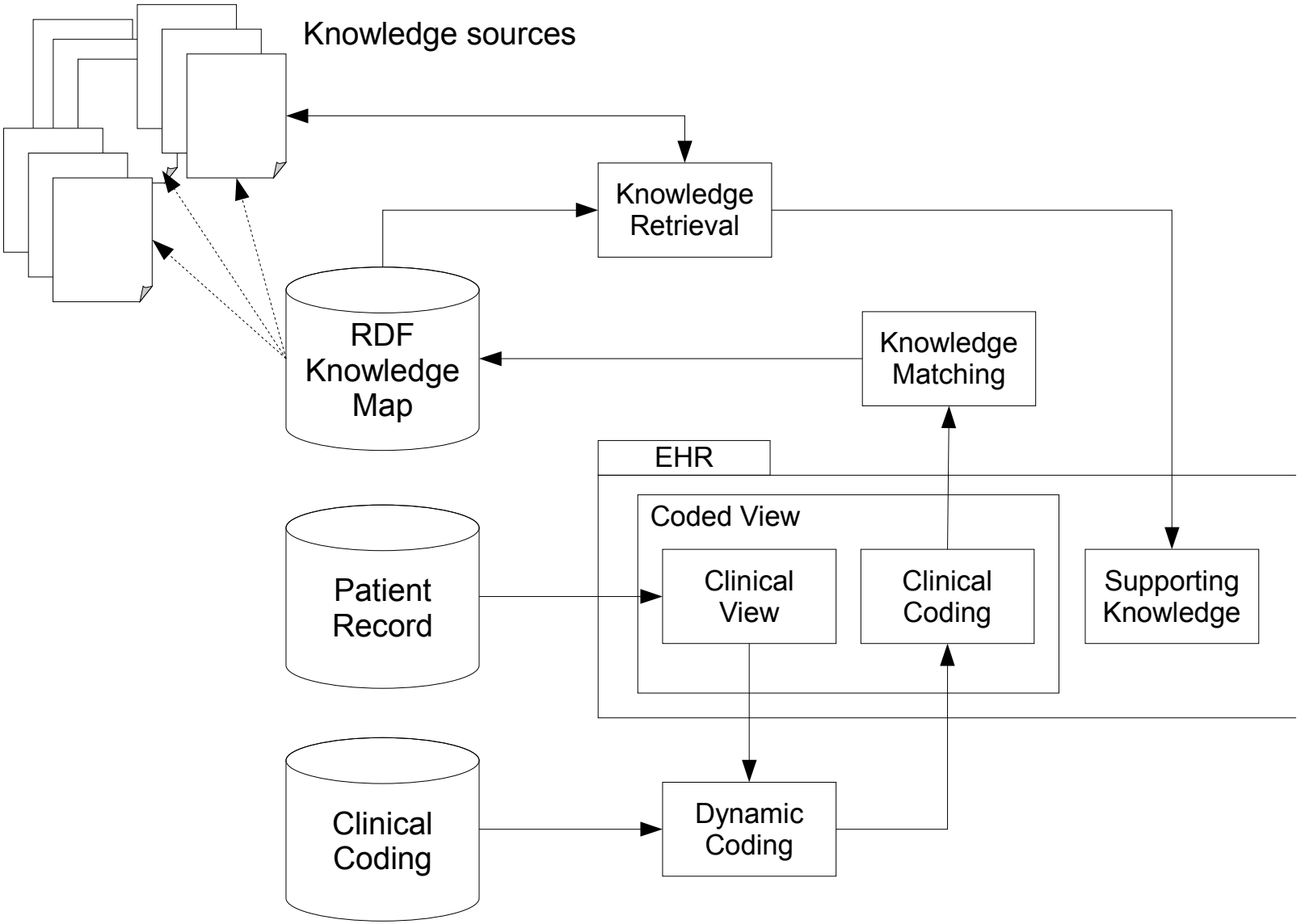
Record Bone Mineral Density



Knowledge Linked Using Clinical Codes



Linking Health Records with Knowledge



Artificial Intelligence Techniques Using Structured Health Records

Linking Health Records with Knowledge Sources

Rules-based Reasoning

Bayesian Inference

Deduction

If a patient has a T-Score (measure of their bone mineral density) less than -2.5 then they have Osteoporosis.

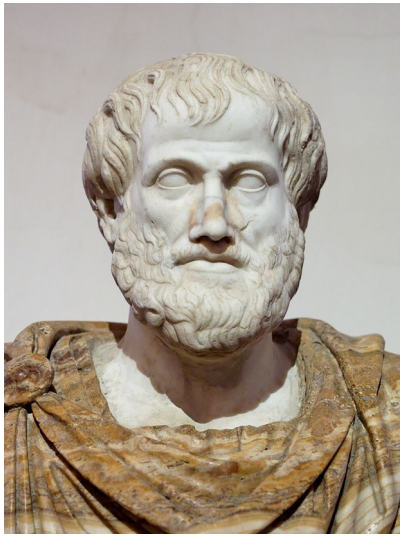
Deduction is generally defined as "the deriving of a conclusion by reasoning". In deduction the truth of the conclusion is guaranteed by the truth of the statements or facts considered.

Merriam-Webster Dictionary

(Actually, Watson doesn't really know, since deduction isn't usually one of Sherlock Holmes' methods.)

Syllogism

A syllogism is a deductive scheme of a formal argument consisting of a major and a minor premise and a conclusion *Merriam-Webster Dictionary*



As introduced by the Greek philosopher, Aristotle

The major premise is a general statement of truth
The minor premise is a specific statement
A common term links the major and minor premises
The common term is eliminated in the conclusion

Deduction as a syllogism

Major premise: Osteoporosis is defined by a T-Score less than -2.5

Minor premise: Doris has a T-Score of -3

Conclusion: Doris has Osteoporosis

Induction

A Fragility Fracture is any fall from a standing height or less, that results in a fracture. Our bodies should be able to sustain a fall from this height without a fracture unless there is an underlying cause that makes the bones fragile.

University of Michigan, Fragility Fracture Clinic

Bone is living tissue that is constantly being broken down and replaced. Osteoporosis occurs when the creation of new bone doesn't keep up with the loss of old bone. Osteoporosis causes bones to become weak and brittle.

Mayo Clinic

Major premise: Weak bones cause fragility fractures

Minor premise: Osteoporosis causes bones to become weak

Conclusion: Osteoporosis may cause a fragility fracture

Induction is a method of reasoning involving an element of probability. In logic, induction refers specifically to "inference of a generalized conclusion from particular instances." In other words, it means forming a generalization based on what is known or observed.

Merriam-Webster Dictionary

Abduction

There is a *causal link* between Osteoporosis and fragility fractures; these types of fracture are caused by weakening of the bones. Osteoporosis weakens the bones.

But if a patient has Osteoporosis then they won't definitely suffer a fragility fracture.

And if a patient does suffer a fragility fracture, this may be caused by Osteoporosis, but there are other possible causes (bone cyst, cancer, Osteomyelitis, ...)

We can explain a fragility fracture by saying that the person may be suffering from Osteoporosis

Abduction provides an explanation for the fragility fracture – a hypothesis

We traverse the inductive rule in the opposite direction

Abduction is defined as "a syllogism in which the major premise is evident but the minor premise and therefore the conclusion only probable." *Merriam-Webster Dictionary*

Major premise: Doris has a fragility fracture

Minor premise: Osteoporosis may cause a fragility fracture

Conclusion: Doris may have Osteoporosis

Chain of Reasoning

Major premise: Osteoporosis is defined by a T-Score less than -2.5

Minor premise: Doris has a T-Score of -3

Conclusion: Doris has Osteoporosis

Major premise: Weak bones cause fragility fractures

Minor premise: Osteoporosis causes bones to become weak

Conclusion: Osteoporosis may cause a fragility fracture

Major premise: Doris has Osteoporosis

Minor premise: Osteoporosis may cause a fragility fracture

Conclusion: Doris may suffer a fragility fracture

Major premise: Alendronate reduces the risk of suffering a fragility fracture

Minor premise: Doris may suffer a fragility fracture

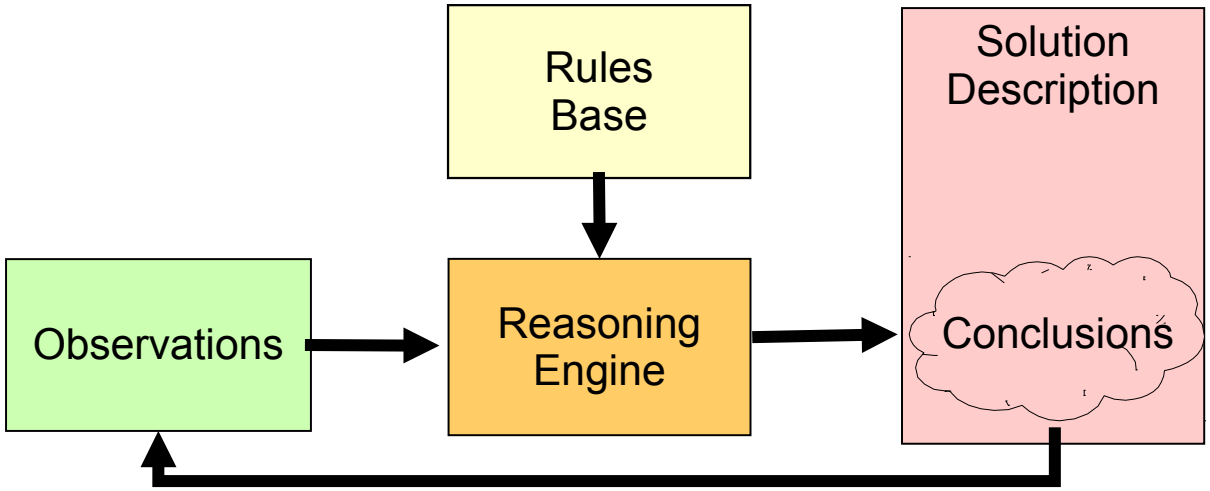
Conclusion: Doris should be prescribed Alendronate

Production Systems

Rules-based production systems as described by Davis and King in 1975.

IF T-Score < -2.5 THEN Osteoporosis

IF Osteoporosis THEN risk of fragility fracture



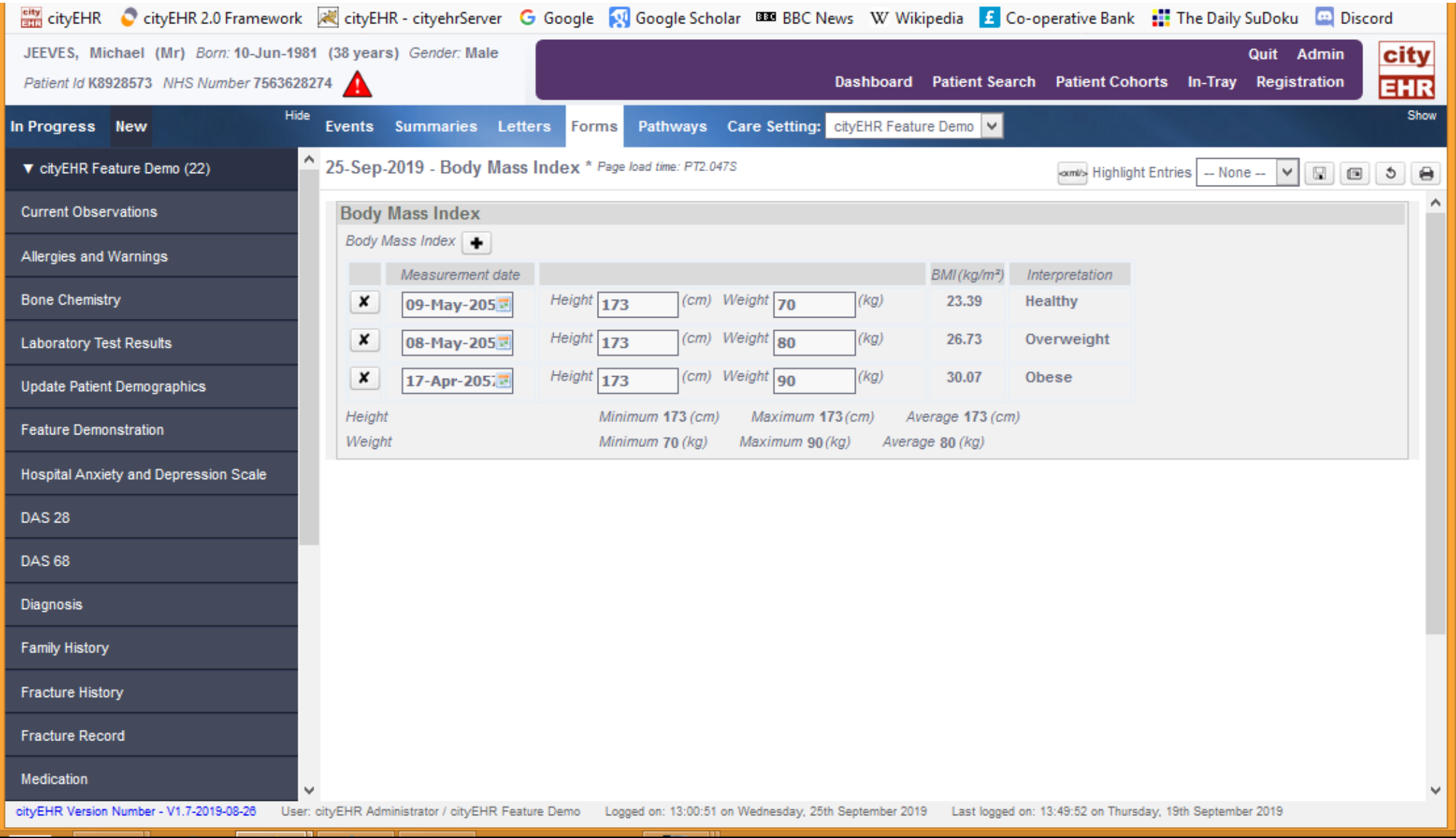
Forward-chaining system
The Reasoning Engine matches observations with the Left Hand Side of the rules, to make new assertions (Induction)

Backward-chaining system
The Reasoning Engine matches observations with the Right Hand Side of the rules, to make new hypotheses (Abduction)

Ontology-based Rules

```
<DataPropertyAssertion>
  <DataProperty IRI="#hasCalculatedValue"/>
  <NamedIndividual IRI="#ISO-13606:Element:CalculatedBMI"/>
  <Literal>
    Round-half-to-even(
        10000*BMI/WeightInKg div (BMI/HeightInCm * BMI/HeightInCm)
        ,2)
  </Literal>
</DataPropertyAssertion>

<DataPropertyAssertion>
  <DataProperty IRI="#hasCalculatedValue"/>
  <NamedIndividual IRI="#ISO-13606:Element:BMIRange"/>
  <Literal>
    if (BMI/CalculatedBMI lt 18.5) then 'Underweight'
    else if (BMI/CalculatedBMI ge 18.5 and BMI/CalculatedBMI lt 25)
      then 'Healthy'
    else if (BMI/CalculatedBMI ge 25 and BMI/CalculatedBMI lt 30)
      then 'Overweight'
    else 'Obese'
  </Literal>
</DataPropertyAssertion>
```

Using X-Forms as a Rules-Based Inference Engine



Artificial Intelligence Techniques Using Structured Health Records

Linking Health Records with Knowledge Sources
Rules-based Reasoning
Bayesian Inference

Diseases and Symptoms

A patient suffering a particular disease may display certain signs or symptoms of that disease.

For a specific observation, o and disease d it is quite easy to measure:

- the probability of o in the general population, P(o)
- the probability of o for a patient with disease d, P(o|d)
- and the probability of d in the general population P(d)

These observations are modelled in ISO-13606 as an Entry.

Bayes' Theorem

For diagnosis, we are interested in deducing the disease, given a set of observations for a patient, P(d|o), which is not so easy to measure.

$$P(d|o) = \frac{P(d).P(o|d)}{P(o)}$$

So if one in a hundred people are currently testing positive for Covid-19

And half of people diagnosed with Covid have a persistent cough

And 5% of people generally have a persistent cough at the moment

Then the probability that someone with a persistent cough has Covid is

P(d|o) = (0.01 * 0.5) / 0.05 = 0.1

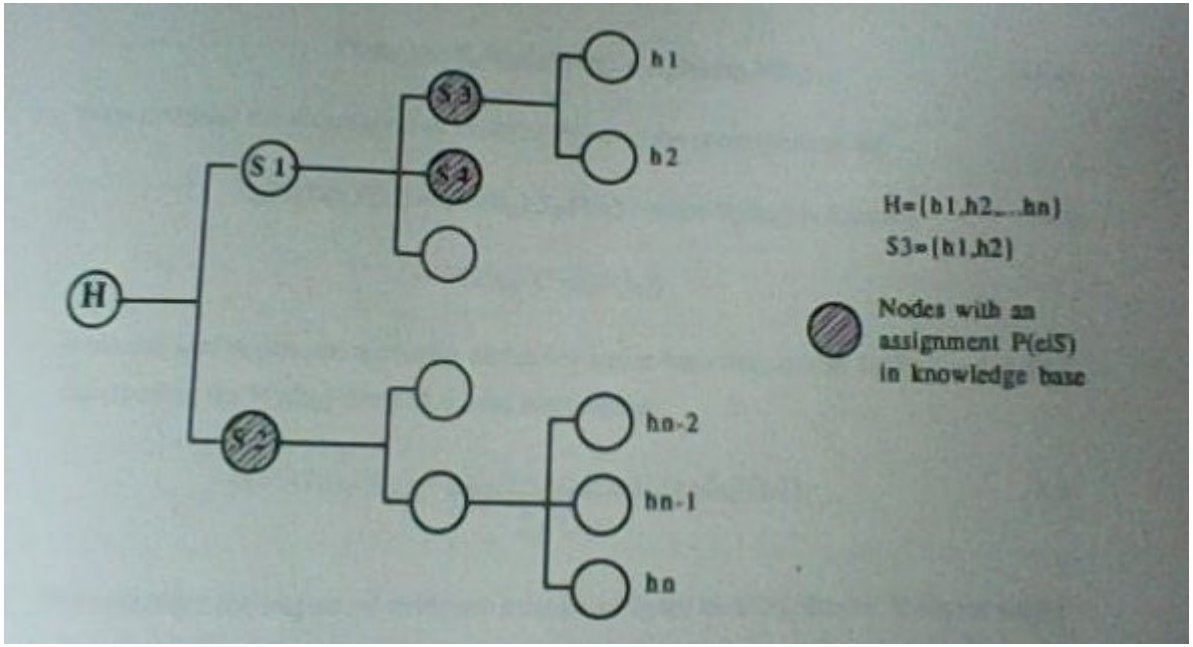
LII. *An Eſſay towards ſolving a Problem in the Doctrine of Chances. By the late Rev. Mr.* Bayes, *F. R. S. communicated by Mr.* Price, *in a Letter to* John Canton, *A. M. F. R. S.*

Dear Sir,

Read Dec. 23, 1763. I Now ſend you an eſſay which I have found among the papers of our deceaſed friend Mr. Bayes, and which, in my opinion, has great merit, and well deſerves to be preſerved. Experimental philoſophy, you will find, is nearly intereſted in the ſubject of it; and on this account there ſeems to be particular reaſon for thinking that a communication of it to the Royal Society cannot be improper.

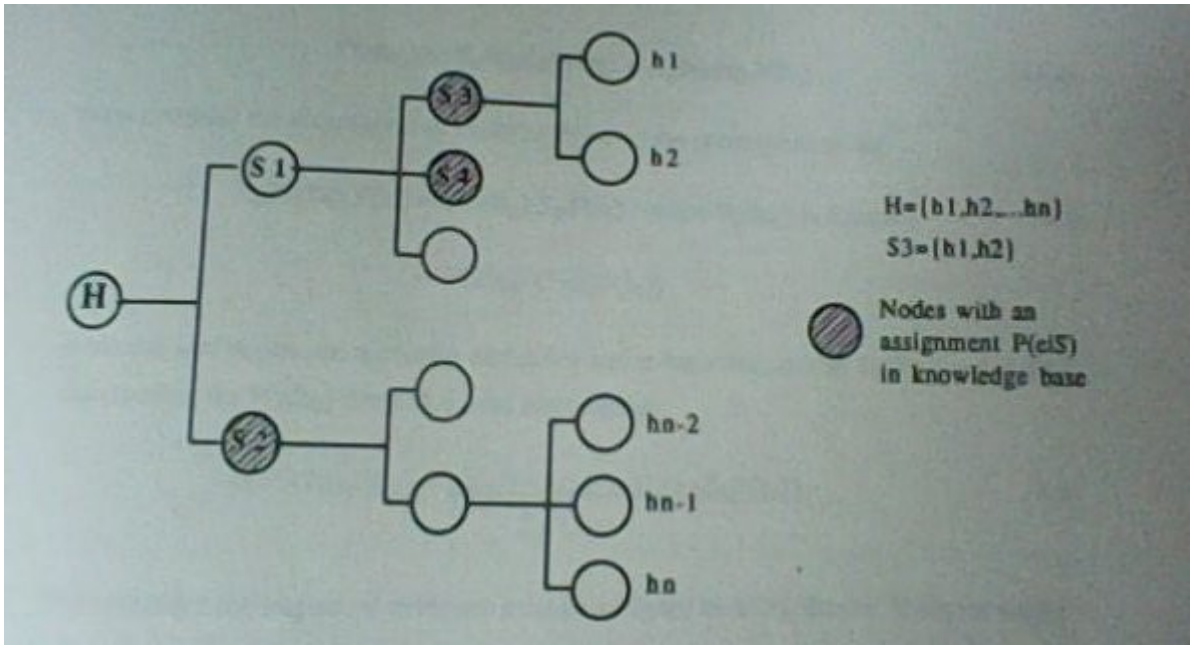
Bayesian Belief Network

Hypotheses organised as a hierarchy, in which the root node describes the class of diagnoses and the leaf nodes form a mutually exclusive and exhaustive set of hypotheses, H, for that class.



The apriori probabilities of each hypothesis sum to 1.

Bayesian Belief Network



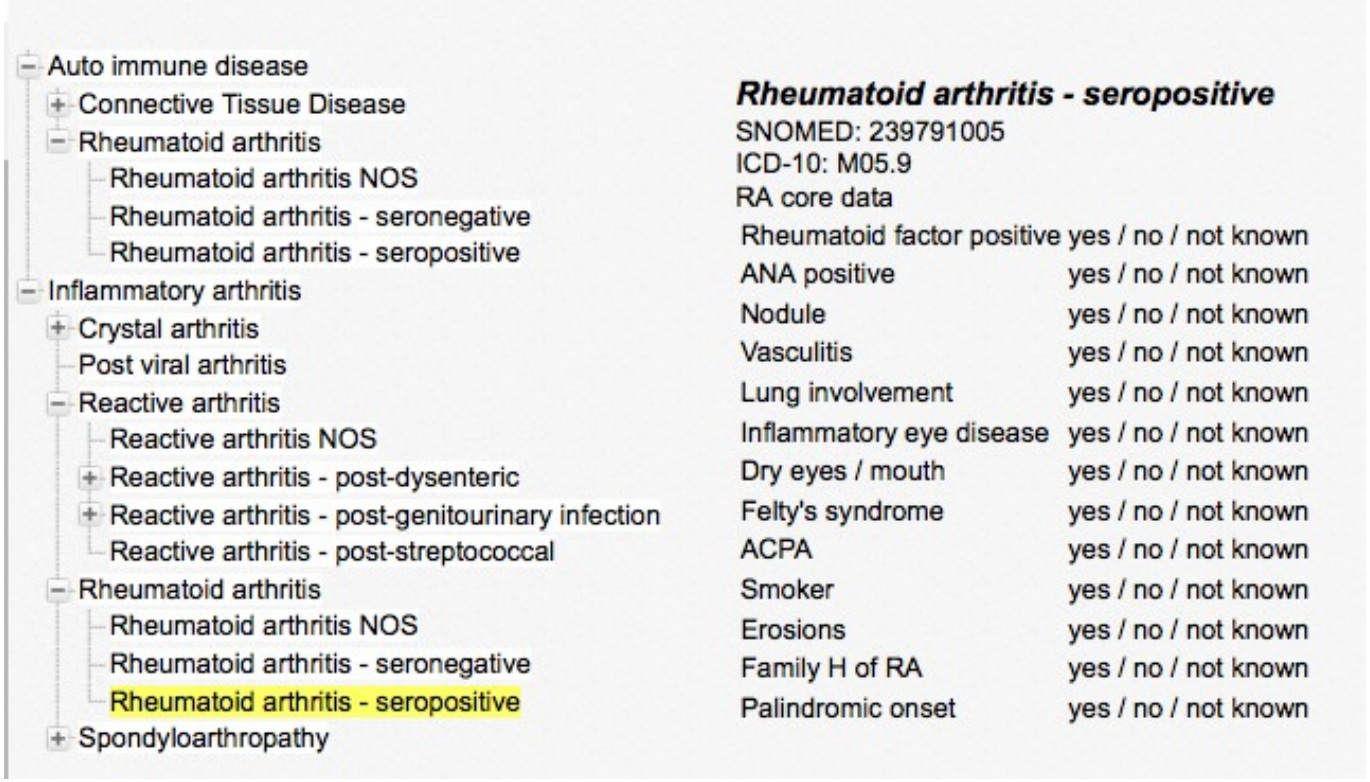
Observations provide evidence for each node in the hierarchy and we can assign the apriori probabilities of each hypothesis h, given evidence e.

So P(h|e)

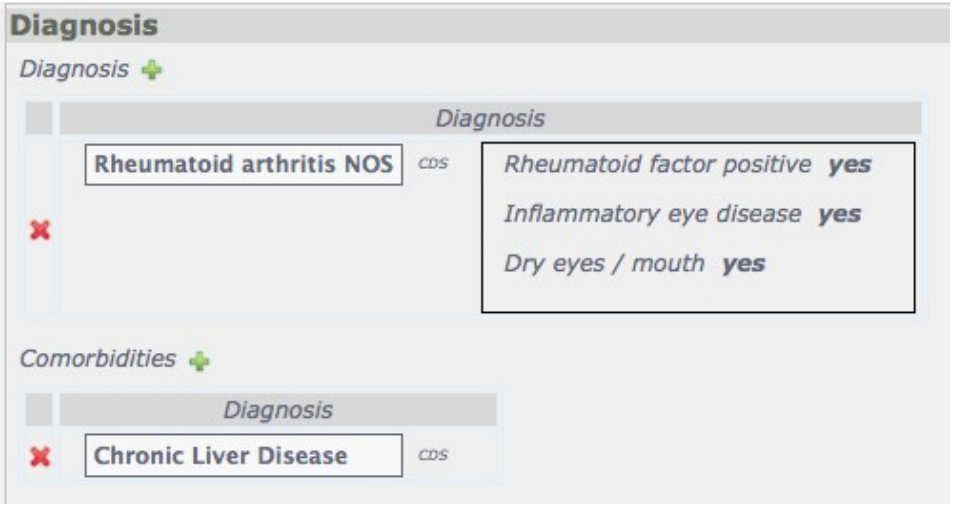
As each new observation is made, a method described by Judea Pearl (1986) allows the belief in each node of the hierarchy to be updated using the prior probabilities (and Bayes' Theorem).

Assign Prior Probabilities of Disease

Create classifications of disease
And Observation sets which define characteristics of the disease
Assign apriori probabilities to each disease class, observation and observation|class
Can then use Bayes' theorem to suggest diseases, given a set of observations



Suggest Diagnosis, Based on Evidence



As evidence is entered, suggestion(s) of the most likely diagnosis can be made, based on all relevant observations recorded to date.

John Chelsom
Seven Informatics
john.chelsom@seveninformatics.com

ST EDMUND HALL OXFORD

FORDHAM UNIVERSITY
THE JESUIT UNIVERSITY OF NEW YORK

12th – 17th September 2021

xmlsummerschool.com

XML summer school

References

Structured Health Records

Beale, T., 2002, November. Archetypes: Constraint-based domain models for future-proof information systems. In OOPSLA 2002 workshop on behavioural semantics (Vol. 105, pp. 1-69).

Camp, C.L., Smoot, R.L., Kolettis, T.N., Groenewald, C.B., Greenlee, S.M. and Farley, D.R., 2008, December. Patient records at Mayo Clinic: lessons learned from the first 100 patients in Dr Henry S. Plummer's dossier model. In Mayo Clinic Proceedings (Vol. 83, No. 12, pp. 1396-1399). Elsevier.

Chelsom, J.J., Summers, R., Pande, I. and Gaywood, I., 2011, June. Ontology-driven development of a clinical research information system. In 2011 24th International Symposium on Computer-Based Medical Systems (CBMS) (pp. 1-4). IEEE.

Dolin, R.H., Alschuler, L., Boyer, S., Beebe, C., Behlen, F.M., Biron, P.V. and Shabo, A., 2006. HL7 clinical document architecture, release 2. Journal of the American Medical Informatics Association, 13(1), pp.30-39.

ISO 13606-1:2008 Health informatics - Electronic health record communication - Part 1: Reference model. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40784

Mayne, J.G., Weksel, W. and Sholtz, P.N., 1968, January. Toward automating the medical history. In Mayo Clinic Proceedings (Vol. 43, No. 1, p. 1).

References

Knowledge and Reasoning

Ackoff, R.L., 1989. From data to wisdom. Journal of applied systems analysis, 16(1), pp.3-9.

Berners-Lee, T., Hendler, J. and Lassila, O., 2001. The semantic web. Scientific American, 284(5), pp.34-43.

Chomsky, N., 1956. Three models for the description of language. IRE Transactions on information theory, 2(3), pp.113-124.

Gruber, T.R., 1993. A translation approach to portable ontology specifications. Knowledge acquisition, 5(2), pp.199-220.

McCulloch, W.S. and Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, 5(4), pp.115-133.

Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review, 65(6), p.386.

Shannon, C.E., 1948. A mathematical theory of communication. The Bell System Technical Journal, 27(3), pp.379-423.

W3C Owl Working Group. 2012. OWL 2 Web Ontology Language. Document Overview (Second Edition) W3C Recommendation 11 December 2012.

References

Artificial Intelligence

Chelsom, J.J., Ellis, T.J., Carson, E.R. and Cramp, D.G., 1988, November. Blood gas analysis: a knowledge-based adviser for the interpretation of results. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 1360-1361). IEEE.

Jain, A.K., Mao, J. and Mohiuddin, K.M., 1996. Artificial neural networks: A tutorial. Computer, 29(3), pp.31-44.

McCarthy, J., Minsky, M.L., Rochester, N. and Shannon, C.E., 2006. A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. AI magazine, 27(4), pp.12-12.

McCarthy, J., 1960. Recursive functions of symbolic expressions and their computation by machine, Part I. Communications of the ACM, 3(4), pp.184-195.

McCarthy, J., 1989. Artificial intelligence, logic and formalizing common sense. In Philosophical logic and artificial intelligence (pp. 161-190). Springer, Dordrecht.

Shortliffe, E.H., 1974. MYCIN: a rule-based computer program for advising physicians regarding antimicrobial therapy selection. STANFORD UNIV CALIF DEPT OF COMPUTER SCIENCE. Available from: http://people.dbmi.columbia.edu/~ehs7001/Shortliffe-1976/MYCIN%20thesis%20Book.htm

Turing, A.M., 1950. Computing machinery and intelligence. Mind, 59(236), pp.433-460.

Werbos, P.J., 1990. Backpropagation through time: what it does and how to do it. Proceedings of the IEEE, 78(10), pp.1550-1560.

References

Artificial Intelligence Techniques Using Structured Health Records

Bayes, T., 1763. LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S. Philosophical transactions of the Royal Society of London, (53), pp.370-418.

Buchanan, B.G. and Shortliffe, E.H., 1984. Rule-based expert systems: the MYCIN experiments of the Stanford Heuristic Programming Project.

Davis, R. and King, J., 1975. An overview of production systems. Stanford Univ Ca Dept of Computer Science.

Lassila, O. and Swick, R.R., 1998. Resource description framework (RDF) model and syntax specification.

Pearl, J., 1986. Fusion, propagation, and structuring in belief networks. Artificial intelligence, 29(3), pp.241-288.

Post, E.L., 1943. Formal Reductions of the General Combinatorial Decision Problem. American Journal of Mathematics, 65(2), pp.197-215.