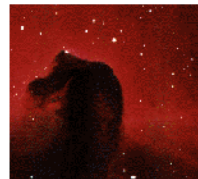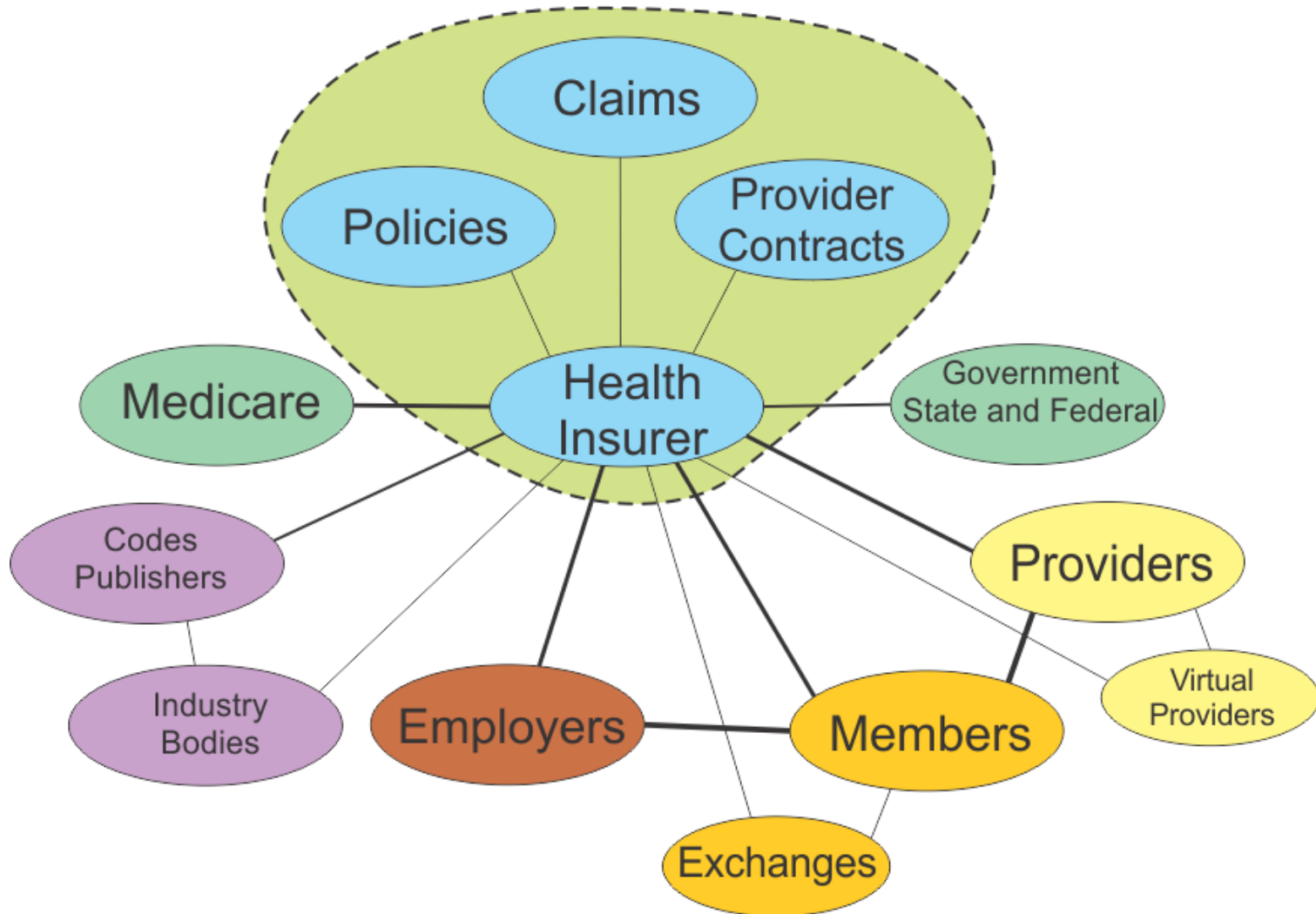# Big Knowledge

US Health Insurer

Knowledge Interactions

Policies

Policies are a combination of mostly simple medical terms – vaccination, dialysis – and legalese. The legalese includes limits and exclusions, some with exceptions.

Most policies are mostly the same, but that can lead to "smearing", where a policy you don't know is assumed to be the same as one you do.

Another problem is Invalid Knowledge Retention – if you have spent years learning what the policies mean, it is hard to forget it when the details change.

Four Pieces

People are limited to no more than four pieces of information in play at once – more than that and they start chunking, which loses the detail.

It seems a very small number, but if you observe their performance when distracted or fatigued (or upset by a disgruntled caller), it can get down to just one piece.

The number has important consequences when expecting people to handle large amounts of changeable knowledge

Reliability

It is easy to understand the need for reliability for an aircraft – the consequences of failure can be dire, but a call centre – not so much.

An insurer gets tens of millions of calls about benefits, and with a reliability of 85-90%, that is millions of wrong answers.

If they are wrongly told the insurer isn't going to pay, people may forgo lifesaving surgery, or if told yes, go ahead with treatment, only to be bankrupted when their claim is denied. Pretty dire outcomes.

Given all the potential sources of error, what would be a reasonable reliability?

Avoiding Limits

So how do we avoid the problems of smearing, invalid long term memory and a limited capacity for handling many pieces of information at once?

We could use something else to handle knowledge –a machine

Can we turn words into active knowledge outside ourselves?

Ontologies

We could set up an ontology to store the knowledge.

Well, no. Much of it is conditional – unless, because, when, except.

Elements of the policy are controlled by properties of the member – sex, age, ethnicity, although some policies allow mammograms for males, so very little is absolute.

Some of it changes every year – for ObamaCare, a lot.

Ontology Problems

CPT Code (Current Procedural Terminology)

> CPT 93229
> Wearable mobile cardiovascular telemetry with electrocardiographic recording, concurrent computerized real time data analysis and greater than 24 hours of accessible ECG data storage (retrievable with query) with ECG triggered and patient selected events transmitted to a remote attended surveillance center for up to 30 days; technical support for connection and patient instruction for use, attended surveillance, analysis and physician prescribed transmission of daily and emergent data reports

A policy has a limited repertoire – office visits, outpatient hospital, inpatient hospital, home care, durable medical equipment
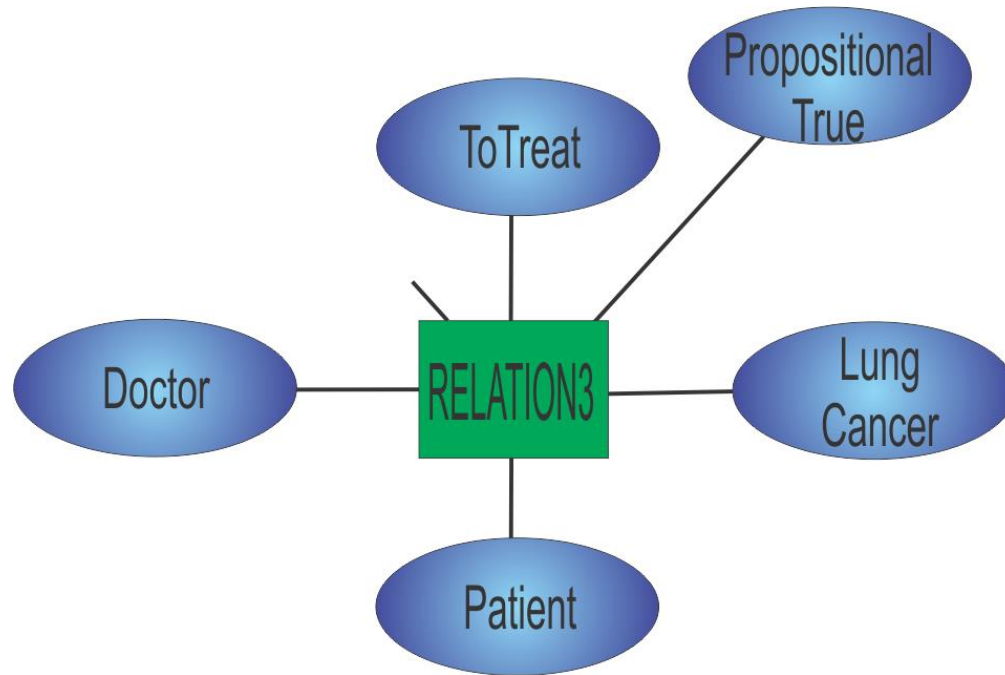
Remote continuous cardiac monitoring through the mobile phone network doesn't fit anywhere, so it has to be bolted on somewhere

Imagine that there are well over 100,000 similar codes, describing every aspect of medical diagnosis and treatment. Provider staff roll up treatments into these codes, and then the system must unroll them into services mentioned in the policies

Semantic Structure

For a machine to use,
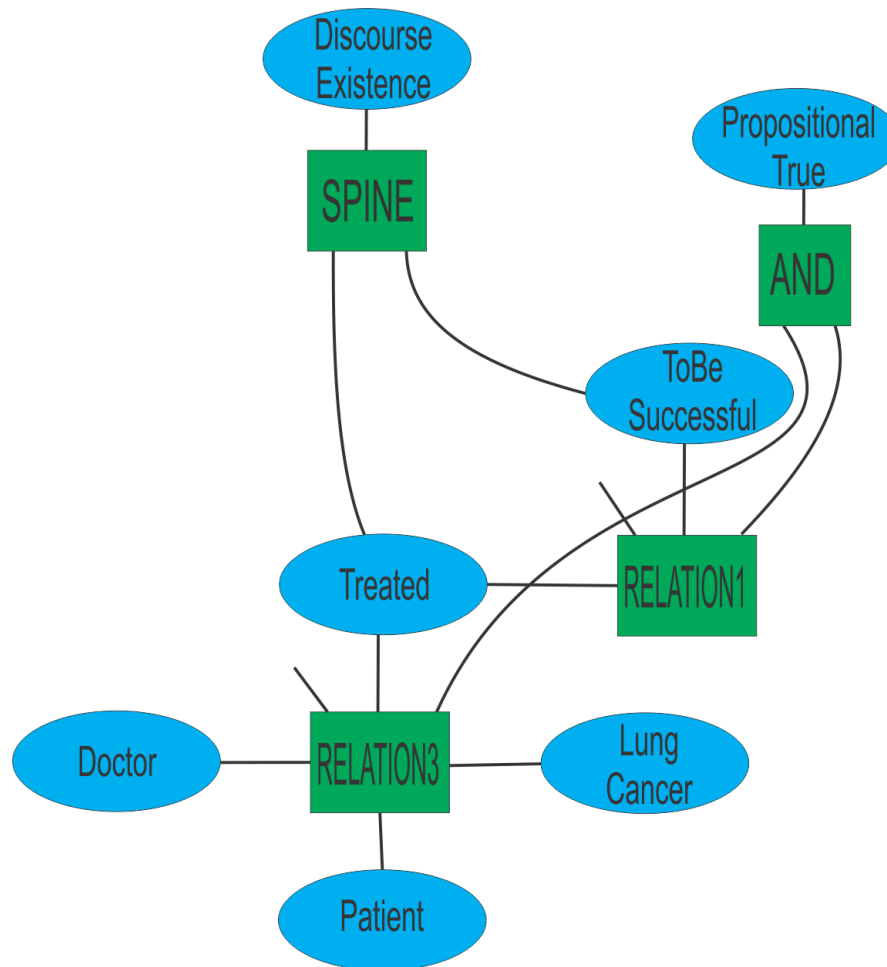not for us to look at

Simple Structure

The doctor treated the patient for lung cancer

A simple true statement

Overwhelming Detail

A simple relation – in the midst of webs of:

- Propositional logic
- Existential logic
- Structure map
- Discourse control
- Inheritance of properties

We can only look at any small piece without being overwhelmed by the mass of connections
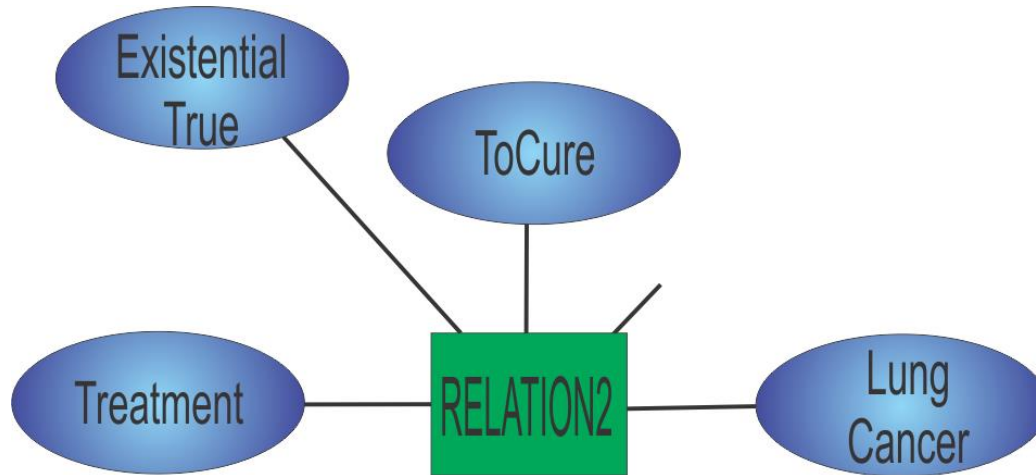
A Growing Discourse

The doctor treated the patient for lung cancer.
The treatment was successful.

Statements in Documents

Statements come in three types:

- Statements about relations among objects

- Statements about the existence of relations among objects – can, is able to

- Statements about the existence of statements

Existence of Relations

The treatment can cure lung cancer

The relation exists, but its truth is not asserted
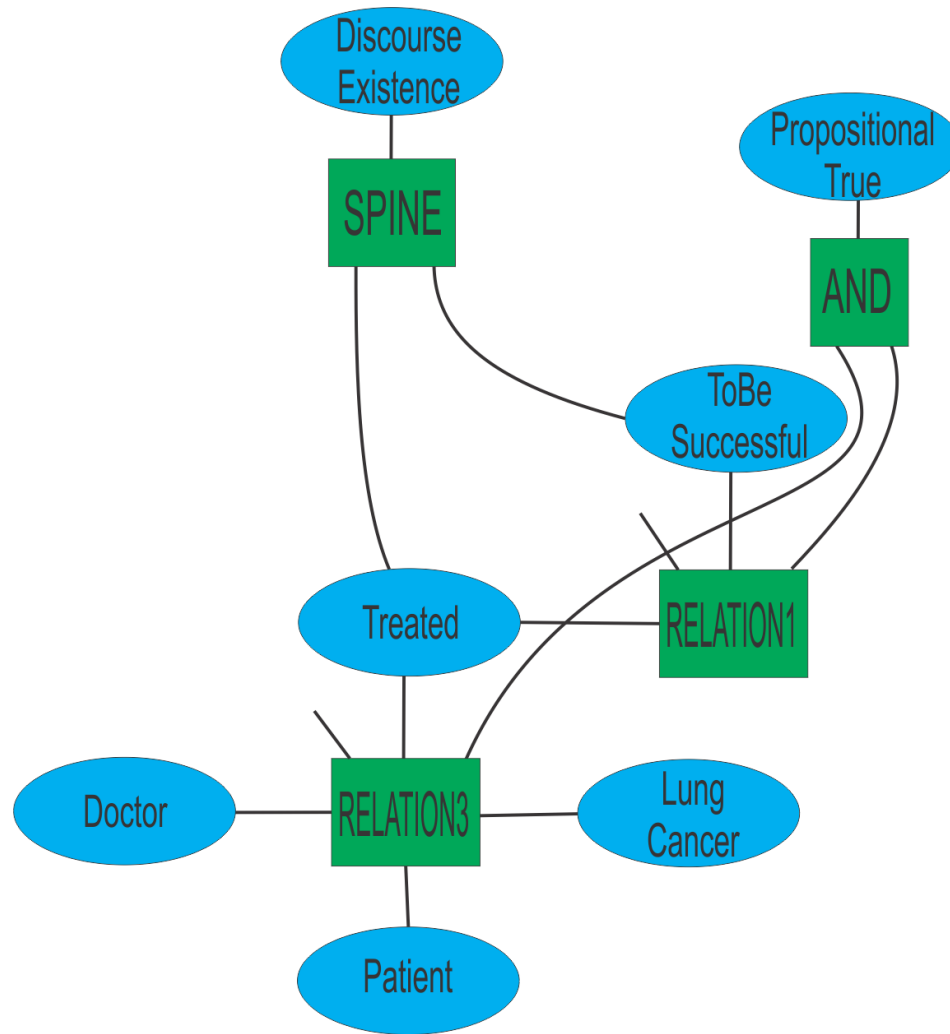
A Statement about Statements

Section 9 is void under the following …

The following item does not apply if…

We need to be able to control the existence of text from a distance – something in one section can control whether some part, or all, of another section has meaning – sometimes even the whole document:
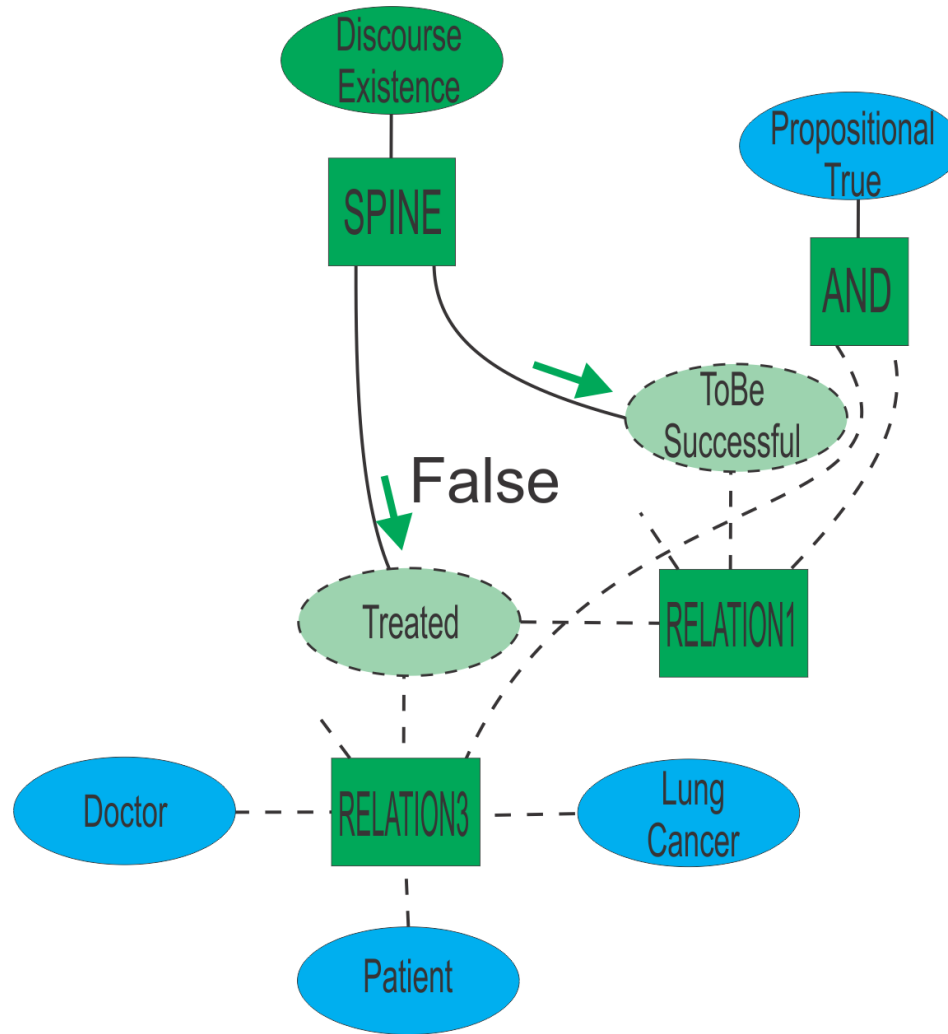
This policy is void if…

Now You See It

The doctor treated the patient for lung cancer.
The treatment was a success.

Now You Don't

Changing the discourse existence to false makes the structure representing the statements disappear (and their truth irrelevant)

Complete Semantic Representation

Now add in all the other aspects of language:

- If, but, except, only, when
- Groups
- Prepositional chains
- Indexed lists

and you have a complete and accurate representation of a policy.

Large documents can be seen as clanking machinery made of a million pieces, with bits switching on and off – they are not just words on a page

Semantic Intricacies
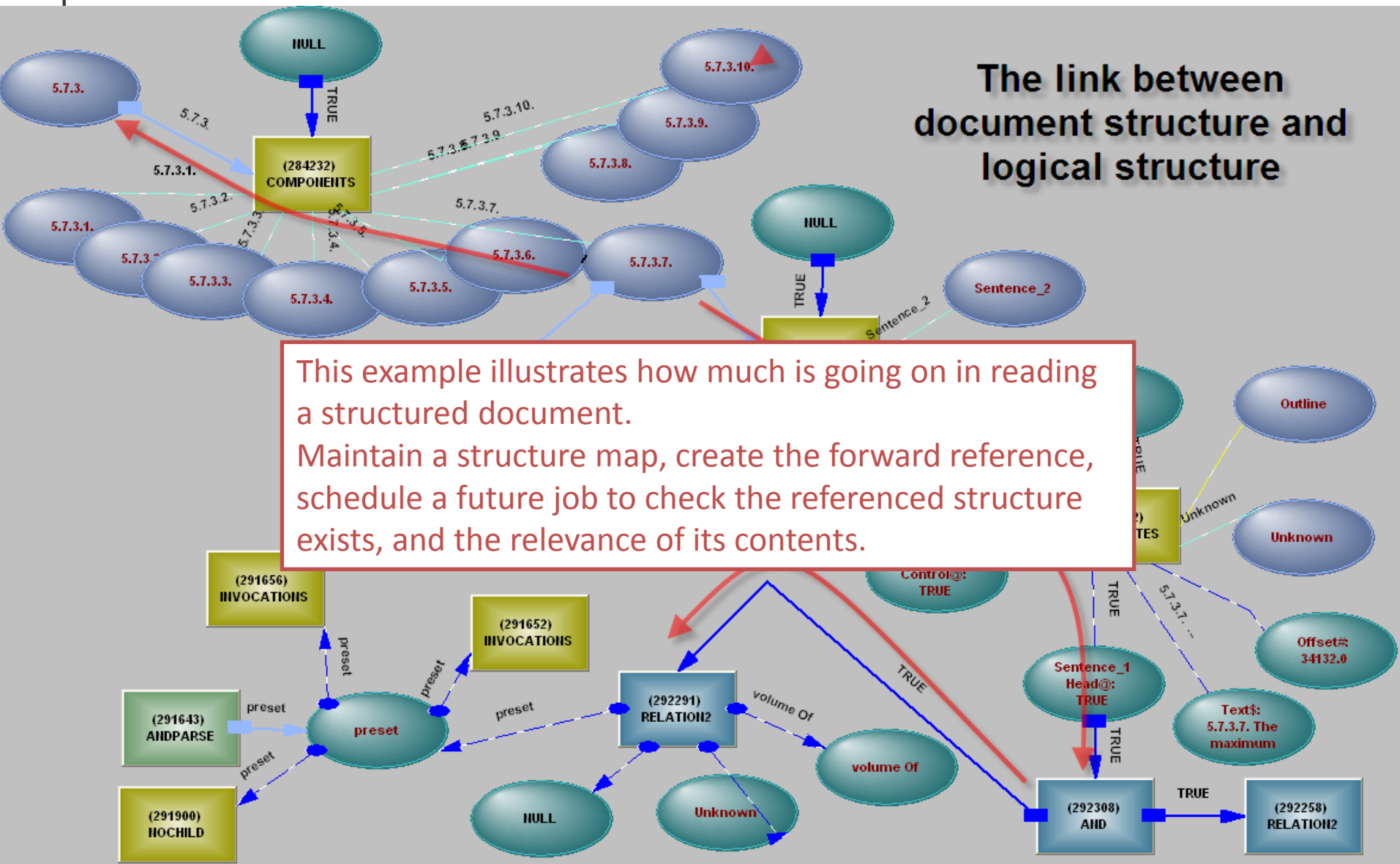
Discourse Structure

A policy is a structured document, with sections pointing to each other and to external materials (Federal and State law, Medicare, authorisation lists). Statements garner applicability from the headings and subheadings they appear under.

Simple word search doesn't work well – there might be six references to Maternity Care, one of which says "Pregnancy is not covered", but you have to know what section that is in, as it is not a general statement.

In other words, you have to know the details of the policy, but there are too many policies to know.

Pinpoint Precision

The link between document structure and logical structure

This example illustrates how much is going on in reading a structured document.

Maintain a structure map, create the forward reference, schedule a future job to check the referenced structure exists, and the relevance of its contents.

Groups

Groups are created everywhere and with anything:

Objects - Jack and Jill went up the hill

Verb Relations – Jack fell down and broke his crown

Clauses – Jack fell down … and Jill came tumbling after

Prepositions – from, to and between hospitals

Subordinate conjunctions – if and when the claim is made

Groups created within the document provide dynamic fluidity to a context

Layering

Clausal relations (think, consider, deem, propose) create a mini-discourse

It was deemed that the claim was paid

Logical control of the mini-discourse can be overridden without error

John thinks he is guilty, but he is innocent

Layering can be without limit:

John thought that Fred wanted him to think that Olga was guilty

Particular Overrides General

Covered expenses include:
>   *Surgery,* including anesthesia.

Great, a general statement saying surgery is covered!

No benefits will be provided for:
- Hyperhidrosis surgery;
- Cosmetic surgery and cosmetic services or devices, unless for reconstructive surgery:
- Dental services, appliances or supplies for treatment of the teeth, gums, jaws or alveolar processes, including but not limited to, any *oral surgery* or periodontal surgery
- Any treatment, including but not limited to surgical procedures:
    - For obesity, which includes *morbid obesity*; or
    - For obesity, which includes *morbid obesity*, for the purpose of treating a *sickness* or *bodily injury* caused by, complicated by, or exacerbated by the obesity.

You have to check everywhere to be sure

Error Handling

Inconsistencies

ROUTINE/PREVENTIVE CHILD CARE SERVICES
BIRTH TO AGE 18
(*Services* Received at a Clinic or Outpatient *Hospital*)

MEDICAL SERVICES	PAR PROVIDER BENEFIT	NON-PAR PROVIDER BENEFIT
Routine Child Care Examination	100%	70% after *deductible*
Vision Screening	100%	70% after *deductible*
Hearing Screening	100%	70% after *deductible*

Policies are long-lived documents – cut and paste them every year and they get into a mess

LIMITATIONS AND EXCLUSIONS

This Plan does not provide benefits for:

51. Routine vision examinations, *services* to correct eye refractive disorders

52. *All services* over the maximum lifetime benefit of $5,000 per person for the treatment of *morbid obesity*. Complications resulting from these services will be covered the same as any other illness.

53. Services in excess of the maximum lifetime benefit of $5,000 per person for you and your covered dependent spouse for the treatment of infertility.

54. The purchase, fitting or repair of eyeglass frames and lenses or contact lenses, unless specifically provided under this Plan;

55. Routine hearing examinations;

56. Routine hearing testing;

Verification and Validation

Verifiability is necessary for a requirement but there are other important issues. A requirement can be verifiable yet incorrect; and assessing verifiability alone will not detect incorrect requirements. Moreover, verification is totally irrelevant with regard to a requirement which has been overlooked. **Mere analysis, inspection, or review alone will find some of these issues but generally is far weaker than usually is realized** – Wikipedia

People don't read stuff

– make it full of legalese and

their enthusiasm wanes even more

Fixing Errors

The system has to find errors automatically – finding them manually is just too hard – and there are about ten million pages, which get changed a bit every year.
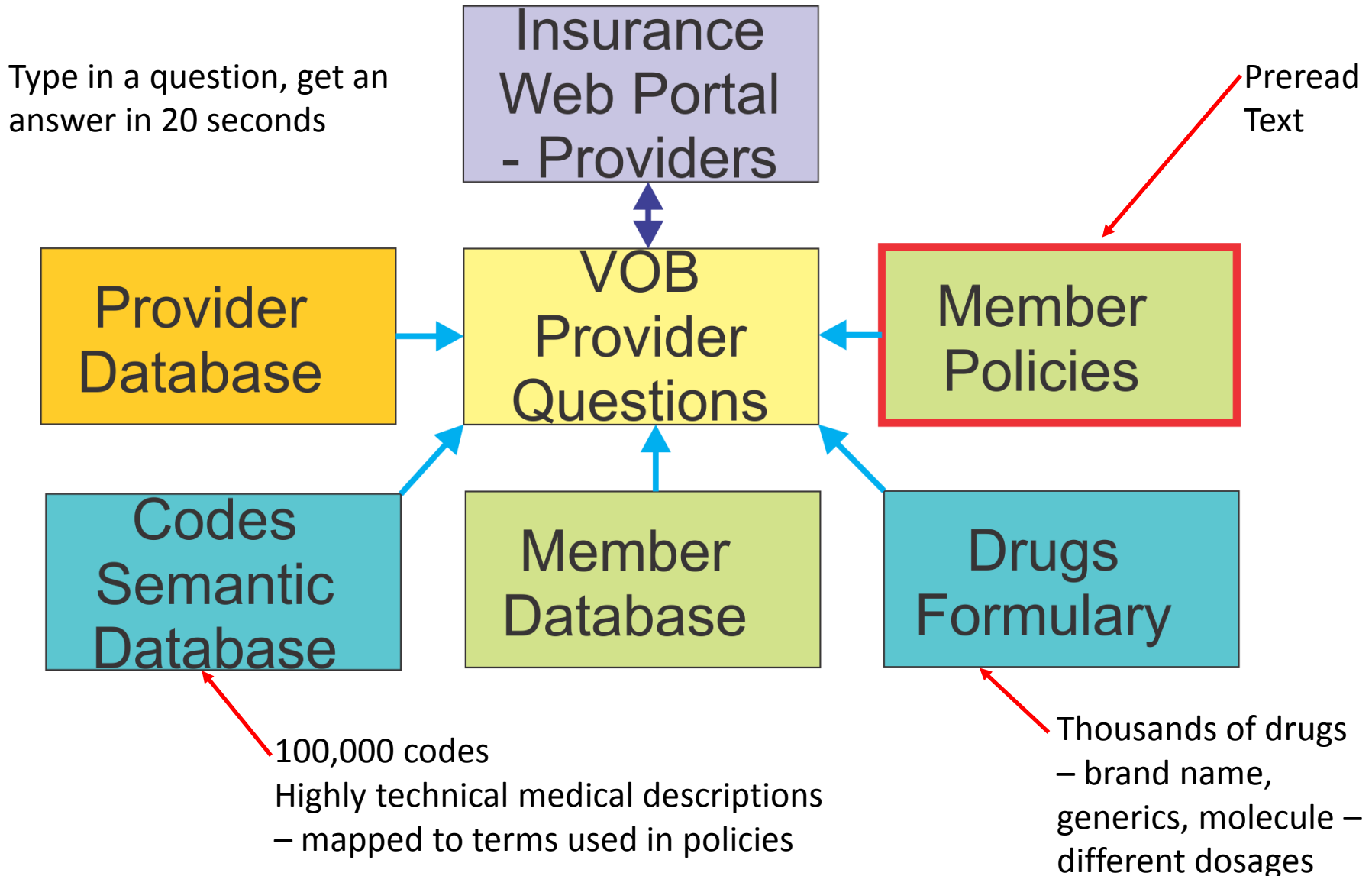
Company-Wide Knowledge Asset
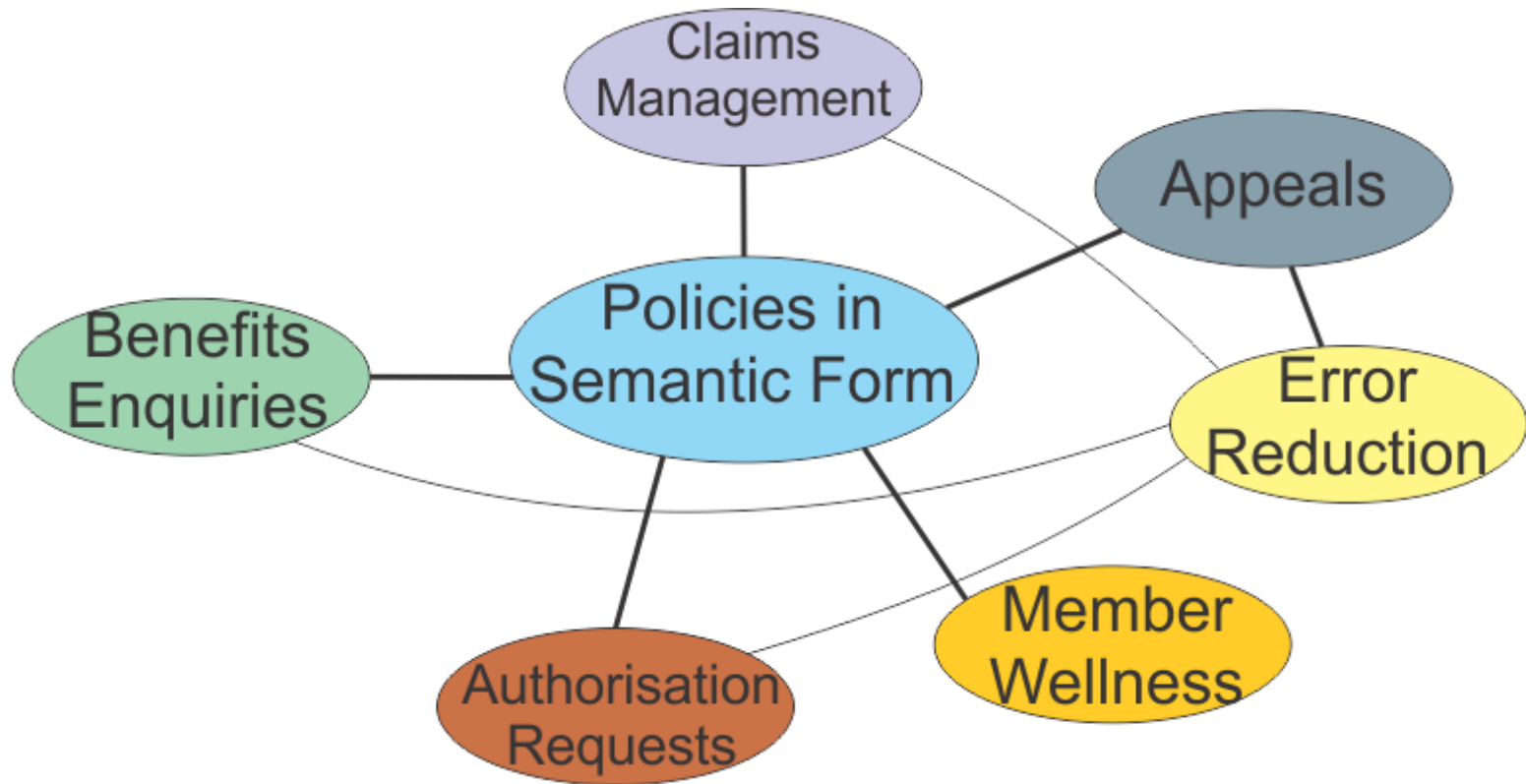
Knowledge held in semantic form can be accessed faster and has fewer errors, leading to fewer appeals

"Big Knowledge" Systems

Web Verification of Benefits Questions

Type in a question, get an answer in 20 seconds

Insurance Web Portal - Providers

Preread Text

Provider Database

VOB Provider Questions

Member Policies

Codes Semantic Database

Member Database

Drugs Formulary

100,000 codes
Highly technical medical descriptions – mapped to terms used in policies

Thousands of drugs – brand name, generics, molecule – different dosages

Handling Questions

The system automatically builds a semantic structure from the policy – for 200 pages, it might take a machine 4 hours to do (so it has to be done long before the question arrives)

A semantic structure is built from the question – it might take 2 seconds

One structure is matched against the other and the logical states or arithmetic values (in the specific policy or database entry) provide the answer

Scale of the US Healthcare System

The US Health Insurance industry turns over $400 billion, and spends about $60 billion on administration costs. But there are costs elsewhere as well

It is estimated that US doctors spend 3 weeks of their year hanging on the phone to health insurers – change that and seven billion dollars of public good would be saved
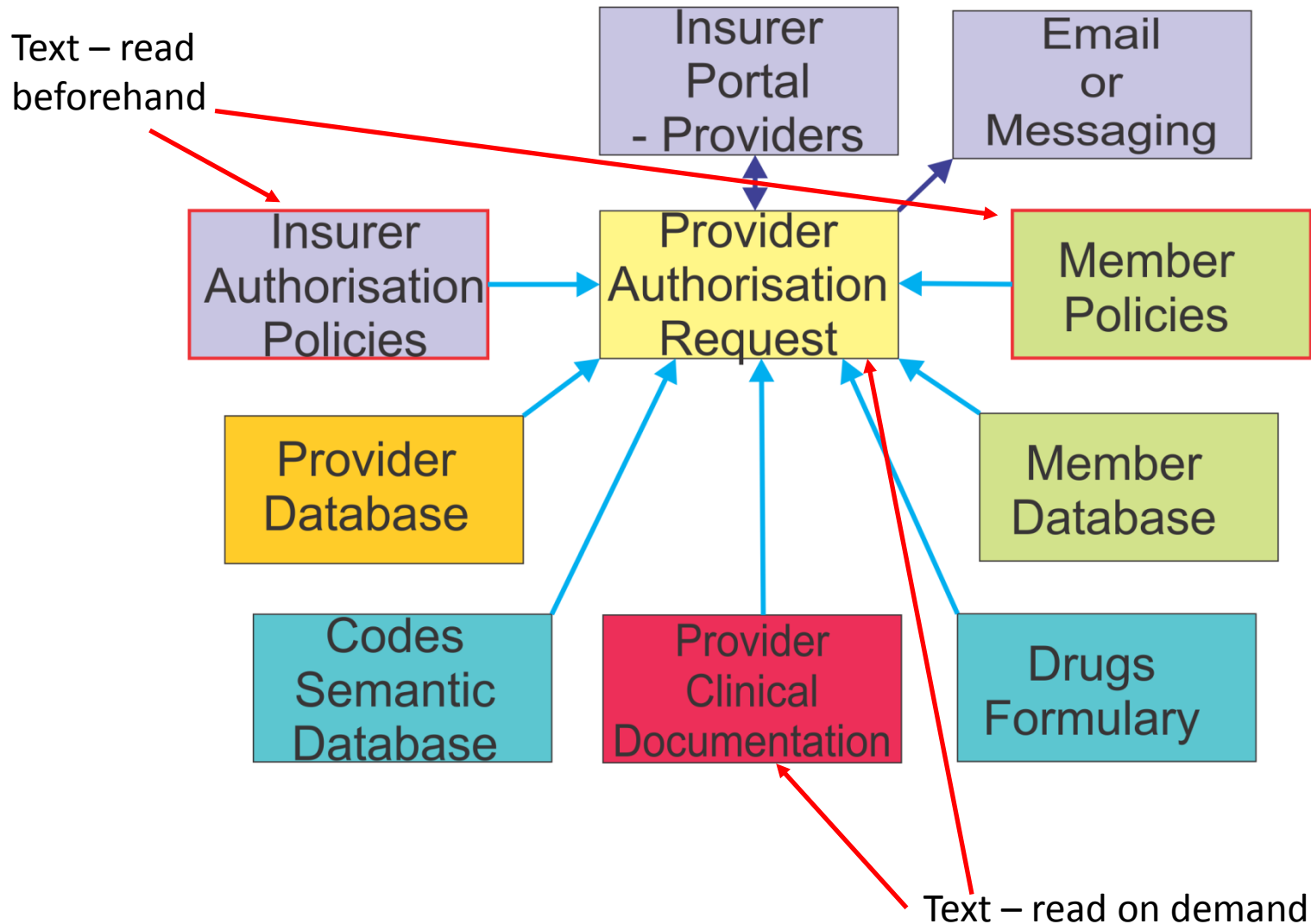
Providers are used to a phone interface, so maybe a text-based web interface won't cut it

Phone VOB Provider Questions
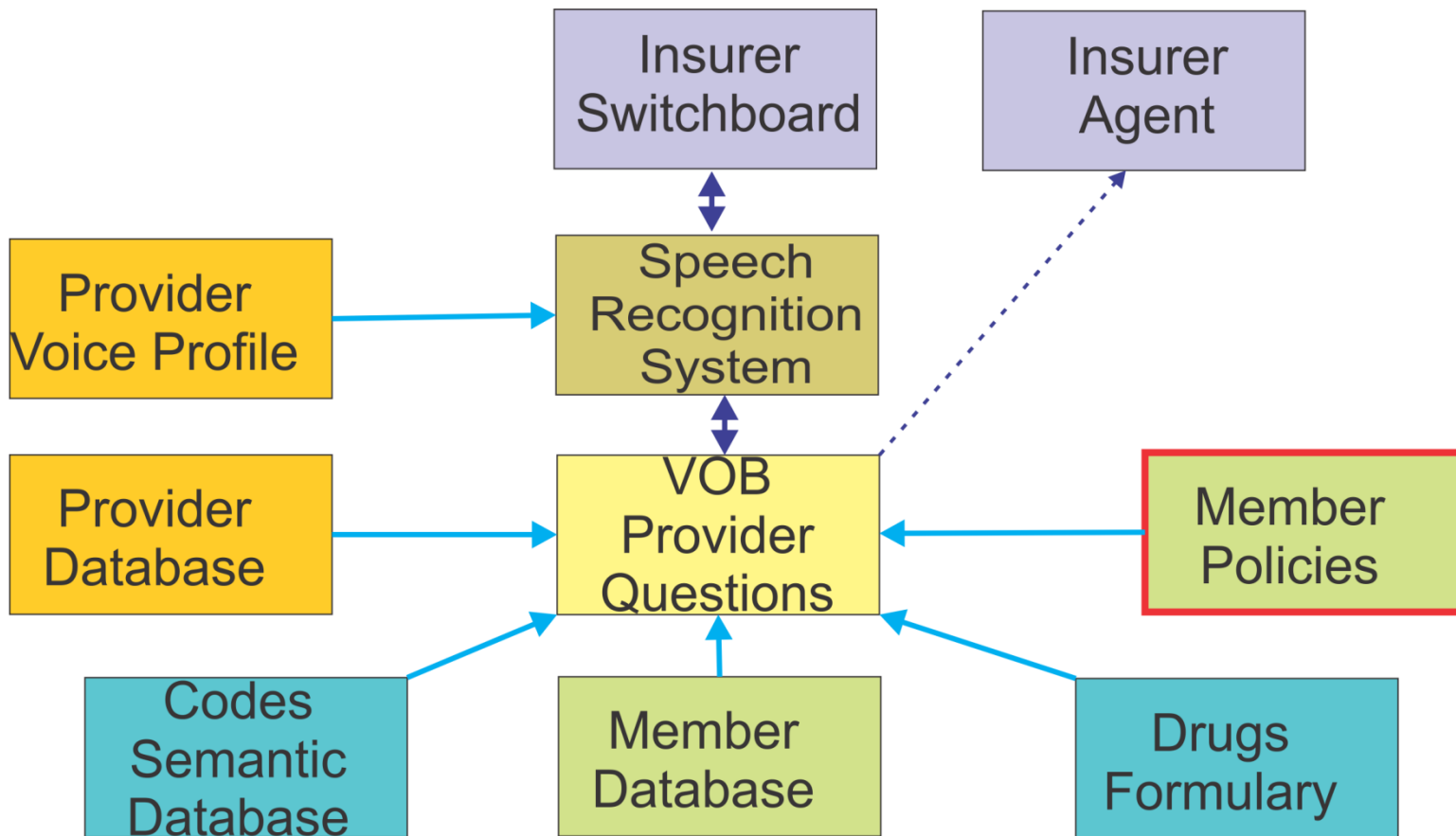
The speech recognition system starts with 75% accuracy, but can get to 98% if it is trained to the person's voice and has considerable post-recognition fixup (its predictable errors, homonyms)

Provider Authorisation over the Web

Clinical Documentation

A medical note:

Sections show a moderately differentiated squamous cell carcinoma with cystic change. The tumour extends to the visceral pleura but unequivocal pleural invasion is not seen in the routinely stained sections. An elastic stain has been requested for further evaluation of the pleura and a supplementary report will be issued with the result. Lymphatic invasion is not identified and blood vessel invasion is not seen. The bronchial resection margin shows squamous metaplasia but no evidence of in situ or invasive carcinoma. The vascular resection margin is negative for carcinoma. Non-neoplastic lung tissue shows centrilobular emphysema and smoker's-type macrophages. A section from the tip of the lobe confirms the presence of honeycomb change and shows associated bronchiolar metaplasia. Anthracotic dust macules are seen with associated emphysema. Additional sections of non-neoplastic lung tissue will be examined for further evaluation. Metastatic carcinomas is not identified.
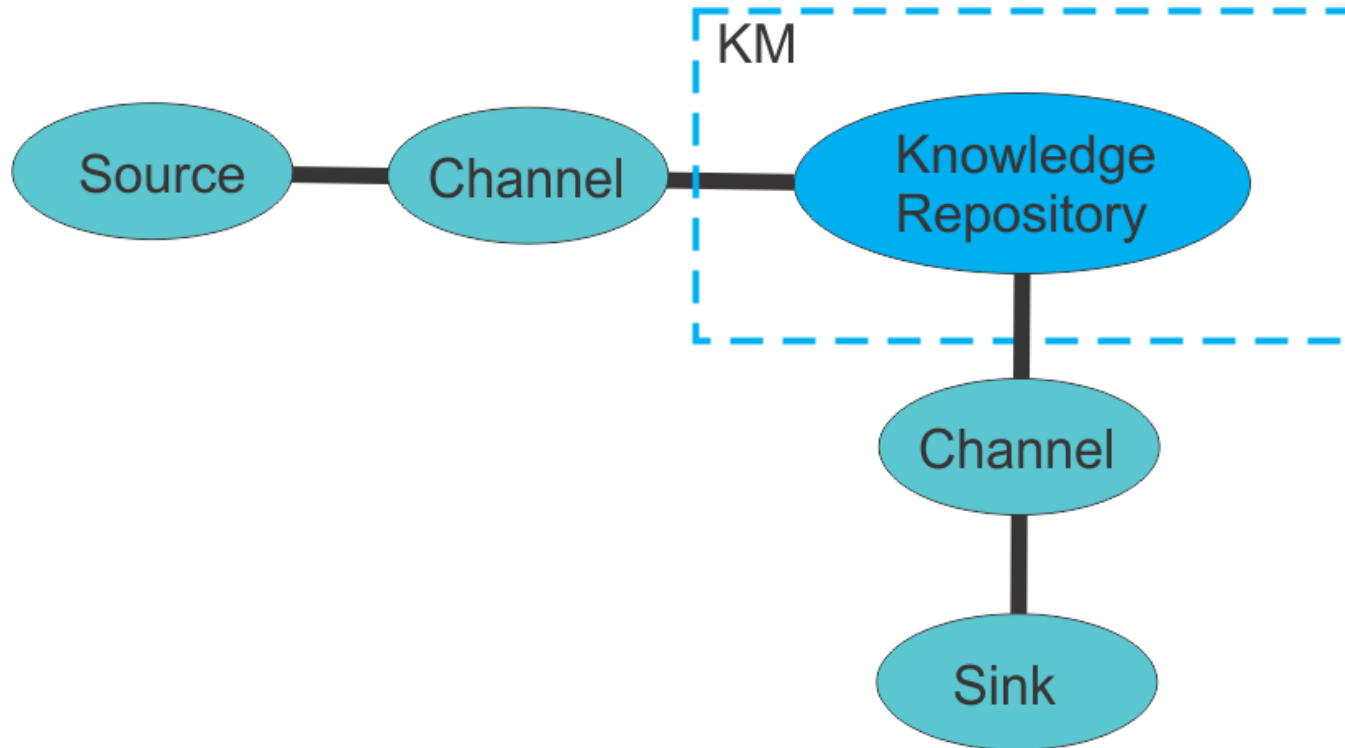
From AEHRC website (this is in range of automatic reading)

Most supporting clinical documentation will be much less complex than this – although a course of drugs might cost $250,000 or a transplant $500,000 - so sometimes "automatic analysis for authorisation" may not be a good idea
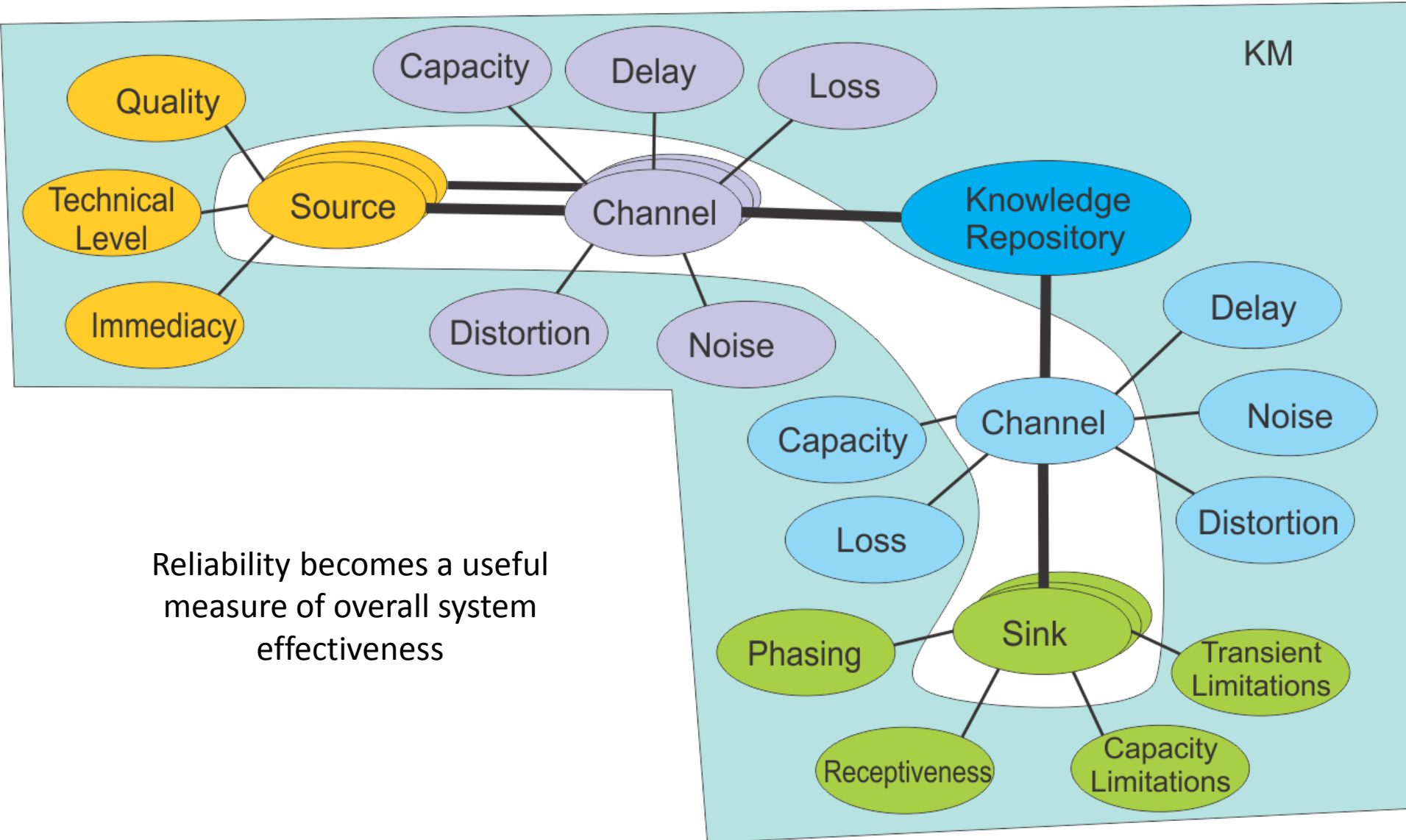
Language Technical Level
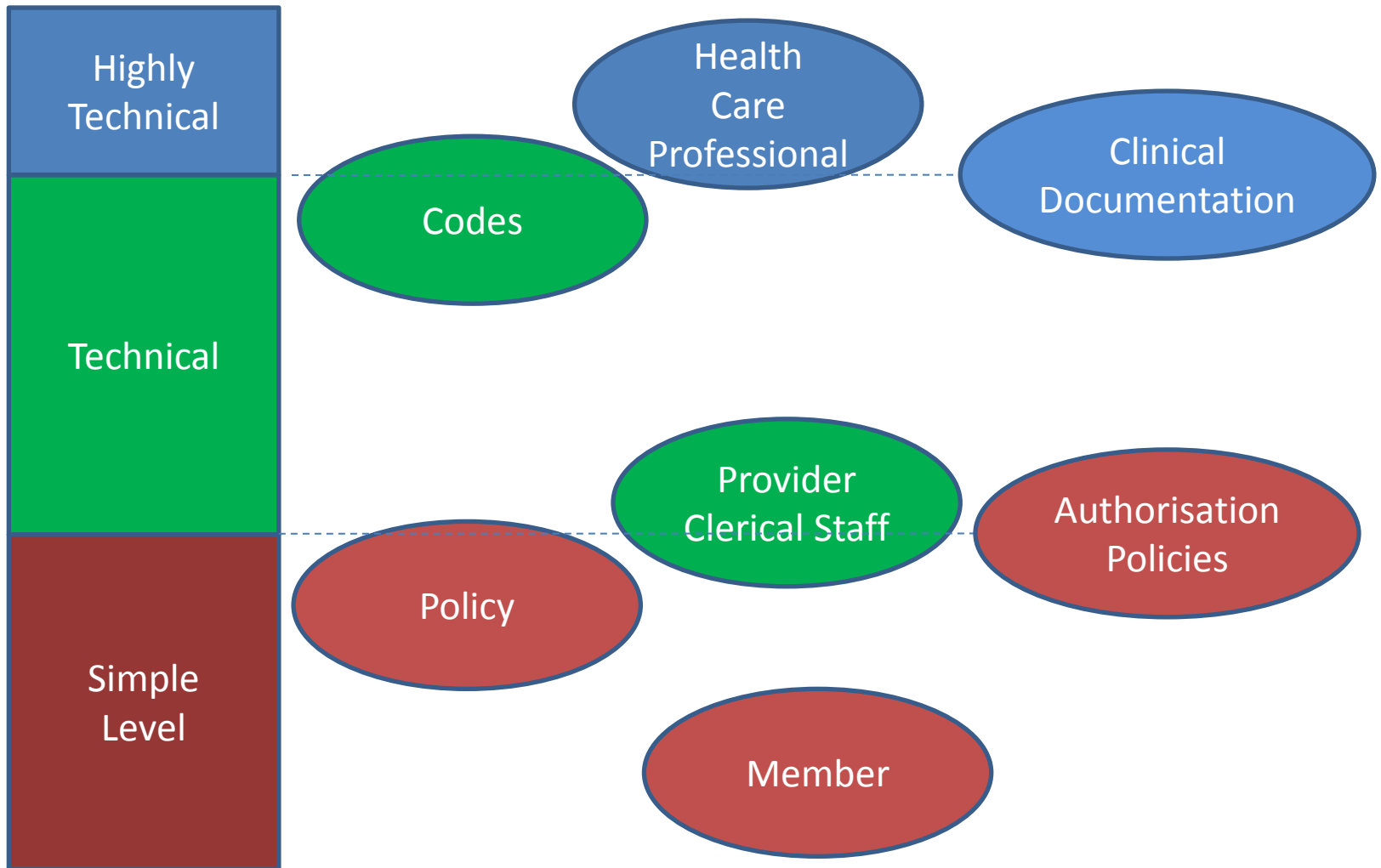
One View of KM

I can get knowledge from somewhere, I can hold it, I can give it to someone
Its quality or level or immediacy is largely irrelevant to me – I am not responsible

Another View of KM

Reliability becomes a useful measure of overall system effectiveness

Areas Touched On

- Large-scale Knowledge Interactions
- Knowledge Representation
- Knowledge Density
- Knowledge Quality
- Language Level
- Reliability of System
- Responsibility