

## FROM....

# Towards Design Principles for Health Information Systems

**Thomas Beale**

Deep Thought Informatics Pty, Ltd  
Mooloolah, Qld, Australia  
(thomas@deephought.com.au)

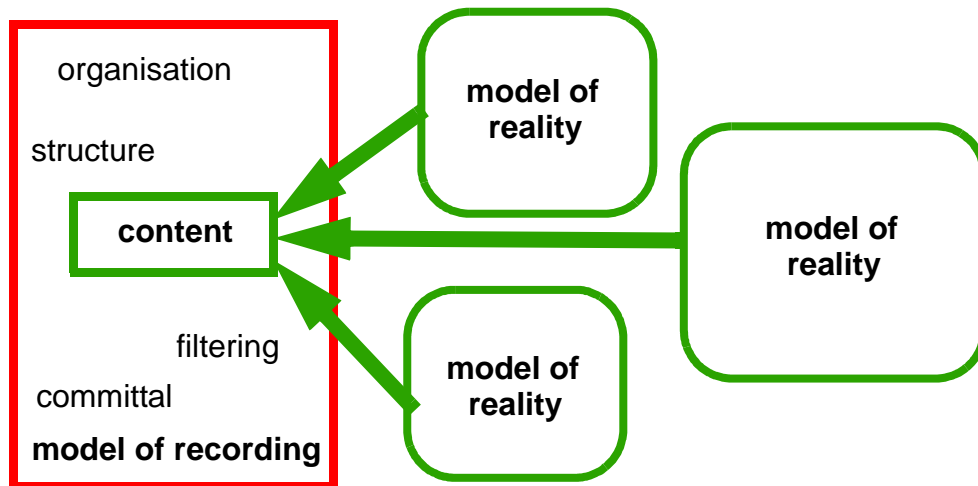
In print: International Journal of Medical Informatics, 2003.

## CONTEXT: Recording versus Reality

A number of models for health information systems have been published as *de jure* standards or *de facto* models by bodies such as national health departments. They contain two kinds of semantics which can be labelled “recording” and “reality”. Models such as the CEN ENV 13606 EHR architecture standard, the GEHR and *openEHR* EHR specifications, and the HL7 Clinical Document Architecture (CDA) Framework are largely models of recording, in the sense that they explicitly define semantics for the activities which occur due to recording phenomena, including: *capturing* the data, *organising* it, *committing* it to a persistent store, and providing some kind of *directory* organisation of the committed items, with appropriate contextual (audit trail) attributes being included at each level. On the other hand, models such as the OMG’s PIDS specification [5] (a demographic model), the HL7v3 Reference Information Model (left hand side) [REF], models of acts e.g. Riche/Nucleus [REF], and the more recent HL7v3 RIM (right hand side), and models of other things such as workflow, guidelines and so on, could be termed models of “reality” in the sense that they try to capture actual state, actions, process, intentions or other aspects of the real world, or thinking about things in the real world. They do not, in themselves, try to define semantics for capturing such information in a recording process such as takes place in the EHR (in fact such semantics, e.g. of audit trailing changes are largely left to implementations).

It should be noted that conceptualising models of recording and reality in this way relies on accepting that recording the results of perception and prehension of things in the real world (including the state of our own minds) is not a null transform; on the contrary, we *choose*, *abstract*, *filter*, and *organise* the fruits of our investigations and thinking to suit our later use of the information.

This broad distinction of model types is useful because it enables the semantics of the object of any recording (some view of “reality”) and those of the recording / organising / persisting process itself to be distinguished, leading to a number of useful consequences. The first of these is that a standard model (or perhaps a small number of models) of recording can be developed, independent of models of such things as acts, workflow, and guidelines. Secondly, models of recording can be defined so as to accept as “content” data whose internal form is defined by the latter kinds of models. It is thus reasonable to try to devise a model of the EHR in the form of a “model of recording” which at some interior level includes the semantics of state, observation, actions, evaluation and planning. The general scheme is shown in FIGURE 1.



**FIGURE 1** Relationship between Models of Recording and Reality

The complexity of a model of recording depends on the requirements. The electronic health record poses some quite strong requirements: audit trailing, the need to be able to visit any previous state of the record, and the need to be able to organise information as required. Recording of other information, e.g. demographic data, has more modest needs, mainly just audit trailing.

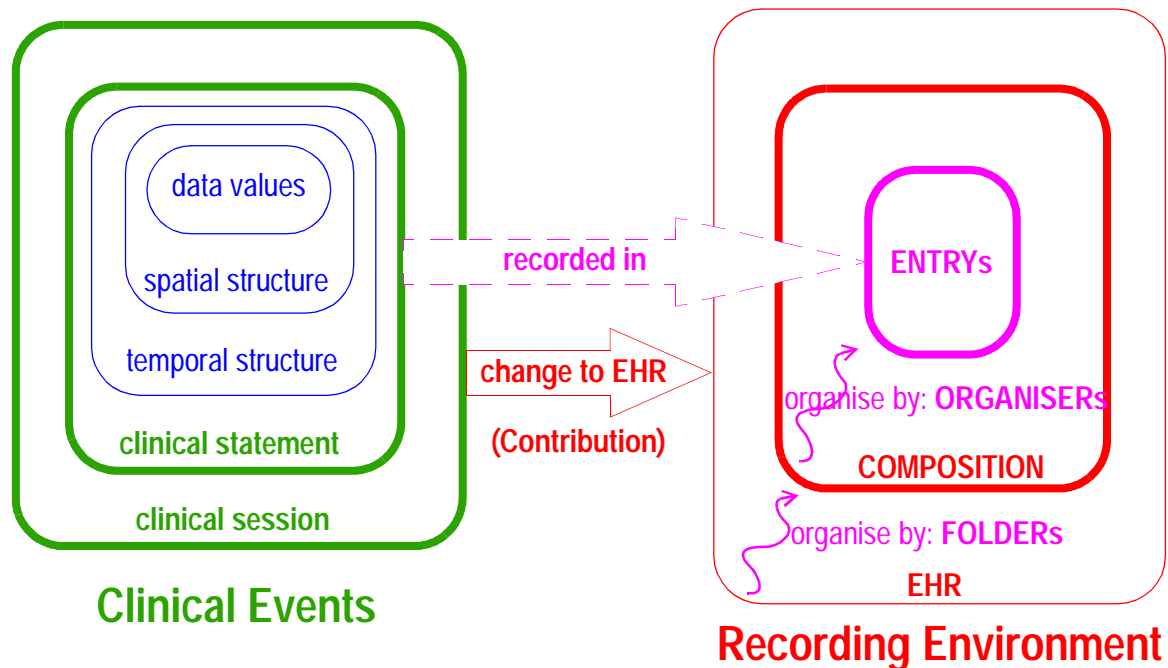
Models developed to date in electronic health records and documents, including CEN 13606-1, the GEHR and *openEHR* EHR reference models, and the HL7 CDA all have a similar structure, which in fact reflects a certain model of recording, although this has not generally been articulated. However, it is worth considering how the generic process of recording clinical information indeed leads to such a structure, and also what the further consequences are.

Consider FIGURE 2. On the left hand side is depicted the context of a “clinical session” in which “clinical statements” are made. A clinical session is defined as any business activity of the health care system for the patient, including encounters, pathology tests, and interventions. A clinical statement is anything the clinician wants to record. Clinical statements are shown in the diagram as having temporal and spatial structure as well as data values - i.e. a data structure of some kind.

On the right hand side of FIGURE 1, the EHR recording environment is represented. We can assume that in general that an EHR consists of separate identifiable parts (“Compositions” in CEN 13606-1), each of which contains one or more clinical statements, in the form of Entries. One of the activities during the recording process is that of organising things; accordingly, Compositions can be organised by Folders, and inside Compositions, Entries are organised by Organisers. Changes are made to the EHR in the form of “contributions”, as described earlier; here we see that contributions are associated with clinical sessions, and that each contribution causes clinical statements to be encoded as Entries in Compositions.

As a consequence, the EHR has the following structure:

- Folders: to organise Compositions
- Compositions: distinct content containers
- Organisers: to organise Entries inside Compositions
- Entries: expressing the actual data being recorded



**FIGURE 2** General Model of Recording

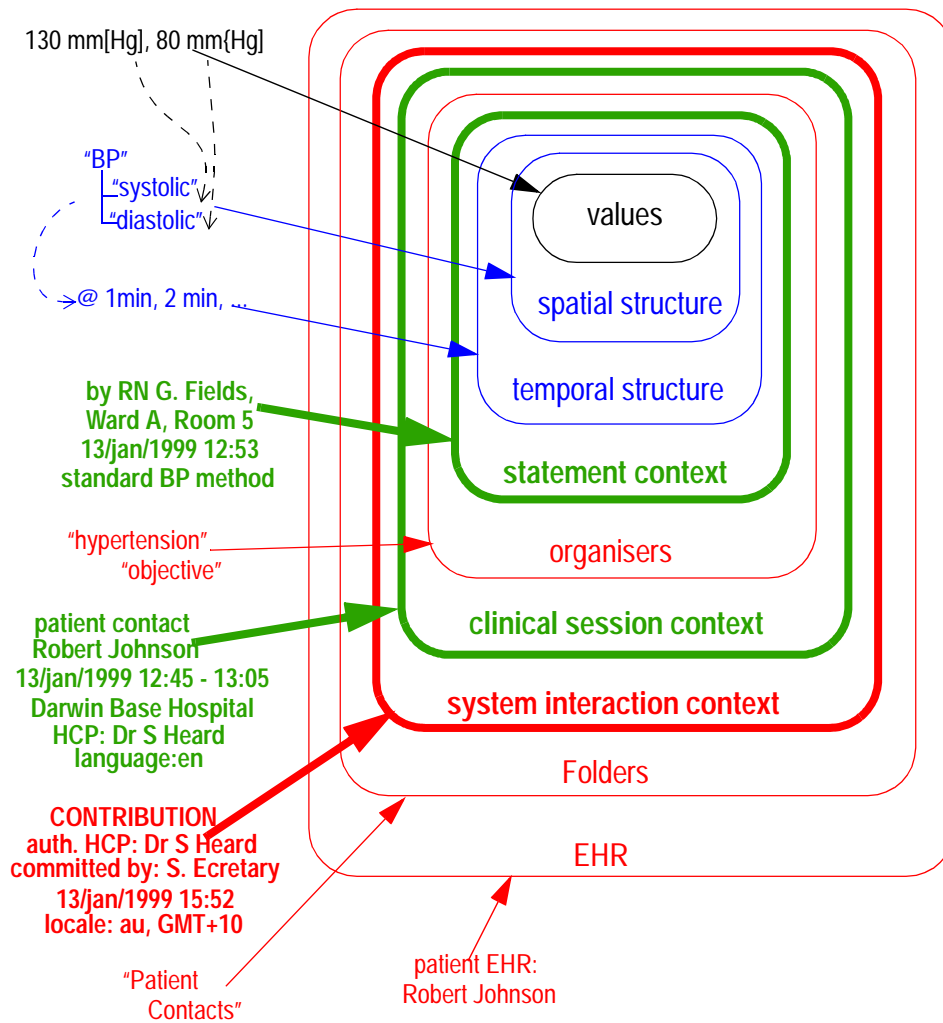
Further, each heavy ring in the diagram corresponds to a "context", for which "context data" (who/how/when/where/why) needs to be recorded. Context information is consequently captured at the following levels:

- the Contribution: attributes capture the context of the interaction of the user with the computer system.
- the Composition: attributes capture the context of a clinical session.
- the Entry: attributes capture the context of a clinical statement.

As a result a clear model for context attributes is possible, superior to the *ad hoc* attempts of the past. FIGURE 3 illustrates an example data item in the EHR including context attributes (in bold), original structure and organising information (organisers and folders). It should be noted that the notional author of EHR content depicted in FIGURE 4 need not be human, indeed, in the future it may well be a decision support application or other intelligent agent.

We can now relate this general model of recording information to models of *what* is recorded. In the CEN 13606 model, the fine structure is completely generic, being made of "Clusters" and "DataItems". In the *openEHR* EHR model, slightly more powerful generic structures are also available, with Entrys being subtyped into three species of information, based on the following general classification of Rector, Nowlan and Kay [32] (the terms used by the latter in parentheses):

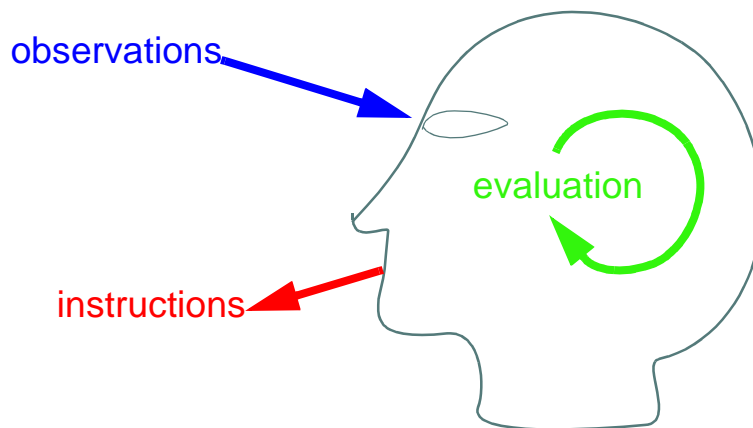
- *Observations (observation statements)*: what the patient says and what the clinician sees. This information is attributable to a source, permanent and factual but may include conflicting, uncertain or negative statements, which may be expressed at an arbitrary level of detail.
- *Evaluations (meta-observation statements)*: what the clinician thinks, including problems, plans, justifications, reasonings and rationales, analyses, groupings of problems and differential diagnoses.



**FIGURE 3** Context Information

- *Instructions (clinical dialogue)*: what now ought to happen, including requests and responses relating to investigations, referrals and treatments, the observations and recommendations of third parties.

These three information types are illustrated metaphorically in FIGURE 1. Essentially they constitute statements recording information coming in to the recorder, statements recording thoughts, and statements recording instructions to act.



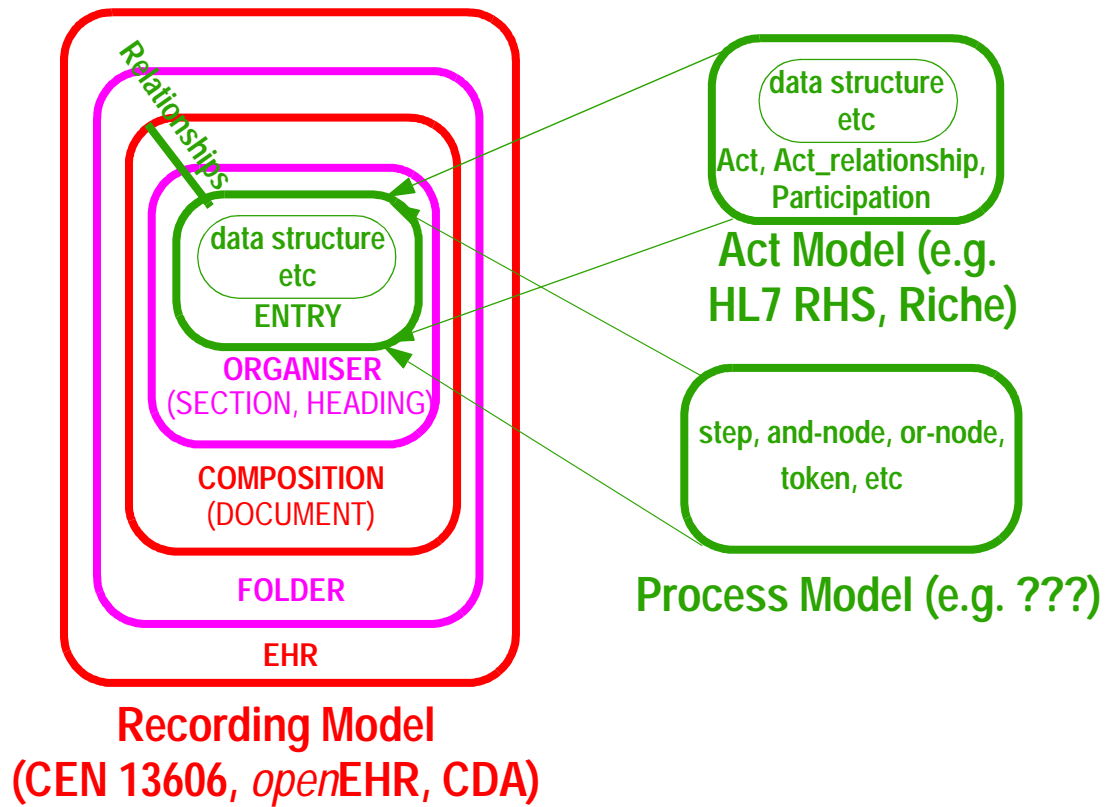
**FIGURE 4** Basis for Clinical Statement Classification

Beneath these general types are generic temporal and spatial data structures suitable for recording events in the past as well as proposed courses of action in the future. In the CEN, GEHR, *openEHR* and CDA models, it is possible to create links between Entries, enabling complex networks of related items to be created as necessary, each with their own internal complexity. In this way, a representation of real world events and relationships is possible. Examples include:

- causal chains: e.g. patient contact (Observations, Instruction for test), pathology test (Observation), patient contact (Evaluation, Plan, Instruction), therapy administration (Instruction...) and so on;
- views: constructed views of existing data for the purpose of presentation as a report or summary;
- thematic networks: e.g. linked networks of items relating to a particular health problem or issue.

We can now see that the relationship between a model of recording such as described above and models of what is being recorded is that the latter kind of models correspond to the Entry level. That is to say, any data recorded which is itself described by a model such as the HL7 RIM, decision data derived from guideline evaluations, data describing processes (e.g. as described in workflow models), should create new Entries in the EHR. This relationship is shown in FIGURE 5 (CEN and CDA names in parentheses on the left hand side), and allows us to understand how to use such models as EHR and models of things in the real world. For example, the HL7 CDA document model is an XML-based model of documenting clinical observations and events which corresponds closely to the CEN 13606 and *openEHR* models up to the level of the Composition (CDA Document). CDA content - i.e. its Entries may come from *ad hoc* sources (e.g. existing clinical documents) or model-driven sources, e.g. messages based on the HL7 RIM.

Further, we can use the principles of two-level modelling described earlier to see how to relate incoming sources of data to Entries in a way that leads to sustainable software and information systems. Consider an HL7 message for a biochemistry test result which is received by an EHR system. It will create one or more Observation Entries, however, we can intelligently



**FIGURE 5** EHR Recording and Models of Reality

manage the information flow in advance by devising domain level mappings between the HL7 message structure, defined by a restricted message information model (R-MIM), and archetypes constraining the valid structure of Entries in the EHR.