Roy Rada, Angelo Rossi-Mori, Ray Paton, Alan Rector, Fabio Magliani, Pieter de Vries Robbe, and the rest of the GALEN team

The GALEN Dream

Rada, R., Rossi-Mori, A., Paton, R., Rector, A., Magliani, F., Vries Robbe, P.de et al: **The GALEN dream.**
Int.Classif. 19(1992)No.4, p.188-191
Outlines the origin, needs and principles of GALEN, the Generalized Architecture for Languages, Encyclopedias, and Nomenclatures as applicable to Medicine. Short-term and long-term plans of GALEN have been elaborated to cope with possible developments. "Milestones" are given indicating what should be reached when and how much funding will be required for each milestone. In two "vision" pictures the situation before and after the introduction of GALEN is shown and the responsibilities at 4 different levels are listed. (I.C.)

1. Introduction

Under the auspices of the Advanced Informatics in Medicine initiative the Commission of the European Community (CEC) solicited proposals from consortia in the fall of 1991. After evaluating the proposals, the CEC decided to support several large projects to advance informatics in health care across Europe. One of the flagship projects, to which the CEC contributes over 3 millions ECUs (the conventional European Currency Unit where one ECU has a value approximately equal to that of one American dollar), is called Generalized Architecture for Languages, Encyclopedias, and Nomenclatures (GALEN) in Medicine. The GALEN consortium includes leading institutions from across Europe, both members of the CEC and non-members:

University of Manchester, England
Hewlett-Packard Ltd, Europe
Hopital Cantonal Universitaire, Switzerland
Consiglio Nazionale Delle Ricerche, Italy
Conser Sistemi Avanzati SPA, Italy
Katholieke Universiteit Nijmegen, Netherlands
University of Linkoping, Sweden
Finnish Hospital League, Finland
GSF Forschungszentrum MEDIS Institute, Germany
University of Liverpool, England

These partners expect to produce by the year 1995 a knowledge base to support the development of medical languages and nomenclatures in Europe. GALEN has important parallels with the Unified Medical Language System project which began in the United States several years ago.

2. The Need for GALEN

Imagine the following scenario, in the not too distant future. A tourist is visiting a Mediterranean island and suddenly goes into shock. Is it the sun, is it the food, is it a violent crime? The local paramedic rushes to the aid of the victim and discovers in his wallet a Smart Card and inserts it into her portable palm-top, computer. The computer reads the patient record from the Smart Card and recognizes that the patient has an allergy. Going through the GALEN knowledge bases, the computer is directed to a Scandinavian database which contains a description of allergic substances and is connected to the relevant information. It proves that the wine in that area contains this allergic substance, and the computer recommends the appropriate treatment plan.

This anecdotal example demonstrates that there is a need to connect information in various forms and places. The different languages across Europe are one barrier to communication and the diversity of databases and knowledge bases is another. Some common meta-language or knowledge base could support connections across the various languages, databases, and knowledge bases.

3. Principles

Standards are often proposed as a solution to the problem of communication barriers. In the health care arena, the International Classification of Diseases (ICD) is one standard which facilitates communication about the mortality and morbidity in a population. The Systematized Nomenclature of Medicine (SNOMED) is another standard which pathologists may prefer over ICD to describe their problems. Medical publishing houses use other languages to describe the literature which they index. Each of these standards has its users and maintainers, but the standards themselves are disconnected and the cost of maintaining them is so great that they typically remain outdated.

GALEN intends to deal with standards in a novel way. Standards will remain a vital part of the health care communication scene, but the GALEN infrastructure will allow people to easily connect across standards and to develop new standards. This approach involves a fundamental paradigm shift for the handling of medical terminological standards. Concepts rather than terms will become the focus of attention in a GALEN world. The

computer will provide support for people to easily move between the conceptual and the terminological.

With a GALEN infrastructure an information system can support users who might otherwise need to manually encode their assessments so that these users can express themselves in ways more naturally than they can with simple terminology systems. GALEN will accept phrases from natural language and encode them into the appropriate terms from selected classifications.

The computer can generate information to cover a wide range of needs for terminology. In such a situation less information needs to be handled by people and the demands on people to remember or generate information are appropriately minimized.

4. Plans

The short-term and long-term plans of GALEN have been elaborated. These plans, of course, can only be realized when the appropriate resources are at hand for implementing them, but the GALEN plans include the acquisition of additional funding.

In the first 3 years, GALEN realistically targets medical coding centers as its customers (see Figure 1: '10 Year Plan'). These centers could use knowledge-based tools to facilitate their work of developing and maintaining classifications, and GALEN will provide such tools. Coding centers are national or international institutions which are responsible for the maintenance and distribution of medical languages and nomenclatures. For instance, the British have a Coding and Classification Center for health care codes which is responsible for maintaining the official, medical codes, called the Read Codes. The Dutch have their WCC, the national centre for terminology and classifications for health. These centres are in close contact with representatives of the GALEN project. In order to effectively liaise with the various coding centers around Europe, members of the GALEN consortium are soliciting additional funding. The expectation is that with a budget about twice that of the current 3 million dollars of GALEN that the European coding centers could be effectively served by the tools of the GALEN project.

Beyond the 3 years 1992-1995 the GALEN team expects to pursue both in-depth medical knowledge base construction and broad application to areas outside medicine. For the medical side the 6 year plan is to have demonstrated the value of integrated knowledge bases in Europe. Prototypes will have been developed that exploit the GALEN knowledge infrastructure and allow integration of different knowledge-bases. By the beginning of the next millenium, 10 years forward from the start of GALEN, the expectation is that some institutions will be responsible for the day-to-day maintenance of the core knowledge of medicine. Given the stability of this core, wide commercial and public-sector development of different domain-specific, health care knowledge bases will be practical — they will all communicate with one another through the core. To achieve this 10 year goal, an investment of about 30 million ECUs will be required.

The parallel stream that develops after the 3rd year involves the exporting of the GALEN paradigm to non-medical arenas. The principle of a knowledge infrastructure which can be gracefully augmented and which supports the interconnection of any other knowledge bases is a powerful one that applies to law, telecommunications, manufacturing, and virtually any other discipline. The importance to Europe of integration across legal, telecommunication, manufacturing and other boundaries is so great that numerous, special initiatives address these needs. The GALEN goal after year 3 is to encourage other initiatives, such as the 'European Nervous System' initiative and the ESPRIT initiative, to fund work throughout Europe that would continue the efforts begun in GALEN.

Ten years after GALEN, well into the first steps of the new millenium, the fruits of GALEN will begin to feed yet more ambitious projects (see Figure 1: '10 Year Plan'). Natural language processing is a task performed well by intelligent agents and intelligent health care computing systems will need this facility. Medical problem solving, whether it be for patient diagnosis, intensive care monitoring, or resource allocation, also will be an area to be addressed.

Medical Plans		
Years after GALEN	Milestone	Funding Needed (times GALEN budget)
3	coding centers using GALEN results	two times
6	prototype of integrated knowledge bases	five times
10	establish institutions to maintain the CORE	ten times
10+	natural language processing, medical problem solving	unknown

Fig.1: '10 Year Plan'. The table shows for various years from the beginning of GALEN which milestones should be reached and how much funding will have been required to reach these milestones.

5. Visualizing the Dream

GALEN is the prototype of a mechanism for a "distributed" standard on medical terminology. As such GALEN together with the Medical Informatics Committee of the Committee of European Normalisation (CEN) will produce the rules and the knowledge to allow everybody to make independent contributions which can then be integrated. If successful, it will facilitate the integration of independent coding systems and user-defined controlled vocabularies among themselves as well as inside other computer-based modules (such as KBSs, medical record and bibliographic databases). In order to illustrate this point, consider the following metaphor.

The current situation in advanced medical informatics may be visualized as a scene in which there are a lot of islands in a terminology bay (see Figure 2: 'Terminology Bay - Before GALEN'). The islands represent current information systems in medical informatics such as, MEDLINE island, Excerpta Medica island and Medical Record islands. We may also recognize various KBS

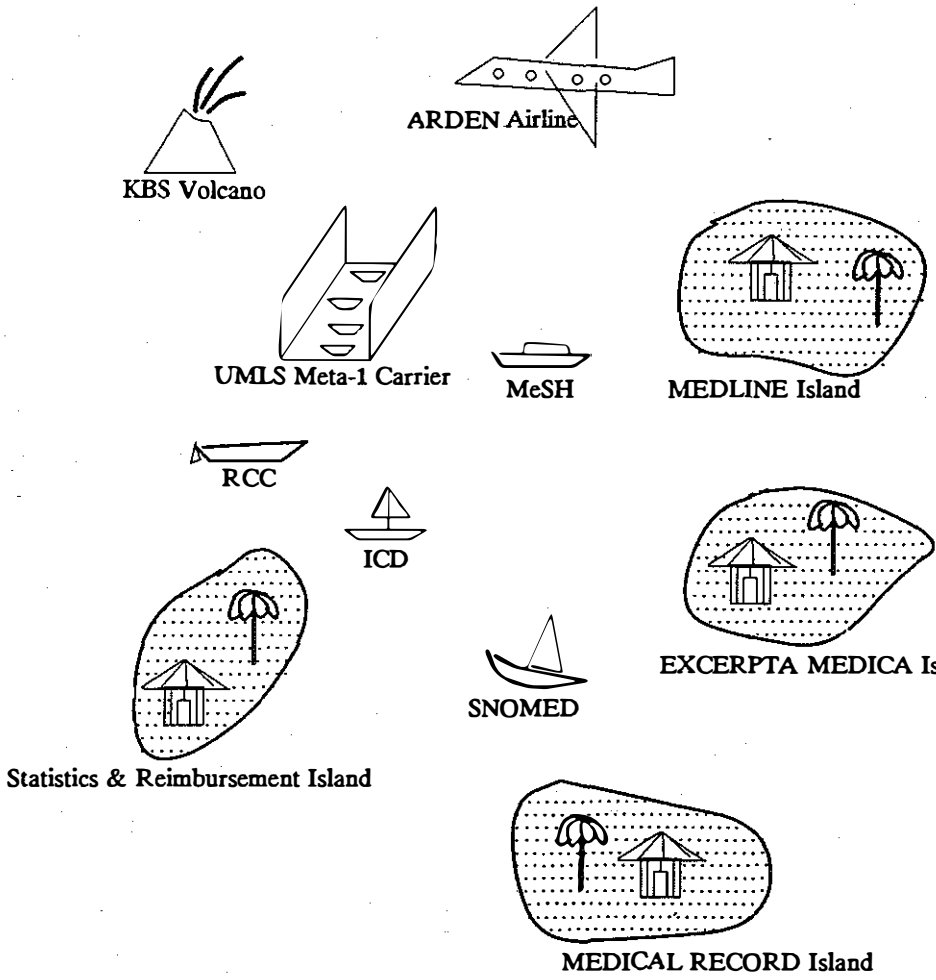

Fig.2: 'Terminology Bay - Before GALEN'. This visualization of the world situation vis-a-vis medical terminology before GALEN shows the lack of connectivity among the various languages and information resources.

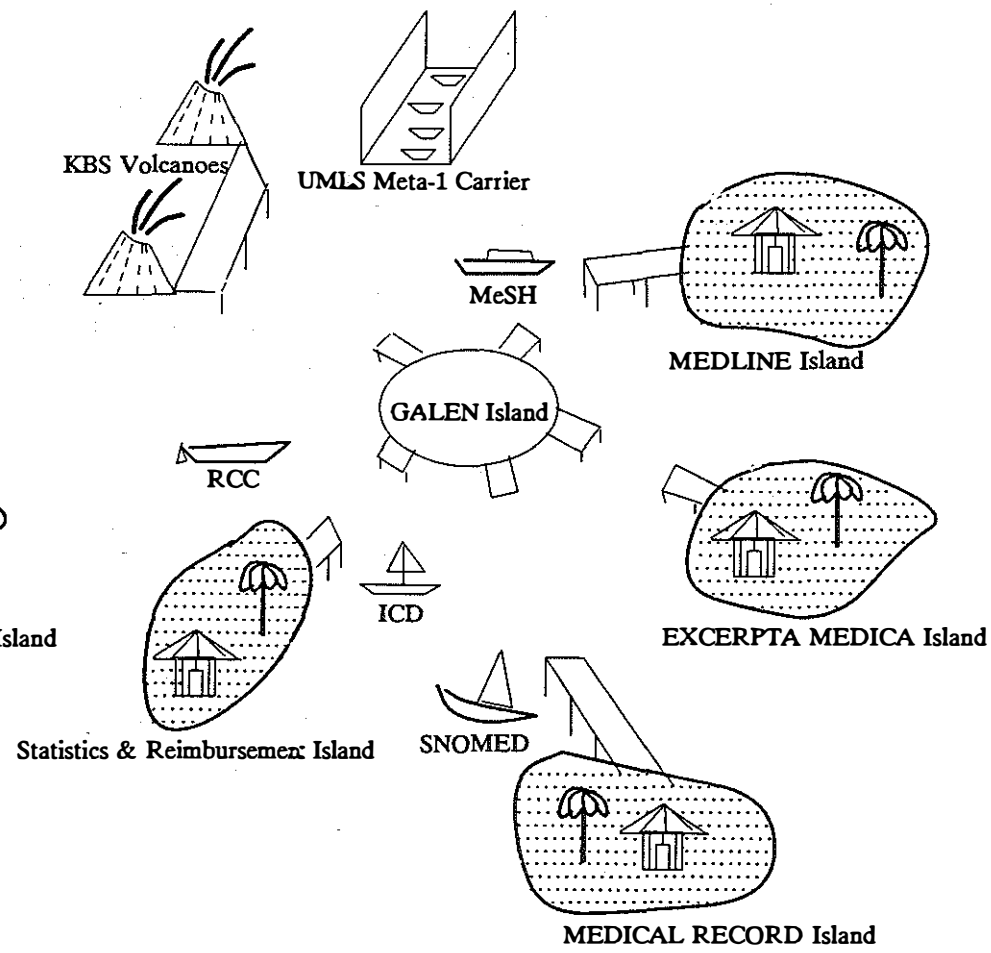

Fig.3: 'Terminology Bay - After GALEN'. Five years after GALEN has been working the long-distance crossings which have been required for traveling in the world's medical terminology scene can be significantly shortened. Ships will cross from harbors on the application islands to harbors on the GALEN artificial island, and on the GALEN island the goods will be automatically transported to another GALEN harbor and then shipped to their destination island.

"volcanos". Is it possible to communicate between the islands? At the moment this is not very easy. The best available method is by pirogue - a kind of single person boat - not a very effective means of communication in the big and dangerous bay. This kind of transportation is typified by single vessels such as ICD and the Medical Subject Headings (MeSH) or by convoys of vessels representing SNOMED or the Read Clinical Classification (RCC). There are also pirogue carriers such as the Meta-1 System from the Unified Medical Language System (UMLS) that allows a safe transportation among islands. Although ARDEN airlines provides some flights (note: ARDEN is a knowledge representation language), landing facilities are poor.

From this we may draw two analogies: How easy will it be to get lost in Terminology Bay? For many people the answer must be "very easy". There are vast amounts of jargon, some of which is synonymous, some may even be redundant. Why is it the case that current means of communication are so poor? Mainly because the terminology bay keeps getting bigger and yet there is no overview that permits the development of a common language or means by which one language may be translated into another.

We can now begin to see how the GALEN vision fits into place. The first version of GALEN may be considered as a floating platform in the middle of Terminology Bay with a set of pontoon bridges which connect to the islands. It is now possible for a great deal more transfer to take place. This floating platform is but a first pioneering attempt to integrate the various geographical locations and soon after an artificial island (see Figure 3: 'Terminology Bay - After GALEN') is constructed, well anchored to the bottom of the bay. Now more solid and well-engineered bridges can be constructed according to CEN rules, the know-how and the engineers to build other bridges and enlarge the artificial island. With this accumulated knowledge, GALEN will then be able to assist other investors in building a set of specialized artificial islands, including a harbour for conventional ships and train stations.

The construction technology will spread. Use of the GALEN technology should ensure that everybody will be able to take a boat from his island to the nearest GALEN harbor. GALEN will manage the transportation on this central, artificial island and together with other investors will ensure the operability of other specialized artificial islands. The rest will be up to the public.

In view of the previous illustration, we may now reconsider the intended purpose of GALEN. At present, it is not possible to satisfy the needs of individual users with a unique, combinatorial, centralized system. We need at least four levels of responsibility that allow exchange between levels:

Level 1

the central nucleus (produced by the GALEN consortium in three years);

Level 2

speciality modules (produced by joint ventures between the GALEN consortium and third parties, after the third year);

Level 3

application modules (produced by software houses and their expert clients on specific applications, according to rules that CEN and GALEN will provide during the third year);

Level 4

user-defined adaptations (made by unsatisfied users on certain items, according to local jargon). They are stored and perhaps transmitted (if really needed) according to CEN-GALEN rules.

There will be a market for Level 3 modules, namely, the modules will be sold at market prices. On the other hand, Level 1 and 2 products are provided to customers at negligible cost.

At this point it is worth noting that GALEN has responsibility for the correctness of Levels 1 and 2 but not for 3 and 4.

6. Conclusion

This brief document has presented the principles and plans of GALEN, as foreseen not just for the next 3 years but also for the longer-term future. GALEN will provide a much-needed conceptual infrastructure to the language and nomenclature needs of health care. Furthermore, the power of the GALEN intentions has been appreciated through a visual metaphor. This power derives from the needs for GALEN which can be seen across the health care field and other fields, as well.

Address: Prof. Roy Rada, M.D., Ph.D., University of Liverpool, Department of Computer Science, Liverpool L69 3BX, England.

Call for Papers: Knowledge Organization - Problems and Trends, Moscow, May 1993

A conference on this topic will be organized by the Russian Chapter of ISKO to take place in Moscow, May 10-14, 1993. Monday, May 10 will be registration day, Friday, May 14, departure day.

The conference organizer calls for a contribution of papers also from abroad concentrating on problems and aspects related to ISKO's areas of interest.

Contributors are requested to submit an abstract of 1-2 pages by January 1, 1993 to the address given below; name, occupation, organization, phone, fax, telex, country, correspondence address should be added.

Working languages to be used during the conference will be English and Russian. Papers should be presented in either one of these languages. Presentation time is 20 minutes, the final text should not exceed 8 pages.

Please contact: Dr.Eduard Sukiasyan, Russian State Library, Ul.Vozdvizenka, 3, Moskva 101000, Russia. Fax 7095-200 2255; Telex 411167 gbl su.