FIVE LAWS OF INFORMATION SCIENCE AS BASIS FOR ARCHITECTING KNOWLEDGE INFRASTRUCTURE IN EDUCATION AND ENTERPRISE

K.R. Srivathsan

ABSTRACT

Information Systems of tomorrow need to integrate seamlessly the organization of different kinds of information for the users. One part of it will be the Digital Library that organizes the classification, acquisition, management and archival of generic library objects for the users. On the other part, institutions and organizations are increasingly deploying Enterprise Integration Portals accessible over the Web. Such portals provide for opportunities where we may present the Digital Library with component systems that manage the dynamic and strategic information to provide knowledge services of value to the customers of the enterprise. Thirdly, we also need to facilitate knowledge capture of internal processes of the enterprise through which decisions are facilitated. In this paper the case is made that the three together are best integrated into one comprehensive Enterprise Applications Integration service for any knowledge-driven organization. This framework is also shown to be in conformity with the five laws of information science that were enunciated by S.R. Ranganthan.

Such integrated portals are shown to be capable of supporting Virtual Knowledge-Driven Enterprises. The case studies of the KISSAN-Kerala and Education Grid portals are presented as examples of information systems capable of supporting such Virtual Enterprises. Based upon this experience, the case is also made for the development of Scientific Portals as extension of the principles outlined in this paper.

Key Words: Five laws of Library Science, Metadata and Classification, Digital Library, Knowledge Management, Enterprise Applications Integration, Virtual Enterprises.

1. COMPLEXITY OF EMERGING INFORMATION SHARING ENVIRONMENT

Traditionally libraries and publishing industry handled the acquisition and dissemination of generic knowledge in the form of books, journals, references, or, literature. These are more in the nature of study and reference materials. With the Internet and electronic form of storing and distributing knowledge, libraries have begun offering some of the generic body of knowledge in the form of objects databases. New information suppliers have emerged like the database designers, aggregators and distributors. Journals and databases in diverse fields such as commerce, economics, agriculture, statistical data, health information and multimedia content are increasingly available in electronic form. This is a challenging and complex new development in several ways for the library community for the several reasons as cited below.

a) IT today is evolving at a rapid pace in both technology and management sense. Issues like security, privacy, fair use, intellectual property issues like copyrights and ownership, virtual organizations management, etc. remain open.

b) There are emerging needs to support virtual knowledge-driven enterprises where members from different organizations pool together the requisite information and competencies to offer services of value over the web to users. These call for appropriate methodologies to manage information resources generation and utilization. As explained later, we may call them as virtual enterprises or thematic communities.

c) The users who are engaged in knowledge intensive work ultimately gain value from information provided it is the *right information* available at the *right time, in the right place, in the right context* and to the *right persons.* Hence, effective value proposition goes with efficient organization of knowledge related activities like interaction, group collaboration, messaging services, etc.

d) At the technology level, the Internet and IT world is swiftly changing over from client – server interactions to seamless peer-to-peer interactions over web accessed information and interaction services. Users may access information from anywhere independent of the client systems they use.

e) With ever increasing proportion of information in complex multimedia and digital form, there are immense challenges ahead concerning archival and retrieval of the same. With sophisticated technology needed for access and playback of the same, the world is yet to intelligently grapple with preservation of human heritage over the coming decades. This calls for reinvention of computer systems, information presentation devices and modes of knowledge creation and utilization. To a good extent, technology in IT should evolve reasonably independent of the modalities of knowledge creation and its utilization in the society.

In the above context, we have to consider not only the information sharing aspects per se, but need to look at the dynamics of how this information is put to use by its users and the society at large. Here it will be useful to identify the different ways of how we use information in the real world when we are engaged in knowledge driven activities. This aspect is developed in the next section.

The Economist Laski stated that the cost of money is that much we forego by not using it in the best possible alternative ways. If we change the word money to information, this statement is as much valid in the context of knowledge enabled wealth creation. Here we use the term information as that, which, when received by a person or a system, either changes, or has the potential to change the state of the person or the system. Hence wealth creation, or, development in the coming knowledge era is about how effectively the members of working groups or organizations use such opportunities provided by information to create value. It is in this context we appreciate Peter Drucker when he writes, "Knowledge has become the key economic resource, and the dominant – and perhaps even the only source of competitive advantage".

What we address in this paper is to look into some key principles and organizational aspects of information systems needed to build efficient knowledge management systems that assist groups engaged in knowledge intensive services to create wealth or value.

2. FIVE LAWS OF INFORMATION SCIENCES AND INFORMATION DYNAMICS

S.R. Ranganathan stated the classic five laws of Library Science [1] as the guiding spirit behind architecting and managing the libraries as knowledge supermarkets. The same may be rephrased as given below with somewhat different relative emphasis to guide us in architecting and managing emerging information systems of the 21st Century.

a) Information is for use.

b) To every user, the information he/she seeks.

c) Information of value to go to the user who may benefit from it.

d) Save the time of the user.

e) The information system with its associated sources of information and the beneficiaries of its services form one integral organism.

Like the way Ranganathan uses his five principles to drive the classification and management of printed information and libraries, we shall attempt to arrive at frameworks and structures that will help us build future information systems. There are fundamental changes in the way information is generated and used in this networked era. Some of the conspicuous ones are:

(i) In volume, information is becoming explosive. Presentation and management of information therefore needs to be specific to the needs of the user or group that uses it.

(ii) The dynamics of generation and usage of information has increased tremendously. The original objective of library as knowledge aggregator and disseminator needs to be enhanced to include links between the way users' or groups' workflow needs and the knowledge base they need to refer to during the course of the work. Understanding such information workflow interactions over applicable organizational behavior framework will lead to the right type of information dynamics that has to be built into the information systems.

(iii) Users need protection from the information flood and glut. We need systems that help them stay focused while working in knowledge intensive activities.

(iv) Unlike the traditional long-life library institutions, we may set up ad hoc information systems and close them on need basis. As an example, an information system for disaster management may be set up at short notice for a target focused group, say for a particular disaster mitigation task, and afterwards wound up with due archival of knowledge generated and audit report of its activities.

A great deal of work has been done and happening in Data Warehousing, Artificial Intelligence (AI) based search/access techniques and Software Agents. The popular Google and other search engines are fine examples of such AI tools. They, together with Net Crawlers, Semantic Web and 'Know-bots' (Knowledge Robots) help in building smart query and attributes-driven filtered view of the open global information space accessible over the web. In a limited way they provide some features for personalized systems for custom-view and access of the open information. These add efficiency by reducing the information glut to more manageable volumes of information that we seek and work with. However, they are mostly passive, stochastic in their behaviour and do not address issues of efficient information generation and utilization.

So we build compound information systems by adding component systems of value like search engines, digital libraries, document and workflow management systems, e-publications and distribution, Groupware and collaboration tools, etc. Portals for Enterprise Knowledge Management Systems, or Enterprise Applications Integration are good examples of such systems. However, the present approach to build such enterprise systems is mostly ad hoc and shaped by contingent needs of the concerned organization. In this paper, we highlight certain core principles of knowledge management that help us understand the interplay between information systems organizations on the one hand, and the need to support the dynamics of information and interaction that result in effective ways of knowledge based wealth creation.

2.1. INFORMATION INFRASTRUTURE FOR VIRTUAL KNOWLEDGE ENTERPRISES

Given the scenario of emerging web-accessed information world, we address the kind of informatics framework of architecting and managing information systems of the future. Today in most portals the information management and access facilities are largely dictated by local expediencies and internal workflow needs. There appears to be ever increasing need to adopt open architectures that are driven by data/metadata standards and protocols of information exchanges. A brief review and requirements towards standards of information organization is given here. Firstly, let us look at the need for different types of classification and associated data structures. There are the following four broad functions in information classification.

i) Organization of generic body of knowledge: Digital Libraries are best organized and managed to support this organization. The Dublin Core metadata standard is currently popular to organize, manage and access digital libraries. Dublin Core is a 'scalable' metadata standard in that we may expand the metadata tree to include pointers to small units of information with subtopics. There are a few other metadata standards in broad subject areas. For example, ACM recommends a classification scheme for Computer Science related subjects.

ii) Subject specific Knowledge Maps: Every subject area comes with its own technical classification, or, taxonomies. Examples are industry specific classification of records, classifications in Botany, Zoology, etc. We need to expand such standards to include subjects like Law, documents classification in E-Governance, Economics and Statistical data in different areas, etc. Each classification system has to be supported by corresponding open generic information systems and made available widely for data exchanges over open standards. This will help drive virtual knowledge enterprises with globally distributed members from different organizations. The classification will serve best if it closely corresponds with the kind of knowledge questions users are likely to ask frequently in the subject area.

iii) Workgroup specific documents generation and publishing: Every activity today is increasingly driven by knowledge. The engine for driving this knowledge space is with workgroups that access, generate and publish information relevant to their function or business. Members of any workgroup or strategy group need tools for regulated generation of information. Document versioning and management tools are one such example. Publishing and Groupware software tools support versioning. They also need support for workflow processes to monitor, control and release of information publications. These are increasingly becoming available as part of Enterprise Applications Integration and Knowledge Management and getting adopted in increasing number of organizations.

iv) Capturing information regarding events and related knowledge generation: When information systems are juxtaposed and used with enterprise systems, we get new facilities that are not available hitherto. Providing alerts to relevant real world events enhances the value of Enterprise Systems. Events based alerts together with the group-specific knowledge base help in effective strategic decision support and crisis management in any organization. Logs of such alerts, associated events and resultant actions provide useful inputs for knowledge capture and building the culture of a learning organization in the enterprise itself. Alerts may be generated by pre-specified conditions or through smart agents. The organization of necessary data structures for events capturing and alerts is not that well studied in traditional information sciences. One subject area where it is well understood is Network Management. For example the metadata structures like Abstract Syntax Notation –1, (ASN-1) [2] for building records of events and actions are used to form the Management Information Base in networks. Similar structures and methods for events capture are also of value in enterprise systems.

The organization and architecture of information systems are influenced very much by the way we design the metadata frameworks for each of the above four component areas. How they influence the architecture may be understood better if we visualize that the environment of any knowledge worker as having the four component areas as depicted in Fig. 1.

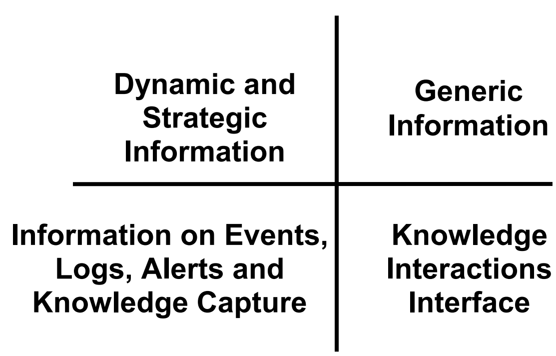

Fig. 1: Information Systems Components for Workgroups

In Fig. 1, organization of generic information pertains to (i) above. Traditionally, this has been well exposited under the metadata organization as in Library Sciences and classification principles like Dublin Core, or Ranganathan's facet classification. However the dynamic and strategic information organization in the second quadrant of Fig. 1 needs considerable study and global efforts in standardization. To a considerable extent it is conditioned by the kind of knowledge questions the system has to ask in applying the information in the corresponding knowledge-driven enterprises.

As a strategy, we need to look beyond information generation, aggregation and dissemination functions. We have to ask how the information systems will result in value addition to the users when they are juxtaposed with their work environment. Such integrated information systems combined with workflow processes of any group engaged in knowledge-based services may be modeled as shown in Fig. 2. The assumption behind the model is that every knowledge intensive service is driven by a close-knit group of users. They need their associated or custom knowledge base for reference and study purposes while they are engaged in their work. An earlier paper by the author [3] made the case that in a Knowledge Society the 'Knowledge Citizen' (or K-Citizen as referred to) is the focused group of knowledge workers that drive specific services of value. Such a view helps in architecting and managing the knowledge driven information infrastructure of the kind shown in Fig. 2. It shows the typical knowledge interactions environment that every K-Citizen, or Knowledge Services Group while doing a variety of tasks. The environment may be referred to as Virtual Knowledge Driven Enterprise, or simply as Virtual Enterprise [VE]. The community or group that drives it may be called as the thematic community.

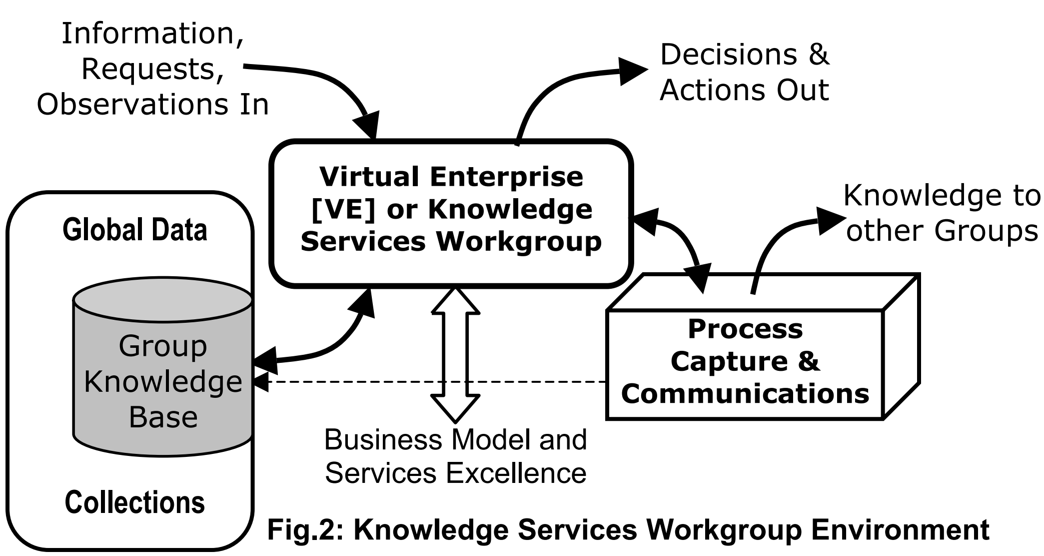

Fig.2: Knowledge Services Workgroup Environment

Every VE has the following attributes.

a) The members of the VE may belong to different organizations or different groups within a large organization.

b) The VE gets inputs in the form of requests, observations, alerts and information that requires service or attention. The VE uses its own business workflow. As it executes its workflow, it generates reports, alerts and information for other groups or thematic communities.

c) Every VE needs its own focused knowledge base partly consisting of a filtered view of the global data and information collections of relevance for the business or services that are rendered to its customers. It will also have the group's own proprietary base that is needed to drive its core competence functions. We call it the Group Knowledge Base.

d) Such Group Knowledge Base has both generic (relatively static, validated) and dynamic information components. The generic part is characterized by the capacity to manage as per traditional library practices that include the subscribed digital information components.

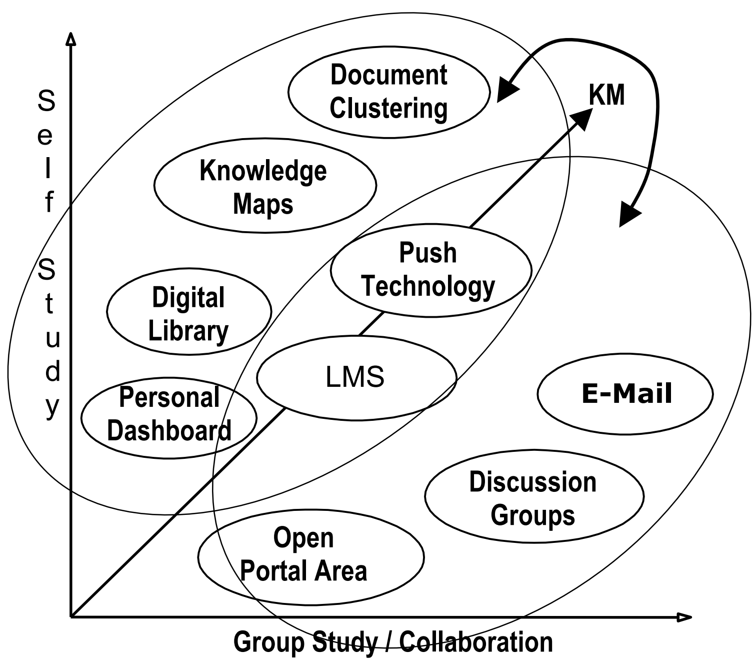

Fig. 3: Personal and Group Interface for Knowledge Interactions

e) The dynamic part of information the GKB pertains to current active data or information that is strategic to the VE's business or services delivery. For example, in agriculture this may be information pertaining to weather, market and farm inputs. Future information systems need to co-link dynamic information that is strategic to the related generic information space. This will add much value like the Decision Support System for the work(s) the VE is engaged in.

f) During the course of VE servicing its customers or engaged in its own internal activities, they create information by way of reports, logs of decisions taken, events, and alerts. It is important to have coherent reports and analysis of the logs of such information. Such reports serve in two ways. First is some of it become knowledge inputs to other VEs or thematic communities. Second is that it serves to direct the organizational learning for the VE's own and its stakeholders' needs.

g) Lastly, the complex mix of such functions as given above needs effective group-oriented user interfaces. We call it as the Knowledge Interaction Interface as stated in [3].

The above mix of information and interaction functions for the VE architecture of Fig. 2 is shown in Fig. 3. Every VE needs systems that support the relatively slow changing generic knowledge base as supported in conventional libraries, except that this may be the relevant part of the Digital Library of interest for the group. In real-life real-time working environment, the same group has to constantly deal with faster changing dynamic and strategic information. For example, a group engaged in agriculture commodity trade will need relevant market information, crop predictions, implications of weather on productivity and prices, and futures trading. These are both strategic and dynamic. Every group arrives at decisions based upon information inputs, requests, and observations through its own workflow. They need to capture the workflow process in arriving at decisions in the form of events log, alerts and intermediate steps. Such captured process builds accountability and helps the group learn from its past. It assists in building the organizational learning culture.

Each group or VE needs a Knowledge Interaction Interface [KII] as part of some Enterprise Integration Portal to deal with the information space outlined in Fig. 1. In [3] a generic structure for a typical KII was developed. The same is reproduced in Fig. 3. The explanations for the components of KII in Fig. 3 are given in [3]. The Indian Institute of Information Technology and Management – Kerala has developed a class of Enterprise Portals with capacity to support the kind of Knowledge Interaction Interface illustrated in the figure. The group that created it has now become a company – Transversal E Networks (see www.transversalnet.com) that further develops and markets these portals. They are now deployed in several leading organizations.

In the next section, we give sample case studies of building the above model of knowledge infrastructure for typical thematic community or VE model is implemented. These are based on the KISSAN (Kissan Information Systems, Services and Networking)-Kerala and the Kerala Education Grid projects.

3. CASE STUDIES OF KISSAN-KERALA AND EDUCATION GRID PROJECTS

The KISSAN-Kerala Project aims to provide the 'Right Information at the Right Time to the Right Persons (the farmers, and persons in agriculture related institutions, industry and trade) in their Right Places and in the Right Context." The project targets the entire State of Kerala. As explained in [4], the methodology is based upon the following four component activity areas.

i) A central interactive Kerala Agriculture Portal (www.kissankerala.net) that serves as the meeting ground of institutions, programs, farmers, traders, scientists and relevant government officials. The portal, besides being a source of growing and rich quality content is becoming a way to disseminate locale-specific information and advisories, market information and opportunities, etc., it serves thematic communities (or VEs) in different strategic studies and advisory areas.

ii) A popular weekly TV serial Krishideepam in the local languages over a commercial TV channel with wide reach. It attracts more a large number of viewers. The KISSAN Project management gets a large number of queries after each episode. The episodes deal with validated agriculture information of current interest, success stories of individual farmer's initiatives and interests, opportunities awareness, and similar topics that address farmers' concerns.

iii) Host an Agriculture Call Centre that handles queries from farmers and others coming over the telephone, portal or email. Useful questions are aggregated and made accessible over the portal under the Frequently Asked Questions section.

iv) The project will soon be launching the Virtual University for Agriculture Trade (VUAT) that will build awareness among the farmers, agriculture based industry, traders, and farm inputs suppliers. The educational focus will be on concerns of productivity and profitability for the state's largely plantation crops dependent agriculture. The objective is to enhance the sustainability and profitability of the state's agriculture in the light of the emerging globally competitive environment.

Based upon the experience in developing the technologies and running the programs related to the KISSAN-Kerala programs, the need for building a ‘Scientific Portal for Agriculture’ has been identified. This is briefly stated in the following subsection.

3.1. SCIENTIFIC PORTAL FOR AGRICULTURE

A major problem in giving knowledge empowerment to any farmer is that he needs timely and correct information of relevance to his farm related needs that is validated and interpreted by associated experts. Such information may come from many sources on a continuing basis. The experts should be equipped to give guidance for right decisions and to take the right steps that enhances his sustainability, productivity and marketing concerns of the farmer, thereby leading to better employment security. Further, with the WTO regime, bio-safety and environmental concerns, modern farmers need larger awareness of global issues of concern that will have deep impact on his trade security and sustainability. Here we propose the concept and architecture of a Scientific Portal for Agriculture to provide such decision support environment for the farmers. This is being built by the KISSAN-Kerala research team in IIITM-K. This portal is designed to provide both farm or farmer specific information and need-based educational and enterprise applications services.

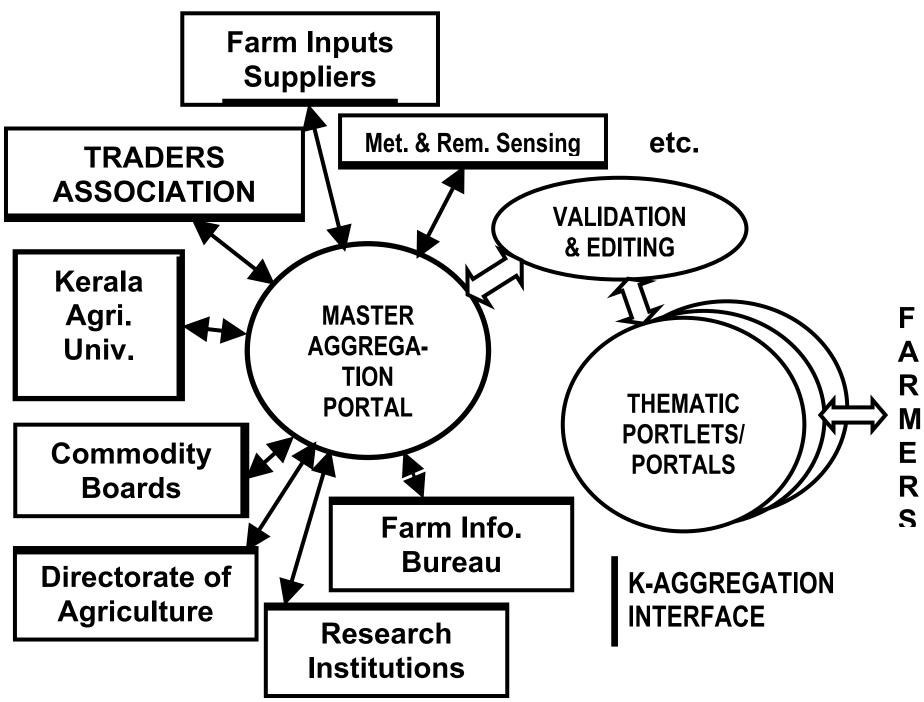

Fig. 4.: Information Aggregation & Flow Management In Scientific Portal For Agriculture

The proposed architectural view of the Scientific Portal is shown in Fig. 4. An objective of the Scientific Portal is to present to each user (e.g., a farmer) such information and knowledge of relevance and interest to him in the context of his farm and farming needs. Such total information has several components each of which may come from different organizations or sources. The architecture is scalable and allows for any number of farmers and other interested stakeholders to build their own knowledge interaction spaces as illustrated in the VE model of Fig. 2. There are several technology and standards related issues that need study while developing the VE information environment. We need development of XML schemas and profiles like the E-Portfolio definition for farms and farmers. We also need VEs that support expert advice and interpretation of dynamic information of value to farmers. These will be addressed as and when relevant VE services get configured and commissioned.

A multi-institutional project to develop and commission the Scientific Portal for Agriculture with capacity to configure and service large number of VEs is being proposed. The project is planned to be coordinated by the IIITM-K with the Kerala Agricultural University, and the Computer Science Department of the Kerala University. The Documentation Research and Training Centre of the Indian Statistical Institute at Bangalore is being consulted for the design of effective schemas and metadata frameworks needed for driving the dynamically configurable portal functions.

3.2. MULTIPLE COMMUNICATION APPROACHES FOR KISSAN-KERALA SERVICES

If we carefully study the five laws of information sciences as stated in Section 2 earlier, we understand that any successful IT enabled service, say for farmers, must have a complete feedback dynamical information system. The KISSAN-Kerala project has the capacity to support a fairly comprehensive feedback dynamical information system for the farmers through the four types of systems as stated at the beginning of Section 3.

The four – Portal based services, TV Serial, Call Centre and Virtual University – together constitute a comprehensive and effective delivery of timely information and empowerment services of significant value to the farmers of Kerala. This is born out by the considerable enthusiasm for the projects' services and almost 10 million viewers for the TV serial. Further, the project services a large number of queries (about 20 to 30 per day) for a variety of services and soliciting information of diverse kinds. Experts from a variety of organizations (as indicated in Fig. 4) provide the services for the different services.

3.3. THE KERALA EDUCATION GRID PROJECT

In several ways the Education Grid project complements the KISSAN-Kerala project. The spirit, approach and methodology of this project may be had from [5]. The project aims at enhancing the quality of education through collaboration across networked institutions. While the Scientific Portal provides the Knowledge Infrastructure for supporting multiple VEs, Education Grid enables academicians and experts from multiple institutions form distributed communities of practitioners over the Web in different subject areas. The approach to maintaining such communities in Education Grid is shown in Fig. 5. It is assumed that each participating college has state-of-the-art education servers that have Learning Management System, Group Collaboration support and ways to organize and access their digital collections like a Digital Library.

In each subject, one of the premier institutions is given the responsibility as a resource institute. The Resource Institute maintains the Digital Library and Resources relevant for the concerned subject. The teachers in the subject and experts of the resource institute along with identified mentors or experts from other places form the community like a VE for the subject. They compile the e-courseware, conduct teacher training and release approved content for use in the different colleges.

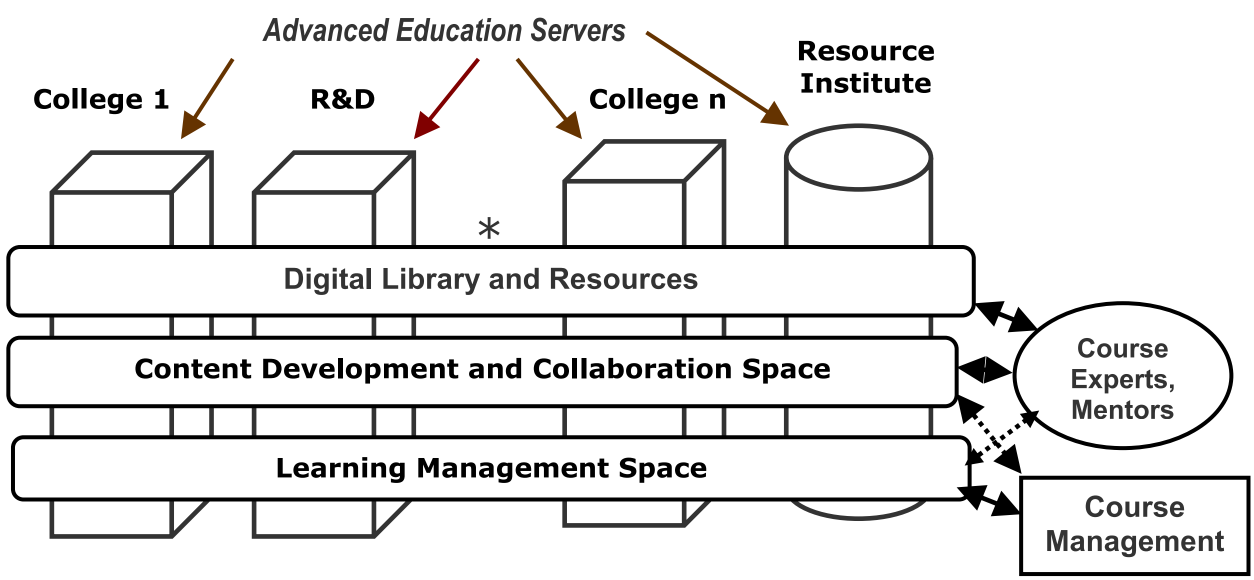

Fig. 5: Course Knowledge and Collaboration Spaces over Education Grid

Processes to monitor quality of content and effective learning delivery are being developed and put in place. The Kerala Government is studying to formalize the Education Grid as the Virtual Institute of Science, Technology and Arts [VISTA] and make it a cooperative knowledge-sharing forum across the colleges. An advanced research and management school in Informatics will drive the core of VISTA's programs and services. The Education Grid project will be augmented by adding educational satellite based remote classroom services and computational portals in different subjects located at institutions.

4. SUMMARY

Building knowledge infrastructure for socially relevant areas like Education, Agriculture, Health, etc. needs several innovative systems, collaboration structures and distributed processes supported by relevant strategy groups with members from different organizations. It also needs a refreshingly different look at the value propositions of the services provided in the light of the five laws of Information Sciences as stated in Section 2. This in turn needs several new information systems organization and management practices facilitated ably by modern enterprise portals supporting web-services enabled group specific activities.

The concept and requirements of Scientific Portal emerges out of the above considerations. IIITM-K in its KISSAN-Kerala and Education Grid projects has built the early prototypes and commissioned such advanced knowledge management systems. The cases of Kerala Education Grid and KISSAN-Kerala projects are stated as examples of how we use the concepts delineated in this paper to provide services to the Higher Education and Agriculture sectors. Scientific Portals like the one proposed for Agriculture in this paper will form the core knowledge infrastructure for thematic web-communities that provide focused services of value to its customers.

References

1. S.R. Ranganathan, "Five Laws of Library Science", Sarada Ranganathan Endowment Publications, Reprinted 1999.
2. The Abstract Syntax Notation – 1 standard is covered in the RFC Documents Series of Internet Standards. Refer to the ASN.1 Site http://asn1.elibel.tm.fr/en/index.htm
3. K.R. Srivathsan, "The 'I' in 'IT' – A Pancha Kosha View", Global Journal of Flexible Systems Management, Vol. 5, No. 1, 2004.

4. K.R. Srivathsan et al, “Strategy focused IT facilitation of Agriculture Extension Services – The KISSAN-Kerala Approach.” Presented at the National Workshop on Alternative Approaches in Agriculture Extension Services, held Feb. 2004 at Mitraniketan Krishi Vigyan Kendra, Vellanad, Trivandrum District.
5. K.R. Srivathsan, “Management Of Refereed Content Generation And Utilization In Formal Higher Education”, Global Journal of Flexible Systems Management, Vol. 4, Nos. 1 & 2, Jan. – June 2003.

Submitted as part of the Festschrift in honor of Prof. A. Neelamegham.

Dec. 2004.