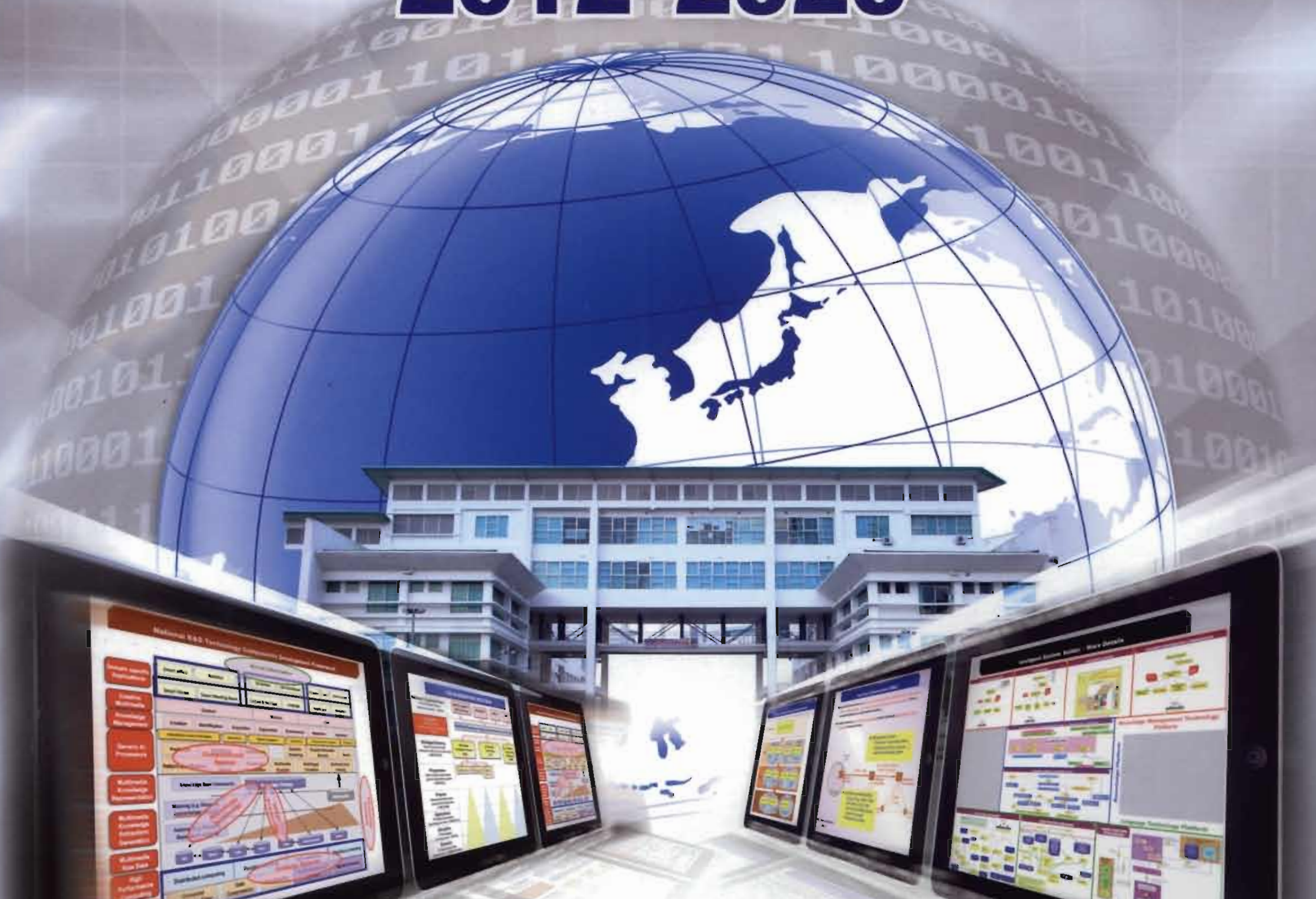
UNIMAS
MOVING FORWARD

UNIVERSITI MALAYSIA SARAWAK

FACULTY OF COMPUTER SCIENCE & INFORMATION TECHNOLOGY

R&D ROADMAP 2012-2020



Faculty of Computer Science and Information Technology

FoCuS IT
focus on individual transformation

22 February 2012

TABLE OF CONTENTS

MESSAGE
FROM THE VICE CHANCELLOR

Professor Datuk Dr Khairuddin Ab-Hamid
Vice Chancellor

I would first like to congratulate the Faculty on their efforts for producing this Research Roadmap document. The Roadmap provides clear directions for potential research and serves as a basis for developing flagship applications. The production of this document has been aptly timed as this year marks the 20th anniversary of UNIMAS. It demonstrates the maturity of the Faculty in shaping its research focus to be in line with developments at both the national and global levels.

I am aware that this document is the outcome of two workshops and a series of brainstorming sessions held at the faculty. I would particularly like to record my appreciation to Professor Dr Zaharin Yusoff for his efforts in coordinating and ensuring that this document is delivered on time.

This roadmap will serve as the blueprint for charting strategic directions and framing the overall research work to be undertaken in the Faculty. There is however much to be done to translate this roadmap into grants and research projects to actually move UNIMAS closer to achieving its aspirations of becoming a research university by 2015.

I hope that this initiative will serve as a catalyst to inspire other faculties to do the same as well. I also wish the faculty all the best in their endeavour to move the Faculty forward to become a leading research centre in the area of computer science in the region.

FOREWORD

Professor Dr Narayanan Kulathuramaiyer
Dean,
Faculty of Computer Science and Information Technology

Making research our way of life

We now find ourselves in a unique situation where "software is going to dominate everything", and is just about to change our lives beyond anything we have ever seen before. The question we pose ourselves is, how can we capitalise on this situation, to become leaders and to play a meaningful part in shaping the world we live in.

In response to this challenge, the Faculty has strategically expanded on its research activities by both nurturing a research culture and by enhancing capacity for knowledge creation. This research roadmap document is thus timely in steering our efforts towards meeting this goal.

This research roadmap highlights both the proposed strategic focus as well as the core areas of our research, based on our targeted research profile. We have charted out our research directions and technology focus for the next 5-10 years in strategic areas, well aligned with the national R&D roadmap. Our emphasis is now directed towards building the foundations for effectively generating intellectual property and value-adds, powered by our enterprising and empowered human capital. Research and continuous knowledge creation will then be at the centre of everything that we do.

We want to play the role of leaders who are able to shape the world, and make a significant impact in this emerging era. We challenge you to share our dreams and vision and invite you to collaborate with us as partners in bringing about a meaningful change to our society and world.

INTRODUCTION

1. Objective

This document is aimed at being the blueprint for the *UNIMAS FCSIT R&D Roadmap: 2012-2020* in high impact and niche areas based on FCSIT's current and future expertise. The ultimate goal for the roadmap is:

> ***Conducting research & development and commercialising next generation infra- and info-structures, content, core competencies, services and products towards affordable ubiquitous broadband platforms for mobile internet running intelligent knowledge management technology systems/applications, with generic toolkits for software architects and developers for these domains.***

2. Overall Roadmap

The ultimate goal as stated will be delivered via three strategies:

- **Affordable Ubiquitous Access (AUA)**
  (subgoal – *developing a computer networks toolkit for providing affordable ubiquitous broadband platforms for mobile internet*).

- **Knowledge Technology (KT)**
  (subgoal – *developing state-of-the-art tools, modules and content for effective knowledge management, culminating in an intelligent system builder that will expedite the development of intelligent applications*).

- **Software Engineering Workbench (SEW)**
  (subgoal – *developing a generic software/application development system for software architects and developers*)

As mentioned above, the R&D Roadmap revolves around the current and future strengths of the faculty (based on staff, R&D grants, postgraduate supervision, etc.), namely in the following 5 domains:

- Computational Science *(more of Computational Modeling)*
- Software Engineering
- Computer Networks
- Image Processing
- Knowledge Technology

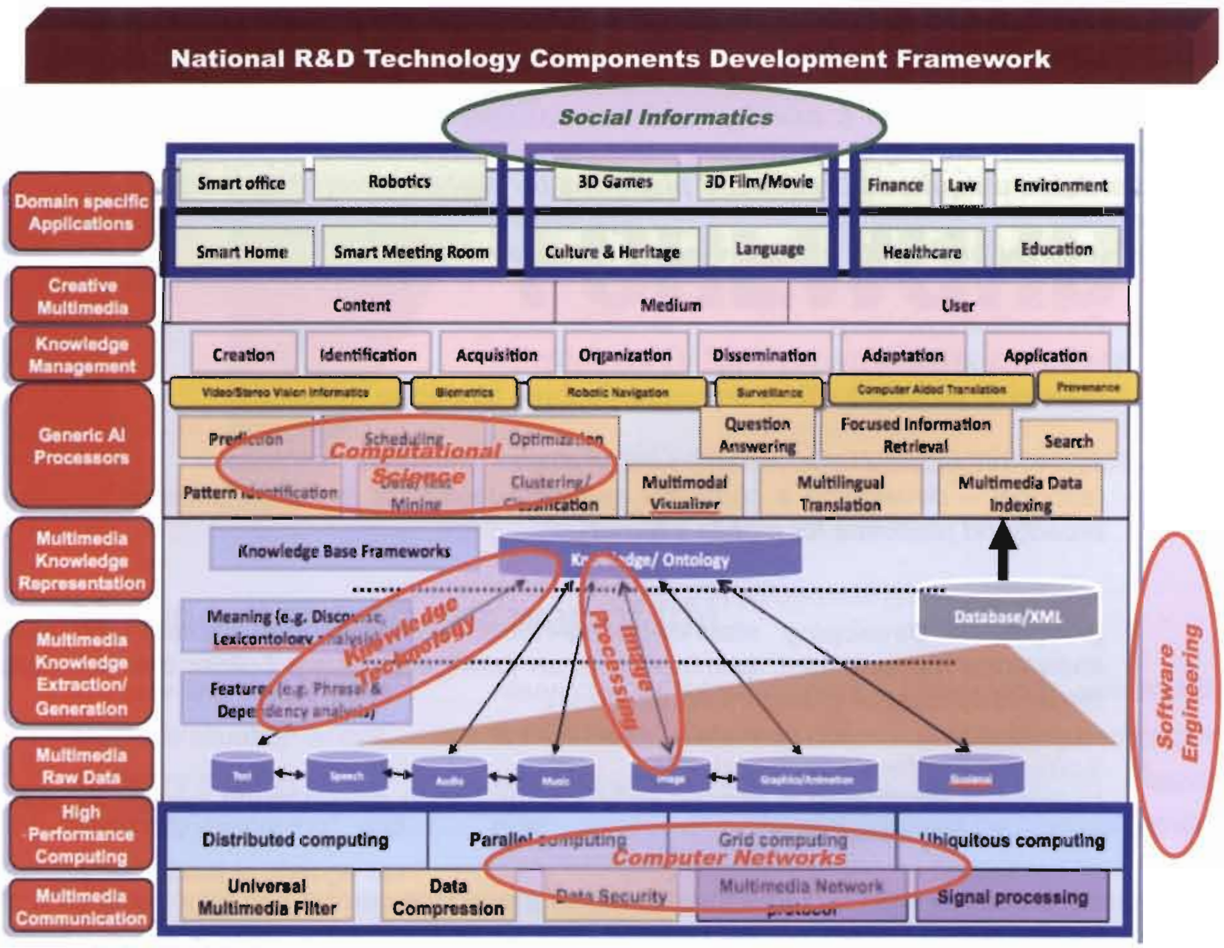
The groups' initiatives can also be mapped onto the **National R&D Technology Components Development Framework** as shown in the diagram below – with the addition of applications in **Social Informatics**, a strength quite unique to the faculty (in collaboration with several other faculties) at UNIMAS, given its geographical location and the many initiatives carried out for a variety of remote and rural communities.

The said National R&D Framework plays a major role in the national ICT objectives in:

- Producing resilient (long shelf-life) ICT graduates,
- **Moving the country towards a producer nation in ICT**,

as was proposed by the MOHE-adopted ***National ICT Human Resources Task Force***:

- involving 4 Vice Chancellors, IPTA ICT Deans, PIKOM, MNCC, MOHE, MOSTI, MAMPU, JPA, MDEC, MIMOS, ...
- documented in *"**ICT Human Capital Development Framework**"* [MOHE Oct 2010]



It is to be noted that this R&D roadmapping initiative is very much in line with the faculty's **Strategic Plan 2008-2015**, where under the section R&D, it is stated that:

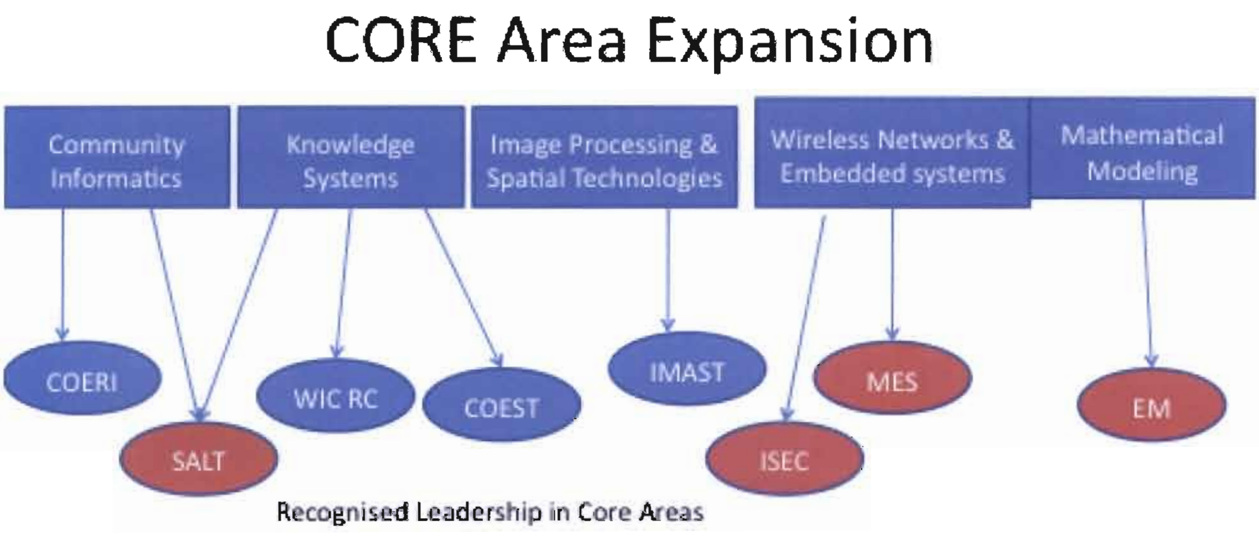
In guiding R&D at the Faculty, efforts will be needed for the following initiatives:

- *Research Portfolios: Career Plan for Academics*
- ***Roadmaps (a separate document is being developed as addendum-see attached)***
- *Idea Bank: Proposals, Plans, Matrices*

Five planned target Centres of Excellence (COE) will drive the R&D and C in the Faculty:

- *Image Processing and Spatial Information*
- *Social (Community) Informatics*
- *Networking and Systems (Mobile and Embedded Systems)*
- *Semantic and Knowledge Management Technologies*
- *Mathematical /Computational Modeling*

The following figure shows the expansion of these niche areas to come up with specific Labs and physical entities.



- *Image Processing and Spatial Technologies*
  - Center of Excellence for Image Analysis and Spatial Technologies (CoEIMAST)
- *Social/Community Informatics*
  - Center of Excellence for Rural Informatics (CoERI) [Now, upgraded to: Institute of Social Informatics and Technological Innovation (ISITI)]
  - Sarawak Language Technology (SaLT)
- *Networking and Systems*
  - Mobile and Embedded Systems Lab (MES)
- *Semantic and Knowledge Management Technologies*
  - Center of Excellence for Semantic Technology and Augmented Reality (CoESTAR)
  - Web Intelligence Consortium Research Center (WIC RC)
- *Mathematical /Computational Modeling*
  - Focusing on Environmental Modeling (EM)

Returning to the mainstream discussion, the R&D Roadmap for the faculty covers a period of 9 years (2012-2020), and encompasses two levels of R&D:

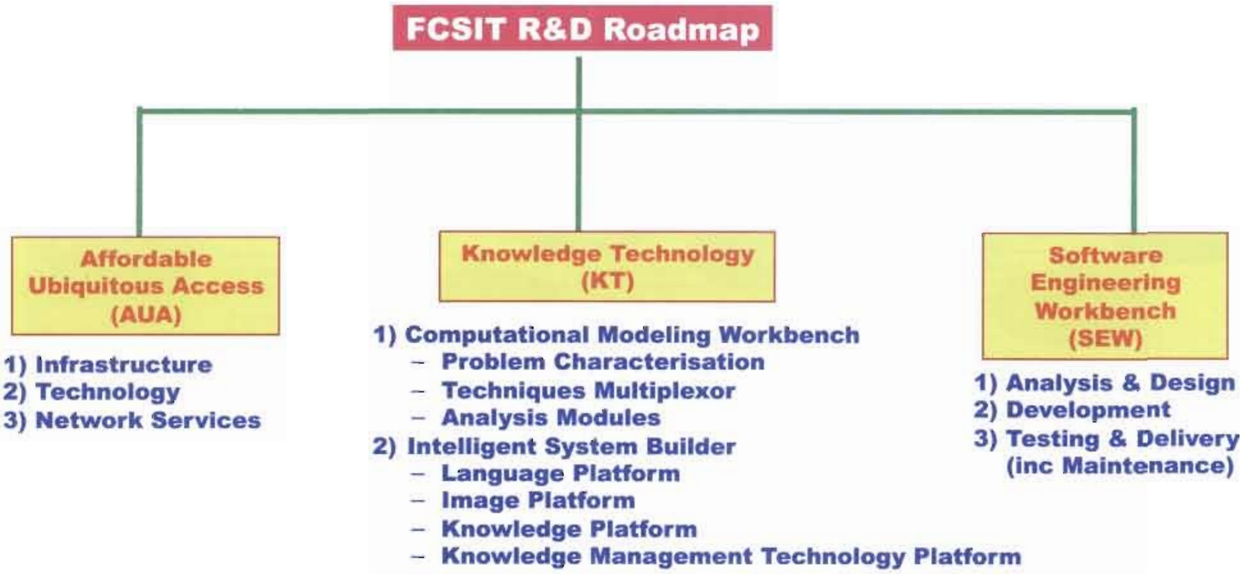
- **Technology Components Development Roadmap (TCDR)**
- **Flagship Applications (FA)**

The TCDRs plan for the development of applications/modules/data in a modular, reusable and incremental manner, such that the results can be readily put together to build various applications or provide solutions to varying problems. The Flagship Applications are designed to use as many as well as to cover a wide range of the modules/data developed in the TCDRs, and in particular with as much significant economic or social impact as possible.

Each of the five R&D Groups has its own internal TCDR and a Flagship Application – hence forming a mini-roadmap. The TCDRs can clearly be synergised across the R&D Groups as well as with external partners. For potential collaborators and/or users, of major interest would be the Flagship Applications, which have clear tangible results and very visible socio-economic impacts. The FCSIT R&D Roadmap is formed by combining the TCDRs from the R&D Groups

(in this case under the umbrella of the Software Engineering's TCDR, since theirs is the most generic), and by providing a bigger picture to cover all the Flagship Applications.

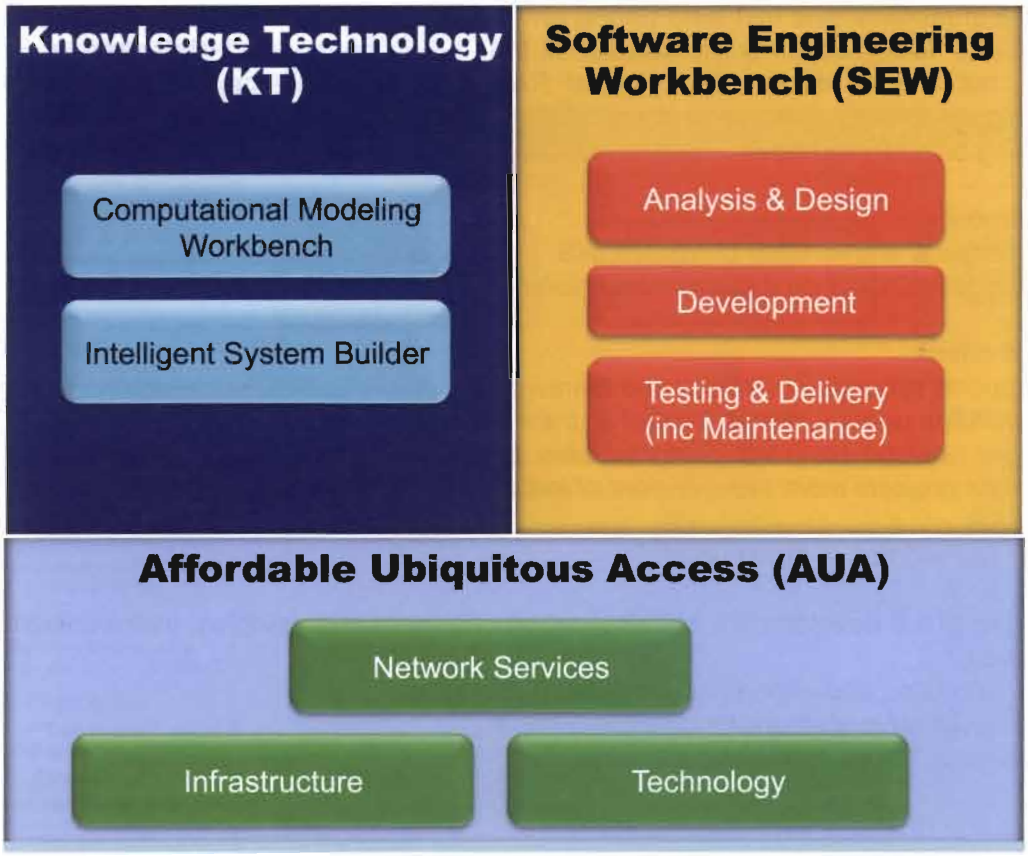
It can be seen that, apart from being the means of delivery of the ultimate goal, the strategies mentioned earlier have also been derived from the TCDRs and Flagship Applications of the 5 groups. As such, each strategy can be further divided into several programmes (a total of 8) as depicted in the diagram below.



Flagship Applications:
- Affordable Ubiquitous Broadband Platform for Mobile Internet
- Within GEOHAZARD: for landslides and floods
- Social Informatics for Minority Groups
- Preservation of Languages & Culture/Heritage (CHERITA)
- Software Development Workbench for FCSIT

A component view of the Roadmap is given in the figure below, whereby **Affordable Ubiquitous Access (AUA)** is seen to provide the delivery network infrastructure (by way of Infrastructure, Technology and Network Services) for **Knowledge Technology (KT)**, which within itself provides the necessary info-structure for the development of content and high level applications/services using the Intelligent System Builder, as well as the usage of knowledge based on the Computational Modeling Workbench. The innovations developed within KT make use of the tools within the **Software Engineering Workbench**.

Component View of Roadmap



This document will detail out FCSIT's R&D plans for the three strategies that comprise the eight corresponding programmes. The general principles within each strategy will be outlined, and for each programme within the strategy, the target projects will be laid out followed by the technical details. The document will also cover the programme's requirements in terms of manpower and costs, as well as its implementation plan. All these will be summarised at the strategy level. There will be several exemplar projects proposed for each programme. Exemplar projects are projects that implement the core components of the roadmap but will also show their application in some chosen domain.

Within the roadmap, five Flagship Applications that cut across the strategies are identified for immediate implementation. These flagship applications do not only reflect the main targeted niche areas but also FCSIT's current strengths. They also implement the core components of the roadmap, as well as demonstrate the feasibility and relevance of the roadmap.

3. Roadmap Document

An R&D roadmap document of an institution lays down the plans for the institution as to what to aim for in the immediate and mid-term future, and is expected to catalyse and boost many R&D initiatives. It is a 'living' document, in the sense that it may be modified according to situations and needs, but it does declare the general R&D directions and focus of the institution. A roadmap belongs to both management and staff. For the institution, and in the eyes of external parties, having a roadmap shows:

- Commitment
  - *Provides & shows clear DIRECTIONS*
  - *Provides FOCUS on strengths and niche areas*

- Planned efforts
  - *Supports synergies of efforts and teamwork*
  - *Facilitates prioritisation of budget & grant applications*
  - *Aligns newcomers (staff, postgraduates, project students)*
  - *Makes projects more independent of individuals*

Furthermore, a roadmap is made up of:

- Big picture of the development and application of generic technologies, methodologies & techniques:
  - *Consolidates and expands current/base strengths*
  - *Looks for advances and/or quantum jumps*
  - *Applies in future applications*

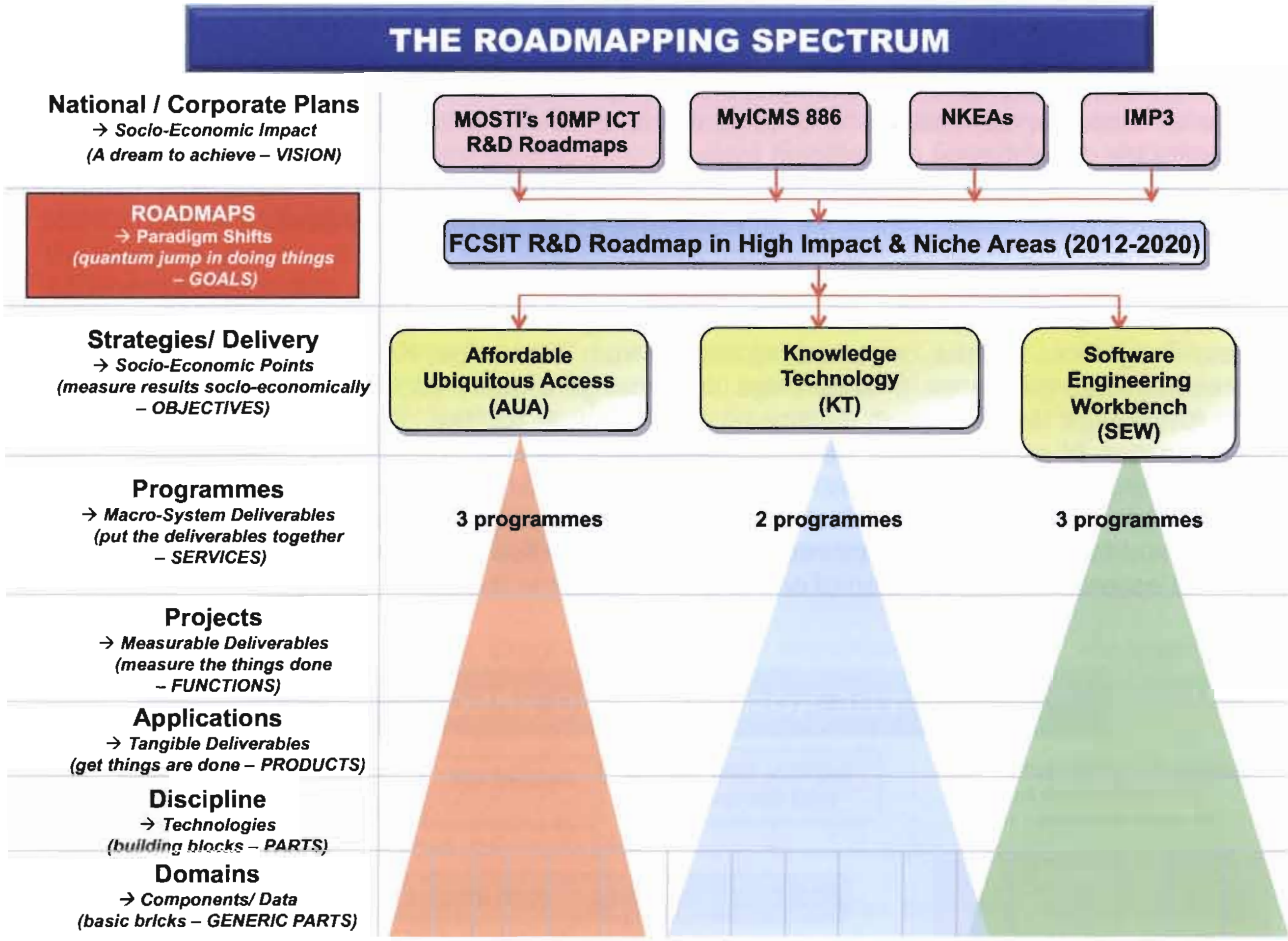
- The W's and H's
  - *What*
  - *Why*
  - *How*
  - *Who*
  - *When*
  - *How much*

A roadmap is NOT:

- ❑ *Description of current work, but of future aspirations*
- ❑ *Set of activities, but of generic technologies, methodologies, techniques*

It is also important to note here that a roadmap is by design very large and cannot be implemented by a single institution/faculty. There should always be a need to collaborate with external parties to gain fresh input and different perspectives. Some components may also be part of a general wish list, which may not be achievable with current resources and expertise, but they will serve as targets to plan for and to work towards very seriously.

FCSIT's Roadmap is based precisely on the principles stated above. The way it has been developed is also from a mixture of top-down planning – to ensure focus on very much required domains and to attain the necessary high level of R&D, and from bottom-up input – to ensure feasibility based on current strengths. This approach can be visualised via the Roadmapping Spectrum in the figure below.



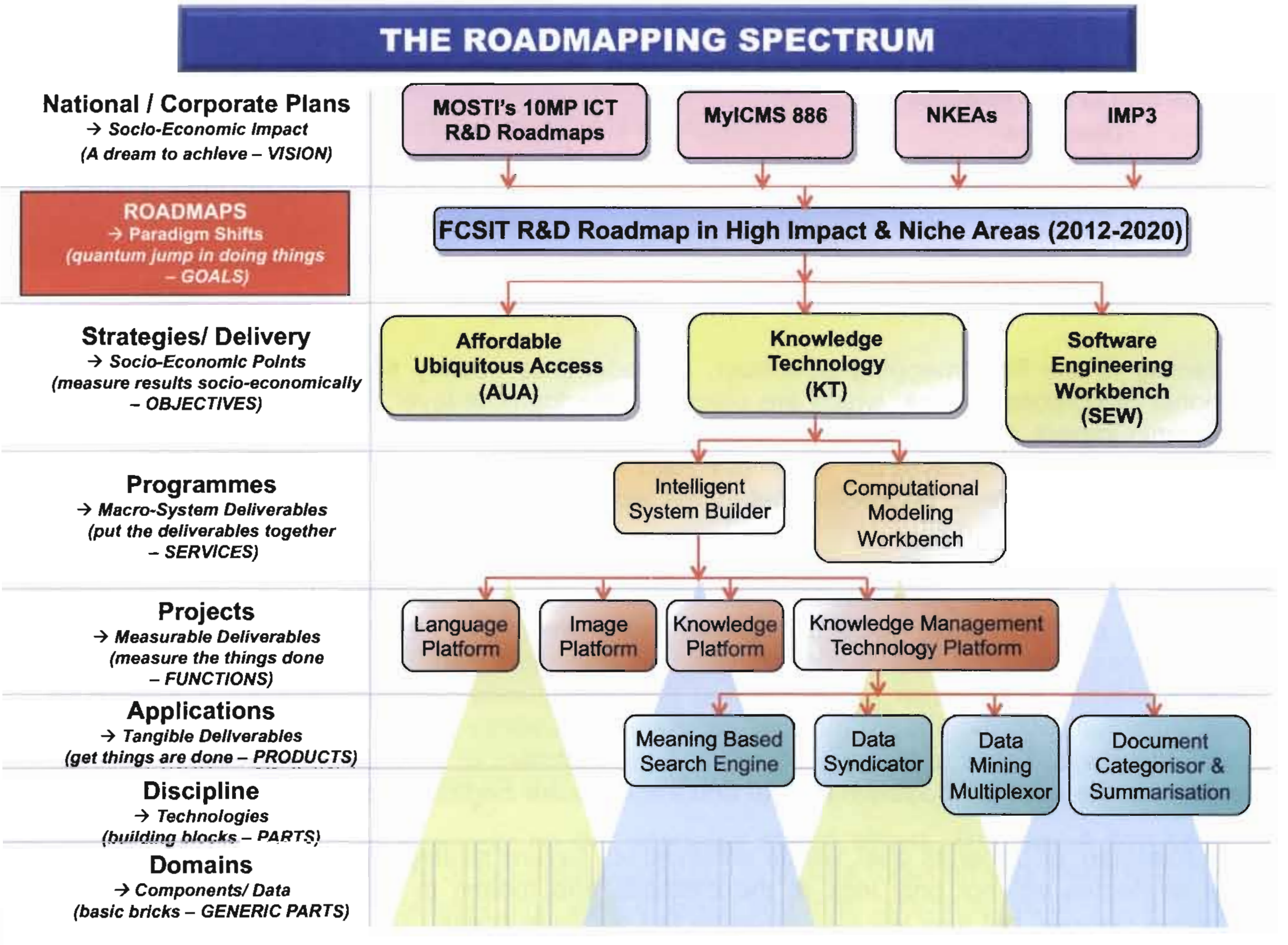
Referring to the Roadmapping Spectrum, a roadmap is usually tied to several higher level national or corporate plans, which are placed at the topmost level. These target some socio-economic impact, and express some dream to achieve – a VISION. In FCSIT's case, its R&D roadmap is tied to the MOSTI's 10MP ICT R&D Roadmaps, MyICMS 886, NKEAs, and IMP3, each of which have their own vision that FCSIT would subscribe to.

The actual roadmap is placed at the second level of the roadmapping spectrum above, which normally targets some paradigm shifts (or a quantum jump in the way of doing things), and it expresses clear GOALS. FCSIT's roadmap has elements of paradigm shifts in most of its programmes. Examples include the emphasis on affordable ubiquitous broadband platforms for mobile internet, in particular for remote and rural communities, with applications in knowledge and language technologies for the said communities towards the preservation of languages & culture/heritage. There is also the design and development of very generic application/system builders in the *Intelligent System Builder* and the *Software Engineering Workbench*.

The roadmap, or rather its goal, will be achieved via several strategies – three, in FCSIT's case. The strategies will not only look at the methods and means to attain the goal in the most effective manner, but they will also measure the results/success using well-defined quantifiable socio-economic points. These points will be positioned as the OBJECTIVES of the strategies.

Strategies will be implemented via programmes – a total of 8 in FCSIT's case. Programmes set their targets in terms of macro-system deliverables, and they put the deliverables together to offer high level SERVICES. For example, the *Intelligent System Builder* integrates the extensive range of modules, applications and content into a functional workbench to (at least) semi-automate the development of intelligent systems.

Programmes will be implemented via projects, where each project implements a FUNCTION of the programme. For instance, the *Intelligent System Builder* has at least four functions which will be implemented as projects: *Language Platform* – for the extraction/acquisition of knowledge from texts, or generating texts from knowledge; *Image Platform* – for the extraction/acquisition of knowledge from images, or generating images from knowledge; *Knowledge Platform* – for the representation/organisation of knowledge and intelligent query into the knowledge repository; and *Knowledge Management Technology Platform* – for the adaptation, application and dissemination of knowledge via generic knowledge application systems. A project in itself may be quite large – equivalent to, say, the e-Science top-down project, or a Techno-Fund project, depending on its size and commercial viability. It is also almost always multi-disciplinary in nature, and has very clear measurable deliverables. These are shown in the diagram below, which is essentially an expansion of one of the strategies in the earlier diagram.



As also shown in the diagram above, projects are put together from results of sub-projects, which are often applications of technologies that are deployed to get specific tasks carried out, and can be seen as PRODUCTS. The results of applications are very tangible deliverables. For example, the Knowledge Management Technology Platform may have several applications/ products, such as *Meaning Based Search Engine*, *Data Syndicator*, *Data Mining Multiplexor*, *Document Categorisation & Summarisation*. Grant-wise, an application is at the level of Science Fund or TechnoFund, again depending on its scientific merit or commercial viability.

Finally, applications are made up of discipline based technologies, which form the building blocks or PARTS of the products/applications. These are in turn made up of domain based components and data, which form the basic building blocks, bricks or GENERIC PARTS. These technologies and components can usually be applied to many different areas. For instance, data mining algorithms are at the technology level and may be used in many different products, such as prediction systems and knowledge extraction engines. These products may be deployed in many different services, such as smart home, healthcare, education, etc, at the service level. At the level of grants, these are clearly for Science Fund or short term grants. Faculty and postgraduate research are usually at this level, and it is considered core and seldom multi-disciplinary.

In the rest of the document, all programmes will be presented using this roadmapping framework and supported by component diagrams to show the relationships amongst projects within programmes.

4. Affordable Ubiquitous Access (AUA)

This is one of the three strategies in the roadmap and it is drawn from one of FCSIT's main strengths, which is **computer networks**. Recall that this strategy has as subgoal:

> ***Developing a computer networks toolkit for providing affordable ubiquitous broadband platforms for mobile internet***

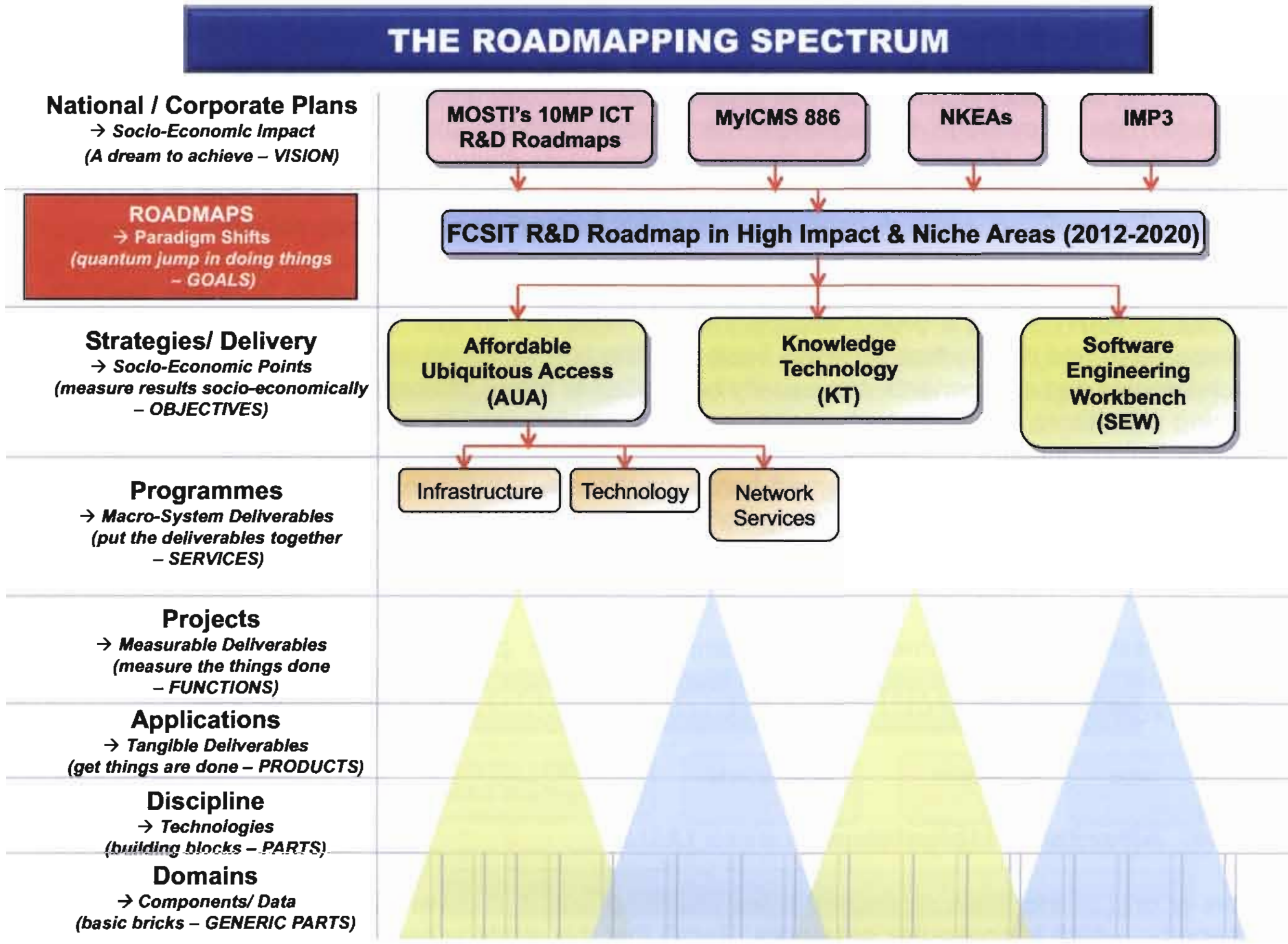
Recall also that this strategy will be delivered via three programmes, which can be seen within the roadmapping spectrum in the diagram further below:

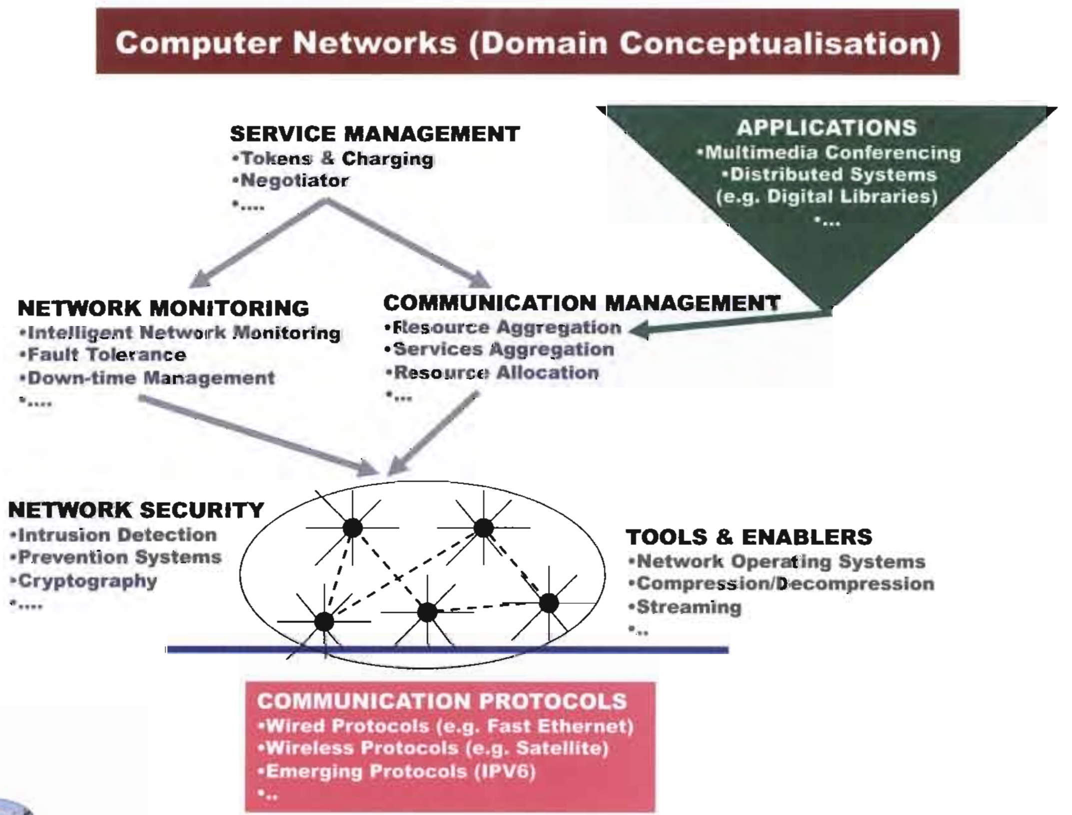
- Affordable Ubiquitous Access (AUA)
  - ❑ Infrastructure
  - ❑ Technology
  - ❑ Network Services

The above breakdown is based on several viewpoints, which are listed here and given as diagrams below in the same sequence:

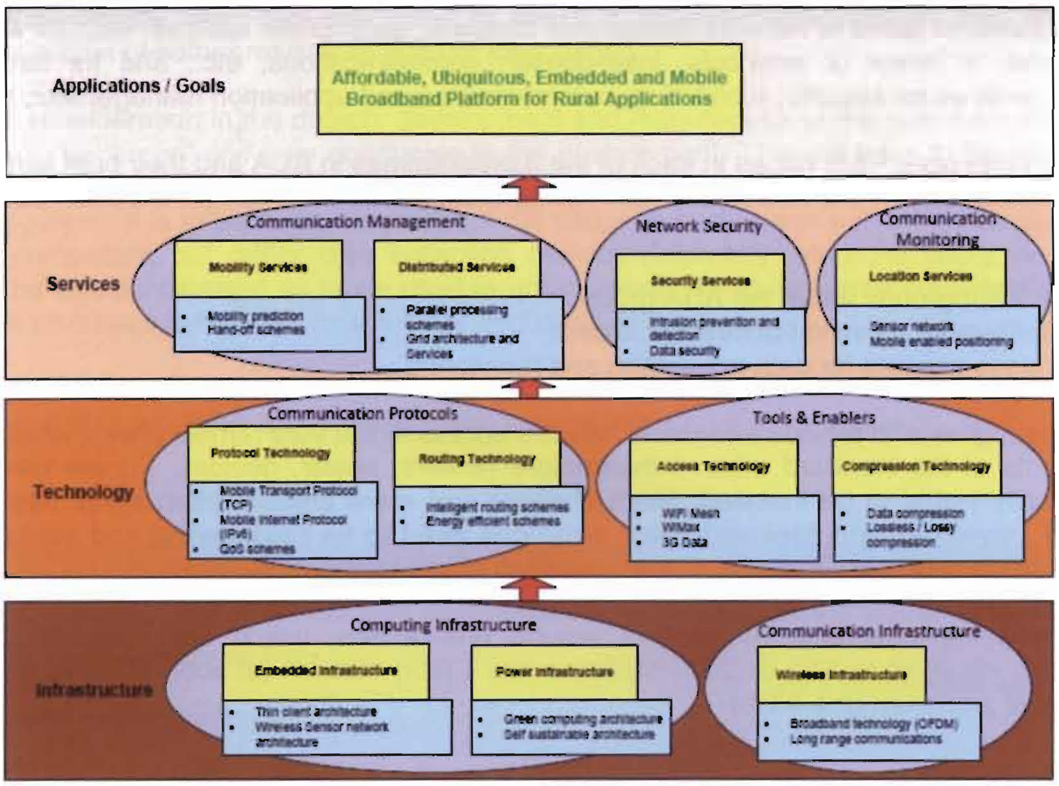
- general domain conceptualisation view of computer networks,
- FCSIT preceding roadmap for the R&D group,

These are then reformulated into a Technology Components Development Roadmap for AUA given as the last diagram in the sequence, which also forms the *Computer Networks Toolkit*.


THE ROADMAPPING SPECTRUM
National / Corporate Plans
→ Socio-Economic Impact
(A dream to achieve – VISION)
ROADMAPS
→ Paradigm Shifts
(quantum jump in doing things – GOALS)
Strategies/ Delivery
→ Socio-Economic Points
(measure results socio-economically – OBJECTIVES)
Programmes
→ Macro-System Deliverables
(put the deliverables together – SERVICES)
Projects
→ Measurable Deliverables
(measure the things done – FUNCTIONS)
Applications
→ Tangible Deliverables
(get things are done – PRODUCTS)
Discipline
→ Technologies
(building blocks – PARTS)
Domains
→ Components/ Data
(basic bricks – GENERIC PARTS)
MOSTI's 10MP ICT R&D Roadmaps
MyICMS 886
NKEAs
IMP3
FCSIT R&D Roadmap in High Impact & Niche Areas (2012-2020)
Affordable Ubiquitous Access (AUA)
Knowledge Technology (KT)
Software Engineering Workbench (SEW)
Infrastructure
Technology
Network Services

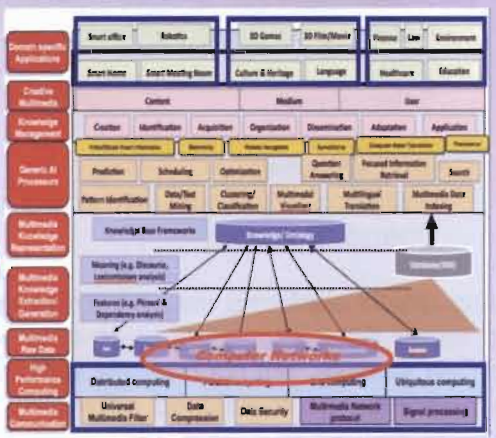

Computer Networks (Domain Conceptualisation)
SERVICE MANAGEMENT
•Tokens & Charging
•Negotiator
•....
APPLICATIONS
•Multimedia Conferencing
•Distributed Systems
(e.g. Digital Libraries)
•...
NETWORK MONITORING
•Intelligent Network Monitoring
•Fault Tolerance
•Down-time Management
•....
COMMUNICATION MANAGEMENT
•Resource Aggregation
•Services Aggregation
•Resource Allocation
•...
NETWORK SECURITY
•Intrusion Detection
•Prevention Systems
•Cryptography
•....
TOOLS & ENABLERS
•Network Operating Systems
•Compression/Decompression
•Streaming
•..
COMMUNICATION PROTOCOLS
•Wired Protocols (e.g. Fast Ethernet)
•Wireless Protocols (e.g. Satellite)
•Emerging Protocols (IPV6)
•..

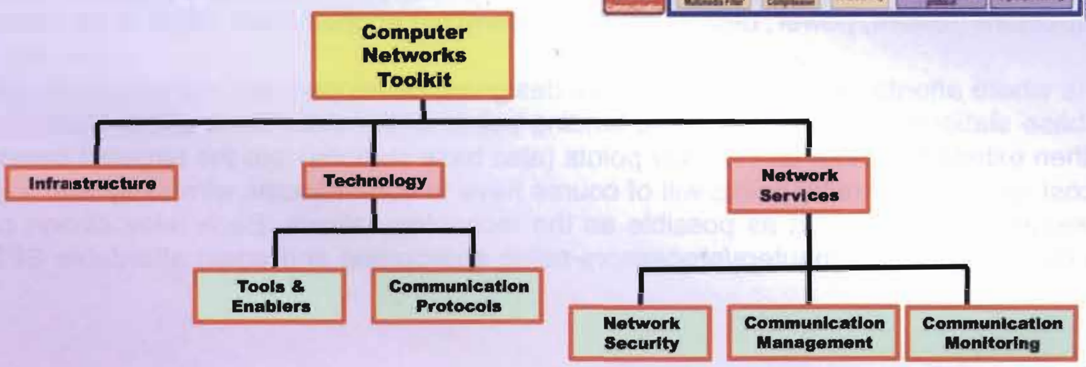
R&D Group Preceding Roadmap: Computer Networks



Technology Components Development Roadmap: AUA

- ***Computer Networks Toolkit* is a kit of tools for network engineers:**
  - ***Insfrastructure* is a toolkit of embedded, wireless and green architecture devices that provides ubiquitous and energy efficient hardware and physical networks.**
  - ***Technologies* is a toolkit of network protocols and access technology that enables seamless and efficient data delivery with wider coverage access and maximum interoperability.**
  - ***Services* is a toolkit of intelligent network management and monitoring services that allows secure mobility and distribution of services.**





The overall AUA programme has a goal of developing a ***Computer Networks Toolkit**, which is* a devices and software toolkit for network specialists, with essentially 3 compartments – Infrastructure, Technology, and Network Services. This will be used for setting up network infrastructures, in terms of network design and choice of appropriate devices, etc.; for enabling the network in terms of protocols, inter-domain communications, etc.; and for setting up services, such as for security, monitoring, communication and application management, etc.

The main innovative R&D values in each of the 3 programmes in AUA and their brief summaries are highlighted below, the details of which will be elaborated in PART I (in the main part of the document).

There are 2 subgroups within the AUA group:

- Rural/remote provision of network access
- Wireless sensors for data acquisition and monitoring

The ultimate goal is to provide affordable network access to the very remote areas, where many components of the standard urban infrastructure (towers, power, devices, ...) are clearly not economically viable to be installed. Much cheaper and more efficient alternatives need to be designed, developed and deployed. **Such solutions need to be cost saving and sustainable, but yet (ultimately) equivalent to urban installations in terms of services**.

The current implementation is centred around tele-centres, which are connected to the mainstream via satellite. The largest installation has 4 relay stations (at about 3 km apart), using the free ISM spectrum (2.4-5 Gb).

(1) Infrastructure

The infrastructure for an affordable ubiquitous broadband platform for mobile internet is made up of the following:

- Devices
  - Relay points (base stations)
  - CPEs
  - Solar power

- Green computing
  - Architecture (off-site, little power consumption, ...)
  - Operating system

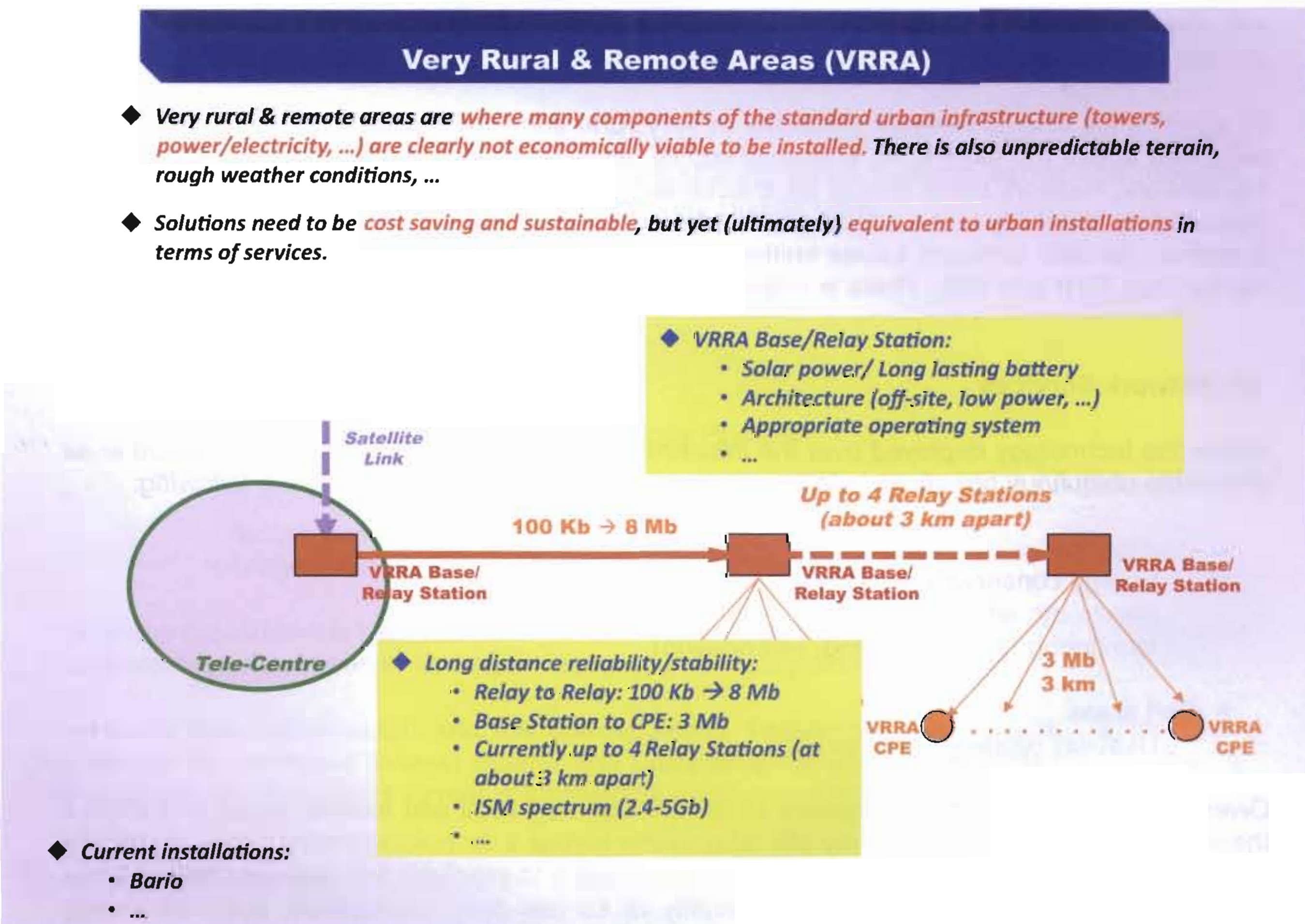
Existing urban platforms may be extended to rural platforms via expensive fibre-optic cables, the cheaper Wimax (or more up-to-date equivalents, e.g. LTE), by microwave, or the cheapest option via satellite, but connectivity will still have to be extended out within the very remote and rural platforms. As mentioned above, this is where many components of the standard urban infrastructure (towers, power, devices, ...) are clearly not economically viable to be installed.

This is where affordable devices need to be designed, developed and manufactured, beginning with base stations to be situated at the landing points of the extensions of the urban platforms, and then extended further on via relay points (also base stations) into the targeted communities. For cost reasons, the relay points will of course have to communicate wirelessly with each other and separated as far apart as possible as the technology allows. Each relay station can then span out to the final computers/processors using appropriate and again affordable CPEs, also

at a distance as far as possible from the last base station as the technology allows. Being in very rural and remote areas, electricity supply will most likely be absent, and so alternative energy sources, in particular solar, will also have to be designed, developed and manufactured for all these devices, and that also need to result in equipment/components that are very affordable, robust (weather resistant), and yet long lasting.

An overall consideration in the design, development and manufacture of the said devices is that they have to be 'green' and very conducive to the environment. This will have to be taken into account for the architecture (off-site, low power consumption, etc.) as well as for the operating system deployed. It is to be noted that most of the usage in such areas will not be humans with personal computers, but rather data collection devices (telemetry, etc.). As such, the CPEs need not be as sophisticated as those used in urban areas, but at the same considerations for 'green' will also have to be used for the other end-point devices.

The following diagram gives an overview of the current infrastructure capabilities.



(2) Technology

In terms of technology deployment over the infrastructure, an affordable ubiquitous broadband platform for mobile internet has to have minimally the following aspects to be considered:

- Network design
  - Terrain (no towers)
  - Minimum relay points
  - Maximum bandwidth

- Long distance stability/reliability
  - 3 Mb for Base Station to CPE
  - 100 Kb → 8 Mb relay to relay

Given that telecommunication towers would not be financially viable to set up in the targeted very rural and remote areas, and that the terrain would be unpredictable (and so is the weather), there is a need for a very good network design tool for the network specialists. In this situation, topology does not refer only to network topology but also to that of the terrain (hence the need for some level of GIS) in order to adjust for heights (hills, valleys, streams, etc.), lines of sight and other considerations. Again due to cost reasons, of major importance is to keep the number of relay points down to a minimum, while retaining the maximum bandwidth available.

Of significance also to network platforms in very rural and remote areas, given the conditions described above (no towers, no electricity supply, etc.), is the long distance stability/reliability of the network. Network traffic should be maintained (or at least recuperable) even if momentarily disrupted by instability in the infrastructure (devices, power, etc.). Currently, some comfort level is attained for 3Mb between a base station to a CPE, but the range from a relay point to another can be from 100Kb to 8Mb. There is indeed a lot of research that needs to be done in this area.

(3) Network Services

Above the technology deployed over the infrastructure, there are certain services required in an affordable ubiquitous broadband platform for mobile internet, being minimally the following:

- Intelligence
  - Energy conservation
  - Bandwidth adaptation
  - Monitoring (self-monitoring, self-healing)

- Soft areas
  - Usability (uptake, idiot-proof))

Given that there will be no technicians at hand in the very rural and remote areas, and even if there are problems reported it may still take weeks before a technician would arrive at the site (although some problems would take only a few minutes to resolve), the deployed network has to have a certain level of intelligence. Quality of service level intelligence, such as energy conservation and bandwidth adaptation to changing weather conditions are indeed a must, but most of all the network has to have a certain level of self-monitoring and self-healing capabilities with online reporting to central facilities.

There is also the need to educate users (or potential beneficiaries) at the remote and rural sites, where concerted efforts have to be made to ensure sufficient uptake to make the initiatives worthwhile. Often, the need may definitely be there (e.g. monitoring of crops), but the locals may not be aware of the many possibilities available, or that information has been provided but not sufficiently customised or localised for their understanding. There are also issues pertaining to logistics, where very simple network problems could be resolved by almost anyone at hand without having to wait for the arrival of technicians, but that know-how will have to be transferred in a simple way and with the use of very simple (idiot-proof) tools.

5. Knowledge Technology (KT)

The need for managing knowledge is very pertinent, receiving major emphasis in all national ICT master plans. It is only appropriate that FCSIT takes on this challenge, whereby its staff already has considerable expertise in language technology and image processing, as well as the usage of knowledge via computational modeling. Recall that this strategy has the following subgoal:

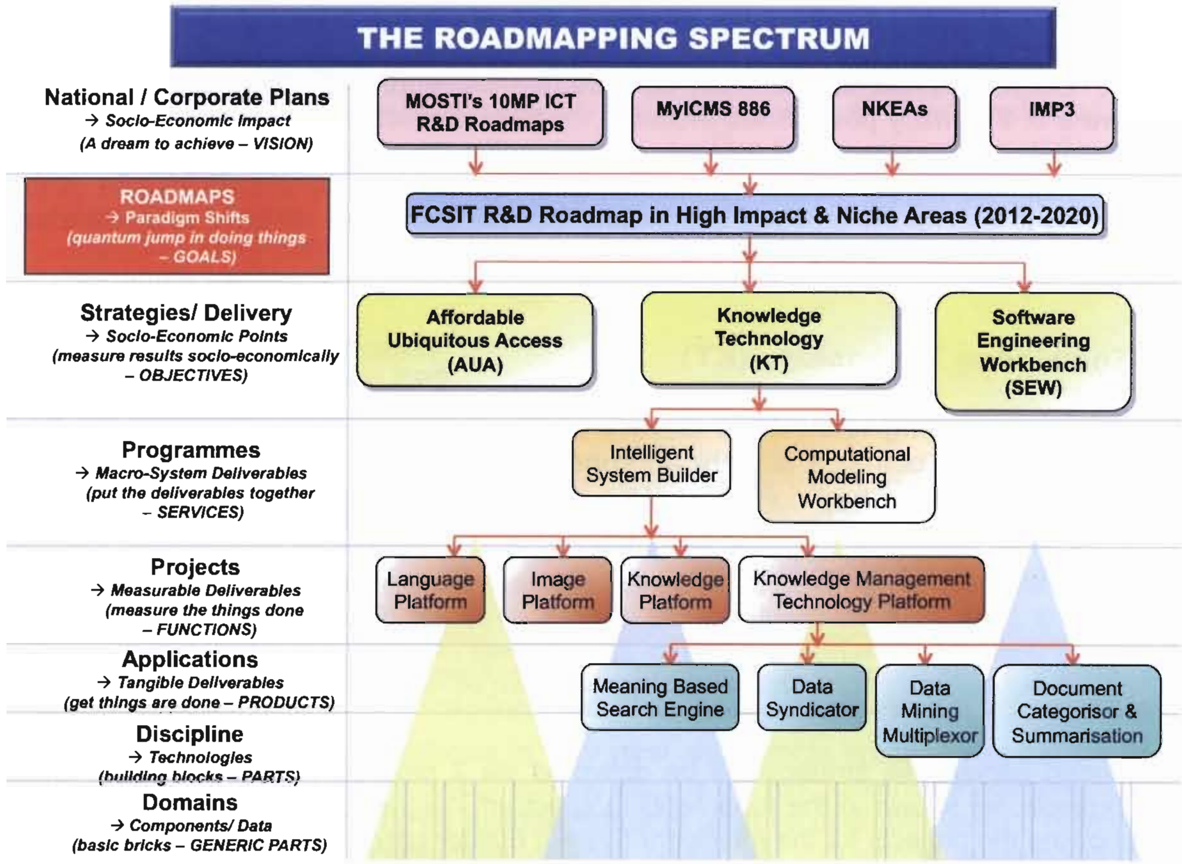
> ***Developing state-of-the-art tools, modules and content for effective knowledge management, culminating in an intelligent system builder that will expedite the development of intelligent applications***

As mentioned earlier, this strategy will be delivered via 2 programmes, and within each are several projects, as shown in the roadmapping spectrum diagram shown earlier and repeated below (showing the projects for the *Intelligent System Builder* only).

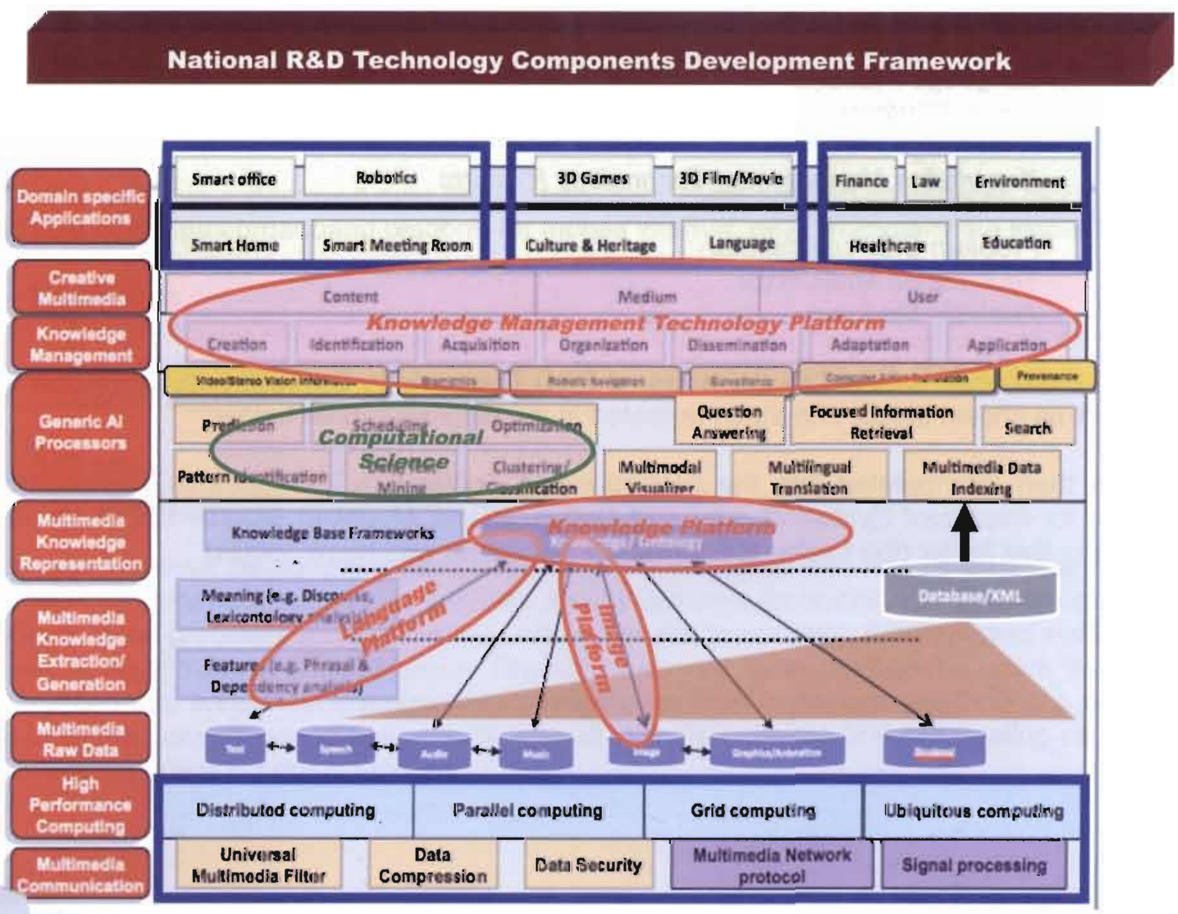
- Knowledge Technology (KT)
  - Intelligent System Builder
    - *Language Platform*
    - *Image Platform*
    - *Knowledge Platform*
    - *Knowledge Management Technology Platform*
  - Computational Modeling Workbench
    - *Problem Characterisation*
    - *Techniques Multiplexor*
    - *Analysis Modules*

The above breakdown is based on the **National R&D Technology Components Development Framework** as shown earlier and again repeated in the diagram further below.

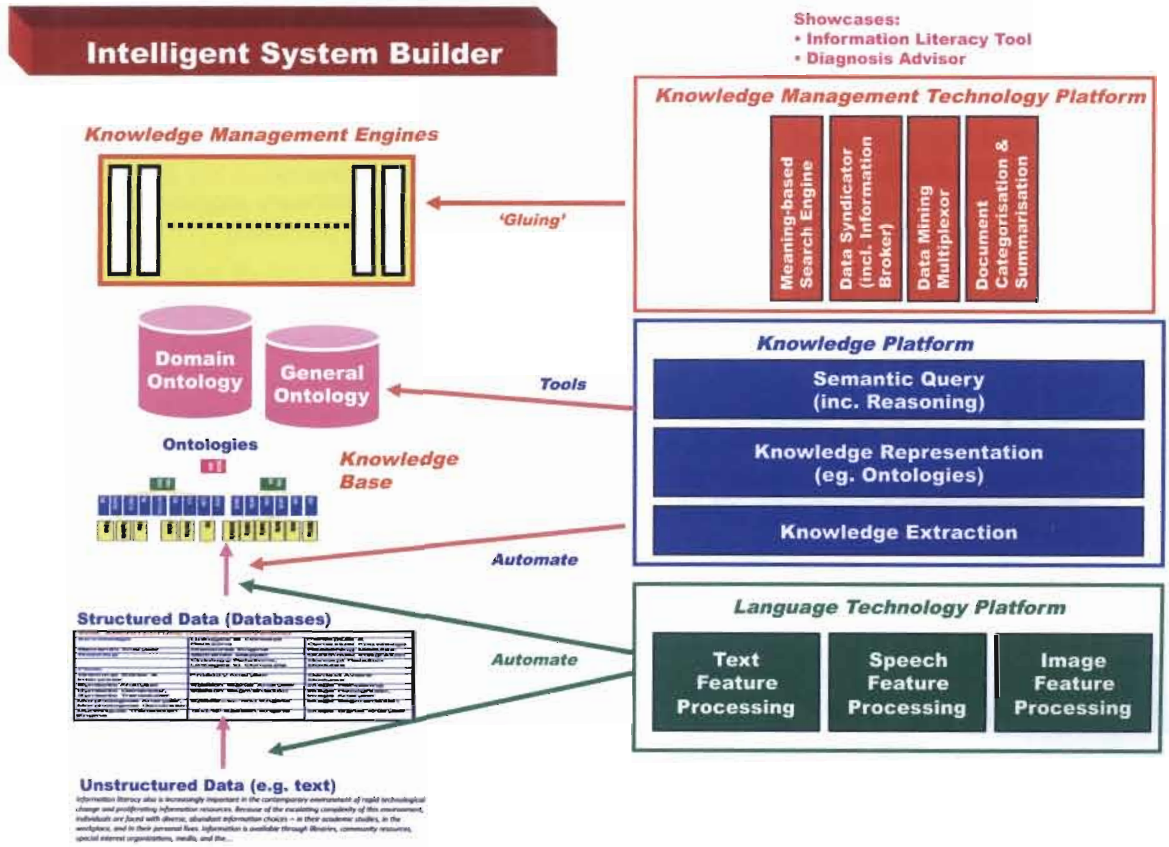
These are then reformulated into the corresponding Technology Components Development Roadmaps for *Intelligent System Builder* and *Computational Modeling Workbench* as given in the diagrams that follow (the former is followed by a diagram with more details).

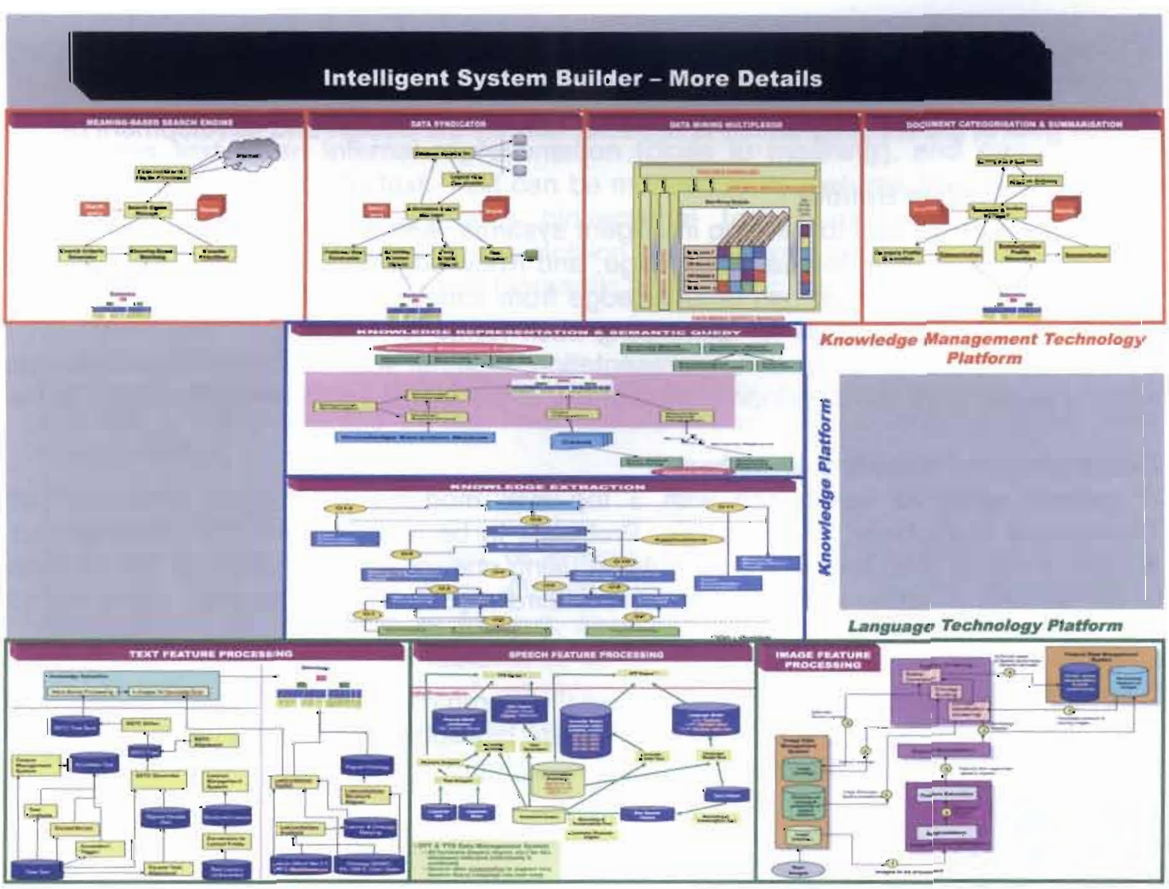


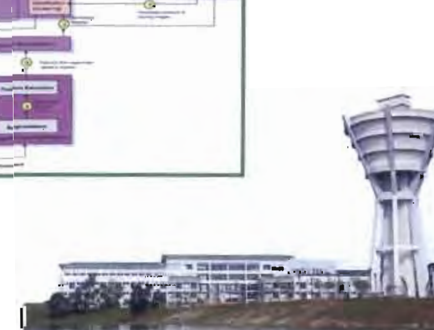
THE ROADMAPPING SPECTRUM
National / Corporate Plans
→ Socio-Economic Impact
(A dream to achieve – VISION)
ROADMAPS
→ Paradigm Shifts
(quantum jump in doing things – GOALS)
Strategies/ Delivery
→ Socio-Economic Points
(measure results socio-economically – OBJECTIVES)
Programmes
→ Macro-System Deliverables
(put the deliverables together – SERVICES)
Projects
→ Measurable Deliverables
(measure the things done – FUNCTIONS)
Applications
→ Tangible Deliverables
(get things are done – PRODUCTS)
Discipline
→ Technologies
(building blocks – PARTS)
Domains
→ Components/ Data
(basic bricks – GENERIC PARTS)
MOSTI's 10MP ICT R&D Roadmaps
MyICMS 886
NKEAs
IMP3
FCSIT R&D Roadmap in High Impact & Niche Areas (2012-2020)
Affordable Ubiquitous Access (AUA)
Knowledge Technology (KT)
Software Engineering Workbench (SEW)
Intelligent System Builder
Computational Modeling Workbench
Language Platform
Image Platform
Knowledge Platform
Knowledge Management Technology Platform
Meaning Based Search Engine
Data Syndicator
Data Mining Multiplexor
Document Categorisor & Summarisation

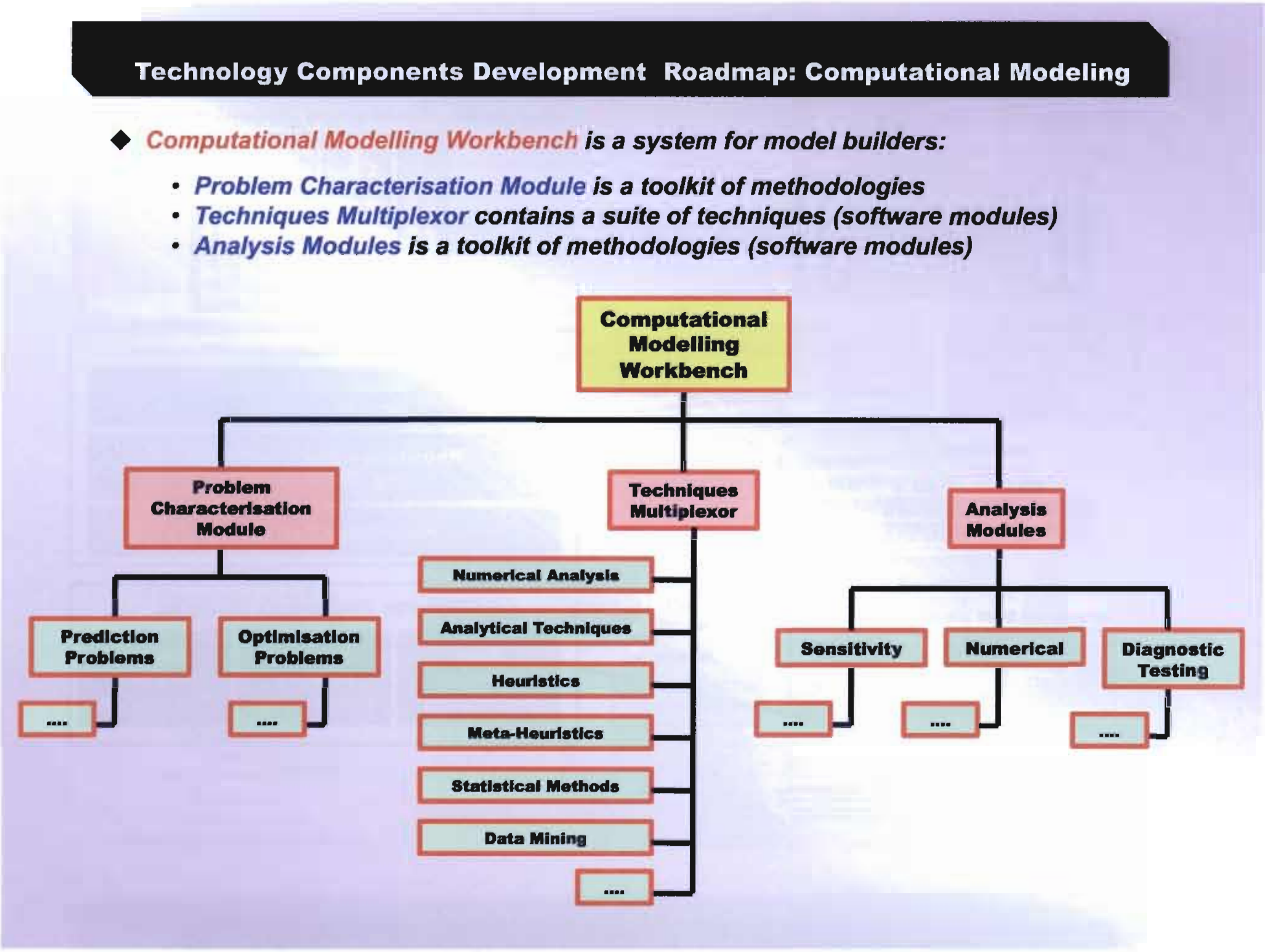


National R&D Technology Components Development Framework
Domain specific Applications
Creative Multimedia
Knowledge Management
Generic AI Processors
Multimedia Knowledge Representation
Multimedia Knowledge Extraction/ Generation
Multimedia Raw Data
High Performance Computing
Multimedia Communication
Smart office
Robotics
3D Games
3D Film/Movie
Finance
Law
Environment
Smart Home
Smart Meeting Room
Culture & Heritage
Language
Healthcare
Education
Content
Medium
User
Knowledge Management Technology Platform
Creation
Identification
Acquisition
Organization
Dissemination
Adaptation
Application
Prediction
Scheduling
Optimization
Computational Science
Pattern Identification
Clustering/ Classification
Question Answering
Focused Information Retrieval
Search
Multimodal Visualizer
Multilingual Translation
Multimedia Data Indexing
Knowledge Base Frameworks
Knowledge Platform
Language Platform
Image Platform
Database/XML
Distributed computing
Parallel computing
Grid computing
Ubiquitous computing
Universal Multimedia Filter
Data Compression
Data Security
Multimedia Network protocol
Signal processing


Intelligent System Builder
Showcases:
• Information Literacy Tool
• Diagnosis Advisor
Knowledge Management Engines
'Gluing'
Knowledge Management Technology Platform
Meaning-based Search Engine
Data Syndicator (incl. Information Broker)
Data Mining Multiplexor
Document Categorisation & Summarisation
Domain Ontology
General Ontology
Tools
Knowledge Platform
Semantic Query (inc. Reasoning)
Knowledge Representation (eg. Ontologies)
Knowledge Extraction
Ontologies
Knowledge Base
Automate
Structured Data (Databases)
Language Technology Platform
Automate
Text Feature Processing
Speech Feature Processing
Image Feature Processing
Unstructured Data (e.g. text)


Intelligent System Builder – More Details
Knowledge Management Technology Platform
Knowledge Platform
Language Technology Platform





The overall goal of the KT programme is two-fold, namely the design and development of:

- ***Intelligent System Builder***
  A generic software tool to develop intelligent systems, with 4 main (very large) Technology Platforms – Language, Image, Knowledge, and Knowledge Management Technology. The first two are for the acquisition of knowledge from various multi-media, multi-modal, multi-lingual sources (as well as for generating such forms for knowledge dissemination); the third is for knowledge extraction, representation, together with intelligent retrieval services; and the fourth is for the development/assembly of the actual intelligent systems for users.

- ***Computational Modeling Workbench***
  A generic software workbench with 3 top-level modules – Problem Characterisation, Techniques Multiplexor, and Analysis. Problems to be resolved are first characterised by the first module, then a model is formulated using one (or a combination) of the techniques in the second module, and analysed by the third module (which contains many statistical but also other types of analysis modules).

The main innovative R&D values in each of the 2 programmes within KT (and their respective Projects) and the corresponding brief summaries are highlighted below, the details of which will be elaborated in PART II (in the main part of the document).

(1) INTELLIGENT SYSTEM BUILDER

Although this segment is presented as a single programme out of a total of 8 within the whole roadmap of 3 strategies, it is by far the largest and perhaps the most complex of all. FCSIT has had collaborations with UNITEN and several other universities in knowledge technology, and this programme is an adaptation of the relevant section in the *UNITEN R&D Roadmap for High Impact & Niche Areas*.

An Intelligent System Builder is a generic software tool for software and application developers to develop rapidly intelligent systems. It has 4 main Technology Platforms – Language, Image, Knowledge, and Knowledge Management Technology. The first 2 are for the acquisition (and generation) of knowledge, the third for extraction, representation and retrieval of knowledge, and the fourth for the application and dissemination of knowledge. Each platform is a collection of tools specific to the given purpose, where they can be alternative tools for the same functions or those that work in unison or in sequence to perform a larger function. The tools may also use sets of specialised data collected for the specified functions, also placed in the corresponding platform (but of course made available universally).

The underlying principles of the programme are the following, and these can be clearly seen in the *National R&D Technology Components Development Framework* diagram above:

A. Text, speech and image are expressions of knowledge, and as such one should be able to extract and represent knowledge from documents in these modes.

B. Once knowledge is captured and encapsulated, it can be utilised effectively by using knowledge management techniques to support large and intelligent enterprise-level applications.

The first point (A) is made possible with the following:

i. *Text Feature Processing*
Analyses text to an internal representation (close to meaning), and conversely from internal representation to text. Text can be mapped to its features (morphology, syntax, syntactic functions) using various bidirectional techniques (structural analysis & generation), and with some more precision, the features may also be bidirectionally mapped to the meaning of the text (semantic analysis & generation).

ii. *Speech Feature Processing*
Processes the necessary data for use by Text-to-Speech and Speech-to-Text engines; and this is complemented by the Text Feature Processing for analysis to the internal representation.

iii. *Image Feature Processing*
Analyses feature representations of images to an internal representation (close to meaning), and conversely from internal representation to feature representations. Although not as advanced as for text, similar processes exist for images, where an image can be mapped to its features (shape, colour, texture) using various techniques (image processing – perhaps not yet bidirectional), and with much more precision, the features may also be mapped to the meaning of the image (semantic analysis – semantic gap).

iv. *Knowledge Extraction*
Another level of extraction from the meaning of both text and images should produce knowledge contained in the sources (knowledge extraction). As meaning should be modality independent (also language independent), there should also be a possibility of translating from one form to the other (multimodal translation – as speech is included within text, or convertible to and from).

The second point earlier (B) is made possible with the following:

i. *Knowledge Representation*
Once the knowledge is extracted, there are now many techniques in knowledge representation, such as ontologies, cases, semantic networks, etc.

ii. *Semantic Query*
Each form of knowledge representation comes with the corresponding retrieval methods, which also allow for reasoning.

iii. *Knowledge Management Technology*
Knowledge management is now a well-studied domain with its 7 components: creation, identification, acquisition, organisation, dissemination, adaption an application. These components have its own technologies inherited from various subdomains of artificial intelligence, and as such many generic modules and tools exist to be reused.

iv. *Intelligent Systems*
With all the above, large intelligent multilingual multimodal enterprise level applications and services can be developed and deployed.

The points (A)(i)-(ii) are encapsulated in the Language Platform, while A(iii) is in the Image Platform. The points (A)(iv)&(B)(i)-(ii) are in the Knowledge Platform, and the points (B)(iii)-(iv) are in the Knowledge Management Technology Platform.

(a) Language Platform

As described above and also given in an earlier diagram, this platform consists of the following components:

- Text Feature Processing
- Speech Feature Processing

There are some current projects that are being carried out, more at the application level, but the results can be readily modularised and further refined to be placed in the platform. The details of the targeted modules will be given in PART II of the main document.

Among the work currently being done are:

1. *Language Tools:*
   - *Generic Named Entity Recognition for Indigenous Languages of Sarawak*
   - *A Generic Text to Speech System for Indigenous Languages in Sarawak*
2. *Repository Development:*
   - *Multilingual WordNet*