Knowledge Management of Biomedical Research Data

S.K.Dey Biswas & Dr.Meera Singh
Scientist F
Indian Council of Medical Research
New Delhi 110029

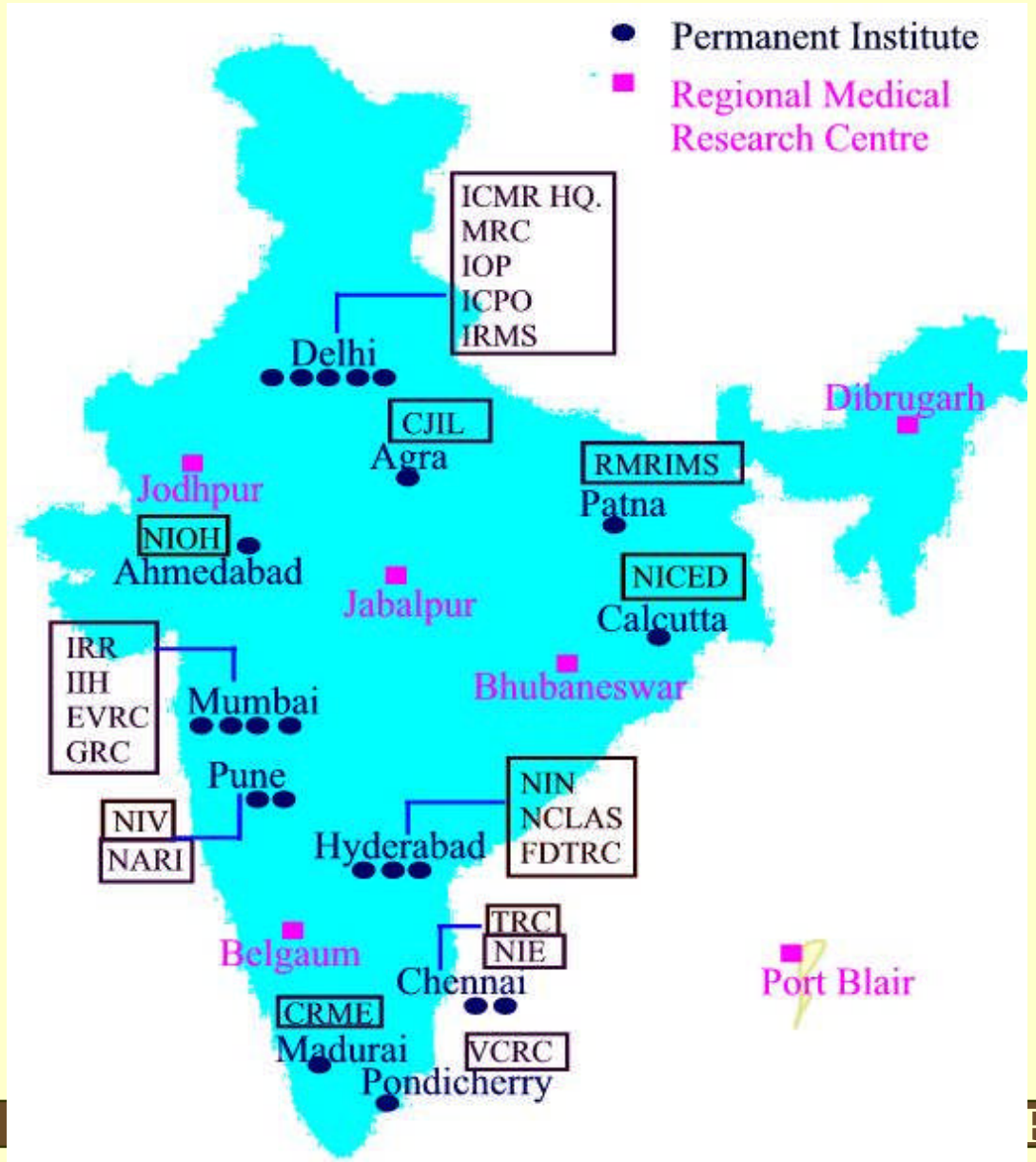
Indian Council of Medical Research

- Founded in 1911 as
  Indian Research Fund Association
- Renamed as ICMR in 1949
- Apex organization to formulate, conduct, coordinate and promote biomedical research
- Network of 26 Institutes/Centres

The Mandate

To undertake and support basic, applied and operational research in the areas of national public health importance using tools including those of modern biology.

Network of ICMR Institutes



Infrastructure/Manpower

- **Number of Institutes/Centres: 28**
- **Number of Field Stations: 70 (approx)**
- **Manpower:**

  Scientific : 792

  Technical : 2850

  Admn/supportive : 1355

Research Activities of ICMR

Intramural

- Permanent Research Institutes/Centres
- Regional Medical Research Centres

Extramural

- Task Force Studies
- Adhoc Schemes
- Centres for Advanced Research
- Fellowships
- Registries
- Other Programmes

ICMR Infrastructure

Epidemiological and Communicable Diseases

Non Communicable Diseases

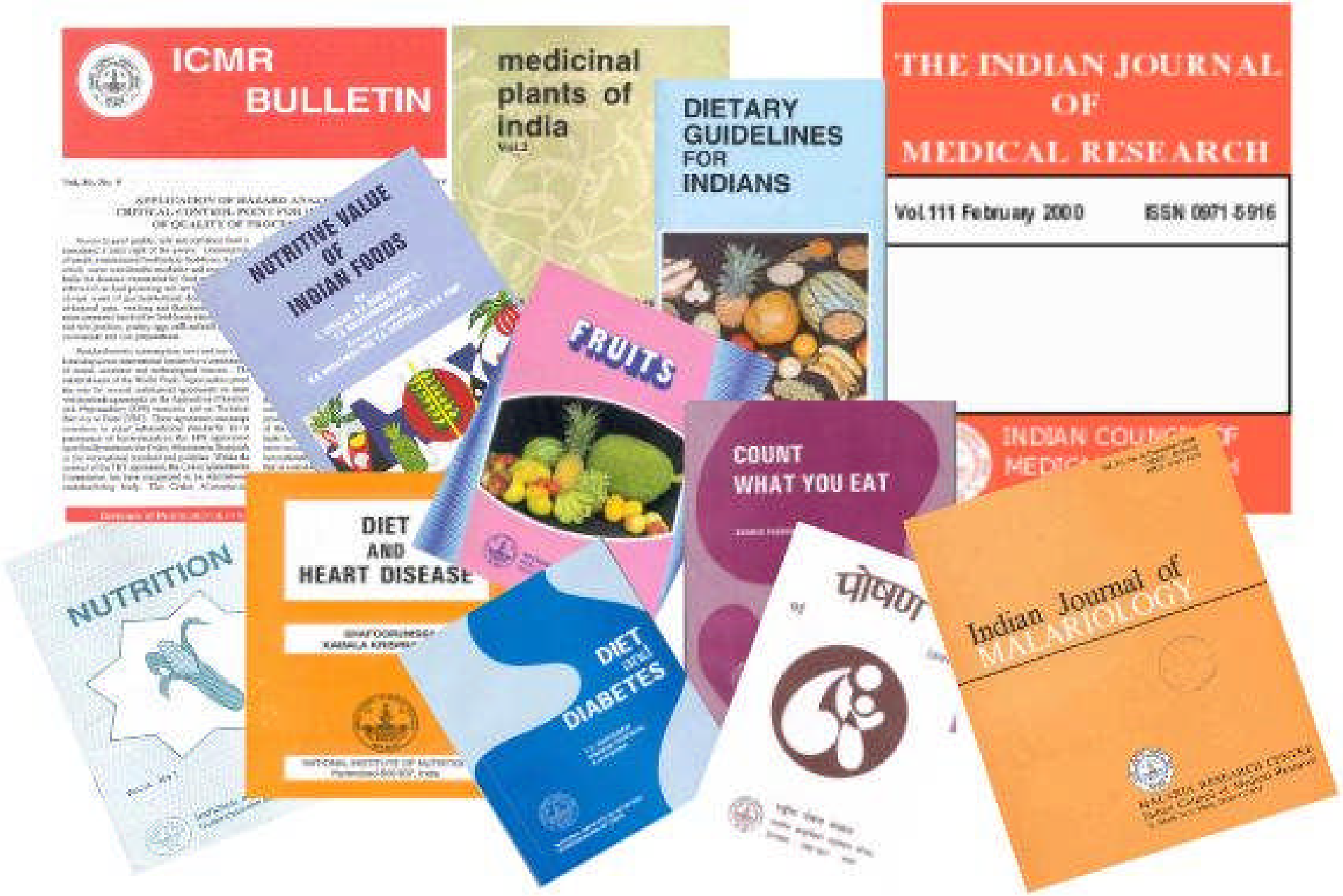
Reproductive Health and Nutrition

Basic Medical Sciences

Publication and Information

International Health

ICMR Publications



Recent Initiatives

Setting up of new Divisions/Units

- **Intellectual Property Rights**
- **Bioinformatics**
- **Alternative systems of medicine**
- **Tobacco free initiative**
- **Nutraceuticals**
- **Bioethics**

Manpower Development & Research Capability

- Junior Research Fellowships
- Talent Search Scheme
- Kishore Vigyanik Protsahan Yojana
- Short term Summer Studentships
- RCS Activities such as FETP, Nutrition, Entomology, Hematology, Diarrhoeal Diseases etc.
- Innovative Hiring Schemes and re-look at qualifications
- Rewarding researchers with good publications through incentives

Research & Data Generation

- **Institutes/Centres conduct their studies independently and are solely responsible for the completion of a particular study/research undertaken by them**
- **Each of these studies/projects generate huge amount of micro & macro level data about diseases, patient population, drug usage, drug efficacy and safety , health practices etc.**

Data Generation & Usage

- **It has been observed that most of these data are never used beyond the initial defined scope of a study or research. Also, repetition of similar or overlapping studies is common, resulting in limited usage thereafter.**
- **There is information loss since not all Centres and studies follow consistent processes and guidelines for gathering, processing, storing, accessing, archiving and analyzing critical data.**

Clinical Studies & Clinical Trials

- **Majority of the ICMR institutes are involved in conducting clinical studies either observational studies or clinical trials.**
- **The data collected are maintained locally or with the sponsors**
- **Data captures vary from manual entries, Electronic data Capture (EDC) or by using data fax technologies**

Data Storage

- Data is stored in whatever systems that prevailed at that time
- Magnetic tapes, Floppies of various sizes, cartridges, CDs
- Scant documentation
- Much of the data has lost relevance due to time
- People who have collected the data have left the organization

Data availability

- **Research data**
- **Published/printed data (Digitized)**
- **Data in databases and MIS systems**

The Information Crisis

Integrated	**Must have a single, enterprise-wide view**
Data Integrity	**Information must be accurate and must conform to business rules**
Accessible	**Easily accessible with intuitive access paths, and responsive for analysis**
Credible	**Every business factor must have one and only one value**
Timely	**Information must be available within the stipulated time frame**

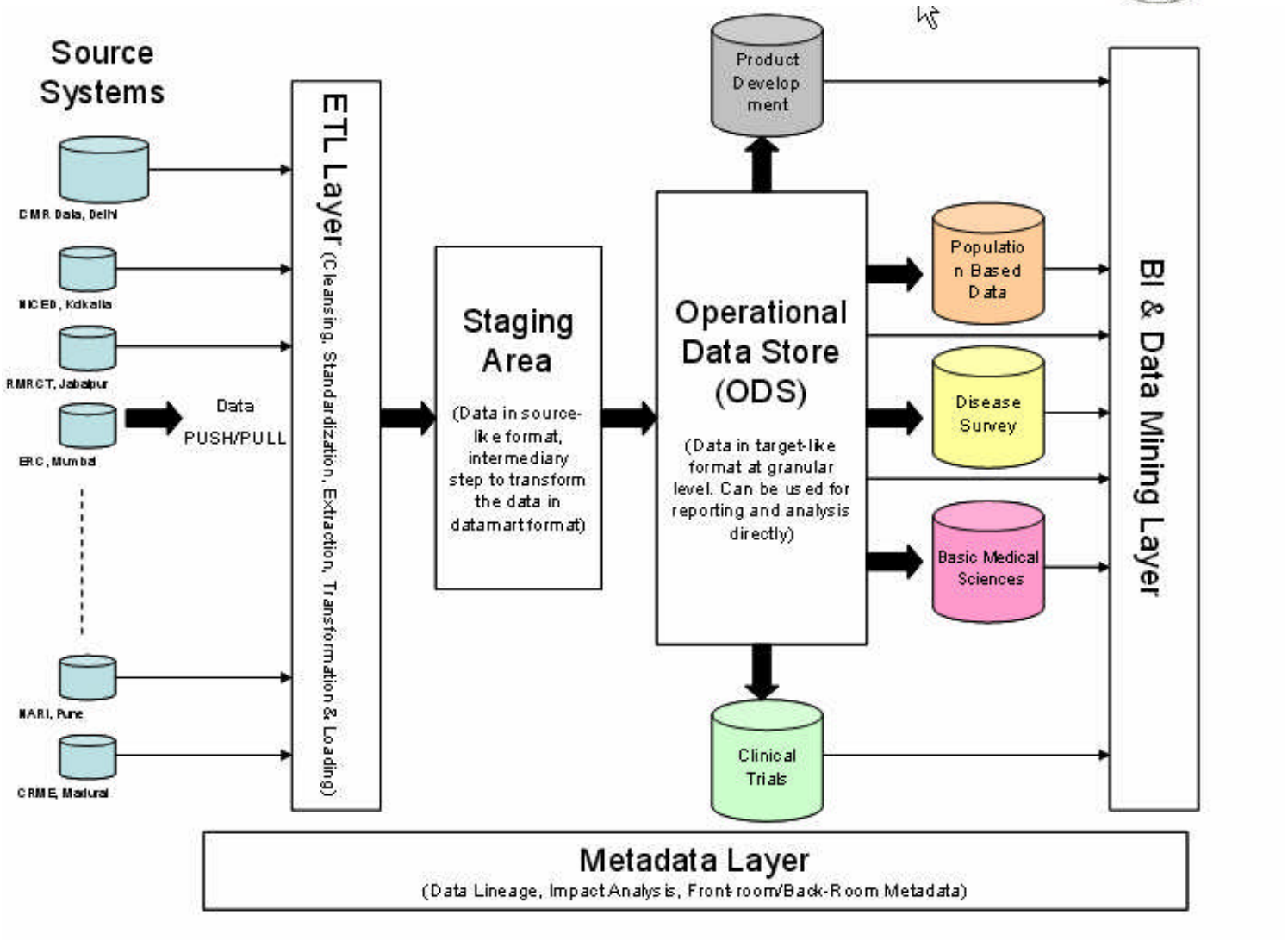
The Compelling need for a Data Repository

- **The need to consolidate data from across institutions/ projects/ reports/ studies that focus on similar disciplines – and make the consolidated data available in a “secured” manner to individual scientists , research Centres etc on a on line basis**

The need to build a Data Warehouse

- given the disparate source systems and their structures, it becomes critical to choose and use technology that streamlines the process of accessing, retrieving, storing, online analyzing and presenting data from all existing systems
- to evolve a system using state of art technology for the benefit of the scientific community across ICMR using Business Intelligence (Data warehousing, OLAP, on line reporting & Data Mining).
- The text mining for the extramural research proposals will help the Project Review Committees to better evaluate projects and eliminate duplicity

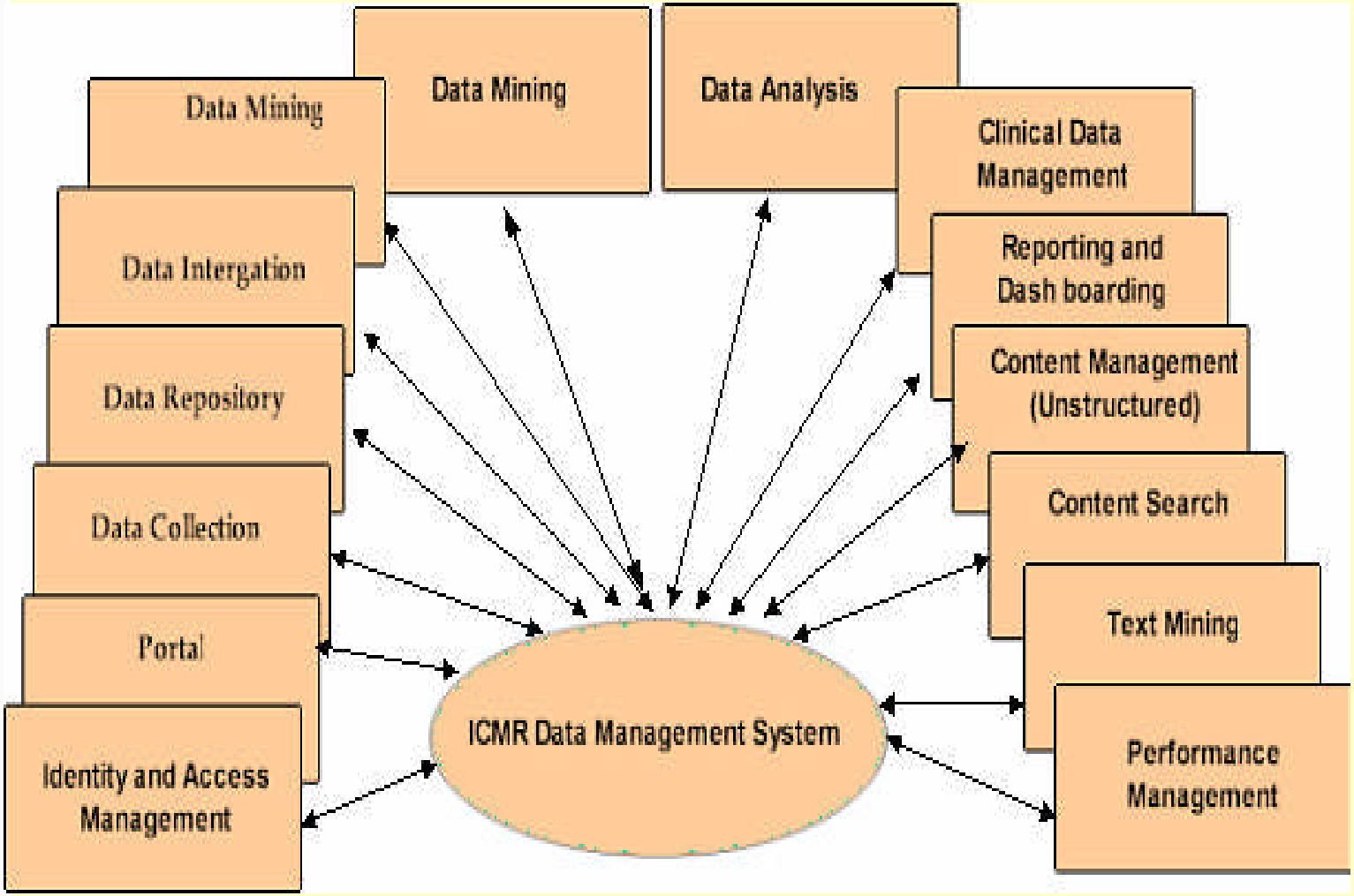
Conceptual Architecture



Objectives of the Project

- To document and archive all research done under the aegis of ICMR for future use to prevent loss of research data
- To create repository of research data as well as published material in order to enable systematic retrieval of information when required.
- To promote ICMR's mission for health research and disease prevention by providing possibilities of intensive analysis of research information to research and public health community
- To promote information-driven decision making by placing timely, useful facts in the hands of public health practitioners and researchers
- To provide the general public with access to specific and detailed information from the ICMR repository (in the form of metadata)
- To develop electronic data capture modules for the various sites involved in collecting the clinical data across the country.

Key Business Services Covered in the project



Benefits of the Project

- A wide range of BI tools will be available to the researcher to submit and analyze his/her own data in a much more meaningful way; this will lead to more publications for the scientist.
- Duplicate and repetitive research will be reduced, thus improving the overall quality of research
- It will assist in decision making, priority setting, program evaluation and resource allocation as also assessing the quality of research done
- The entire workflow of preparatory research, through regulatory approval, to obtaining a data set will drop from months to weeks
- The data warehouse will allow healthcare professionals to work more intelligently, speeding the development of treatments for disease
- It will provide a massive storehouse of clinical information, procedures and research, enabling rapid analysis and reporting to foster best-practice care
- It will enable extensive, diverse, medical information to be used as the basis for medical research, treatment and life-saving breakthroughs
- It will link the clinical data of all the sites to the Central data repository at ICMR and facilitate monitoring of the data and workflow at Central level to react quickly when changes are made as well as data validation at the site level.
- It will enable clinical trial managers and research team members of ICMR to allocate resources to ensure the success of their respective clinical trial projects.

Overall scope of work of DR & BI project

- **Development, implementation and rollout of data Repository Portal**
- **Pilot Study and Road Shows**
- **Preparation of SRS and its modification on basis of Pilot Study**
- **Creation of Data Repository and populating it with ICMR data**
- **Resolution of Data Quality Issues**
- **Creation of Data Marts/Warehouse**
- **Business Intelligence Reports, Dashboards, OLAP Analysis, Data Mining**
- **Document Management and Text Mining**
- **Clinical Data Management**
- **Implementation, training and support**

Development, implementation and rollout of data repository portal

- **Design and Graphical User Interface (GUI) of the portal**
- **Displaying the contents**
- **Providing interactive interface for depositing and accessing data**
- **Management of content including naming convention, file structure etc.**
- **Providing state of art search facility**
- **Running applications/services through portal**
- **Providing enhanced security at different levels through authorization**
- **Preparing user and technical manuals of the portal**
- **Data must not be susceptible to data-theft by hackers, the data must be available to the authenticated users at the right time, and the system must keep a record of activities performed by them.**

Pilot study

- Create prototype covering full range of data, using data from at least three institutes/headquarters on minimum four of seven data types.
- Validate application requirements
- Determine the core functions of the system (the functions) that should be part of the prototype
- Design and build preliminary database
- Build temporary structures
- Identify and develop key modules and functions
- Design and develop core functions
- Find missing pieces & discrepancies in the requirements
- Test selected DBMS and other tools for efficiency
- Create dashboards for viewing results

Preparation of SRS and its modification on basis of pilot study

- Gather information
- Classify the information
- Identify scope and out of scope requirements
- Define and document business rules
- Prepare top-down model representation of major processes
- Prepare logical model by integrating data that support processes and business rules.
- Reconcile business requirements with models
- Merge interfaces, processes and data to describe systematically how ICMR will use the application and how data will be retrieved processed and stored
- Collect and collate feedback from road shows of Pilot Study and modify SRS

Thank You