



Preserving, Sharing & Re-using Biomedical Research Data: Policy & Practice

**DCC-DPC Workshop
3-4 July 2006**

**MRC Data Sharing & Preservation Web Pages:
http://www.mrc.ac.uk/strategy-data_sharing.htm**

Why Preserve or Share Research Data?



- **Reduce Duplication**
 - Effort of Data Creation and Analyses
- **Secondary Analysis**
 - Comparative & Confirmatory analysis
 - Testing of new & alternative hypotheses, tools & methods
 - Meta-analysis e.g. systematic analysis of clinical trials data
- **Data enrichment**
 - Linking datasets
 - Large-scale pooling for complex or comprehensive analyses
- **Asset management**

Asset Management



- **Assume nothing lasts**
 - Fragility and corruptibility of storage media
- **Obsolescence**
 - Storage media & hardware to access data
 - Dependence on proprietary or unsupported software
 - Unsupported file formats
- **Storage and Management**
 - Costs e.g. format migration, metadata creation, storage
 - Space e.g. physical (paper), digital (server)
 - Management of deposition, access and use = governance
 - Ease of use: preservation vs. archiving vs curation
- **Good practice**
 - 10 & 20 year good practice for (non-)clinical data storage

International Policy



- **Organisation for Economic Co-operation and Development (OECD):**

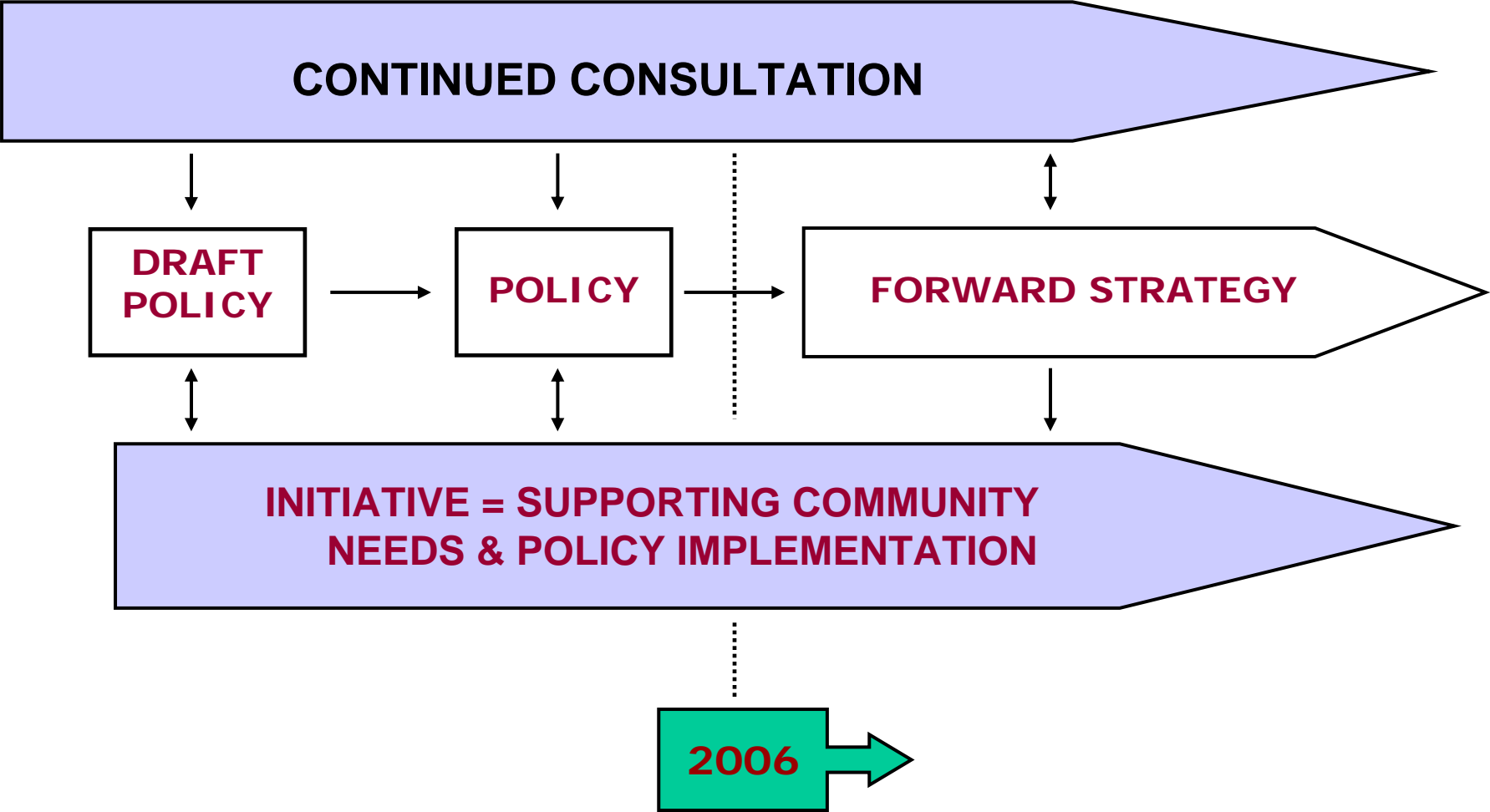
“Publicly-funded research data are a public, produced in the public interest and should be openly available to the maximum extent possible”

- **National Institutes of Health (NIH) Policy:**

“Data should be made as widely and freely available as possible while safeguarding the privacy of participants and protecting confidential and proprietary data”

Data sharing is essential for expedited translation of research results into knowledge, products and procedures to improve human health

Initiative Overview



Attitudes & Practices...



- **Wide-Scale Survey (2001)**
 - MRC intra – and extramural research
 - Review of key data assets & DSP practices; views on future policy
- **Case Studies (2002)**
 - Conducted by UK Data Archive
 - 18 data collections: epidemiology /cohorts/clinical trials
 - Informed draft policy and implementation strategy
- **Joint Data Standards Study (2005)**
 - Large-scale data sharing across life sciences
 - Sponsors: MRC, BBSRC, NERC, Wellcome Trust, JISC, DTI
 - http://www.mrc.ac.uk/pdf-jdss_final_report.pdf
- **Joint MRC- WT Data Access Project (2006)**
 - Review of access arrangements across co-funded data & material collections
 - http://www.mrc.ac.uk/pdf-access_report_march_2006.pdf
 - Next steps – identifying principles for governance of access

Attitudes & Practices...



- **Different communities, different:**
 - Attitudes
 - Awareness
 - Considerations & Needs
 - Knowledge
- **Early surveys - little attention to data preservation:**
 - beyond primary data
 - for sharing beyond research team
- **Practical issues : migration; resources; standards & tools**
- **Ethos and Awareness Challenge**
PI-led research -“my data!” “where do I start?”

Quality matters



➤ **Creator/Custodian concerns**

- **Quality of secondary research**
 - inappropriate or unethical associations
 - data raiding & devaluation
- **Recognition & rewards**
 - reputation – personal; the data resource
 - timing
 - intellectual investment

➤ **User concerns**

- **Data utility**
 - Quality control of (meta) data - integrity
 - Documented curation 'history' – provenance
 - Access to data

MRC Data Preservation for Sharing Policy



- **General principles that apply to all MRC-supported research**
- **Promotes DSP as an essential part of good research practice**
- **Promotes sharing of data for new research purposes to bring added value**
- **Balances interests of data creators and other users**

Some Policy Principles...



- **Data creators and custodians have a duty to make publicly-funded research data available in a responsible manner for further research;**
- **Data providers should legitimately benefit from their intellectual investment and effort;**
- **Prolonged exclusive use of data, if not justified, is not in the interests of scientific advancement**
- **Robust arrangements are in place to manage any risks identified relating to IP, consent & confidentiality**

More Policy Principles...



- **Do not prescribe when or how researchers should preserve and share data**
- **Responsibility of investigators to propose a costed preservation and sharing plan – including an access policy – which will be reviewed at the time of funding requests and renewal**
- **Provide explicit reasons for not sharing data**
- **Applies to all proposals from 1 Jan 2006**
- **Focus initially on funding proposals to support large-scale data collection and use**

Policy Support: Current Work Streams



Engaging with wide range of stakeholders to...

- **Develop practical guidance to support MRC researchers in the planning and execution of their data curation activities**
- **Identify costed options for long-term preservation for sharing to support major population data assets**
- **Deliver a 'route map' through current processes regulating use of personal data for medical research**
- **Commission research into public awareness of, and attitudes to, medical research using personal data**
- **Working with funding partners towards joint policy and governance principles**

Good Practice Guides to Data Curation



- **Three web-based guides responsive to MRC community needs**

Guide 1

Development & Assessment of researchers' DSP plans

Guide 2

Data & Metadata Management for Population-based Medical Research

Guide 3

Strategies & Solutions for Data Preservation for Sharing

- **Target audiences involved in the shaping and testing of guide content**

Cohort Support Project



Initial focus: 2 large-scale cohort studies

Avon Longitudinal Study of Parents & Children (ALSPAC)

<http://www.alspac.bristol.ac.uk>

National Survey of Health and Development (NSHD)

<http://www.nshd.mrc.ac.uk>

- **Assess current & prospective (meta)data assets**
- **Costed options for enhanced preservation for wider access and re-use through sharing**

Bronze → **Silver** → **Gold**

- **Inform data management plans**

Cohort Support Project



"SILVER"

- **Focus on Enabling Wider Sharing & Use**
- **Up to 3 years to implement**
- **Main Features:**
 - **Preservation strategy**
 - **Complete metadata & documentation**
 - **Data catalogue & index**
 - **Data & metadata in suitable release formats**
 - **Governance in place - access and secondary use**
 - **Significant Visibility - enhanced online information**
 - **Data available on-line for "bona fide" users**
 - **Basic but dedicated user support functions & resource**

Route Map:

Personal Data and Human Tissue in Medical Research



•Goals

Visual guide through current ethical, legal and institutional processes regulating use of personal data and human tissue for research

Promote principles of good practice supported with real-life exemplars

•Outputs

Easy-to-Navigate Information Resource e.g. <http://www.ct-toolkit.ac.uk>

Real-life examples and case studies illustrating key principles

Annotated links to information sources – ‘informative pointers’

‘How To’ Material e.g. downloadable pdfs from regulatory sites

[Home](#)[Route Maps](#)[About this site](#)[Study Scenarios](#)[What's New](#)[FAQs](#)[Glossary](#)[Useful Info](#)

Search:

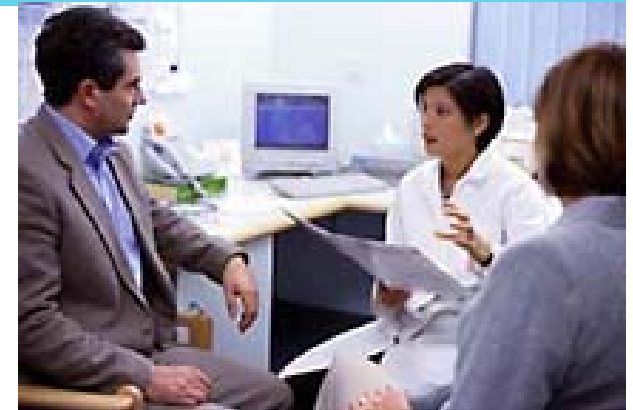
Welcome to the Data and Tissues Tool Kit

On this site you will find practical help with implementation of the Data Protection Act (1998), Section 60 of the Health & Social Care Act (2001), the common law of confidentiality, the Human Tissue Act (2004) and Human Tissue Authority codes of practice, and all associated guidance.

The site has been developed primarily for health researchers and research managers working with personal data and human tissue samples, in the academic sector, but will also be of use to other health professionals.

To help you to navigate the regulatory environment, much of the information is organised in Route Maps. Your feedback on the content and presentation will help us improve the site.

The site has been developed by the UK Medical Research Council for use in all academic research studies. Further Info...



Access Toolkit Bibliography

Find links to relevant documents on the web, plus specially commissioned pieces to help you in the approvals process.



Provide feedback on Toolkit

This is the first draft of the Toolkit. We need feedback on which areas are useful and which are not – and where you may need more help.

First time on this site? Visit the [Getting Started](#) page

Ready to use a Route Map?

Planning & Funding



Seeking approvals



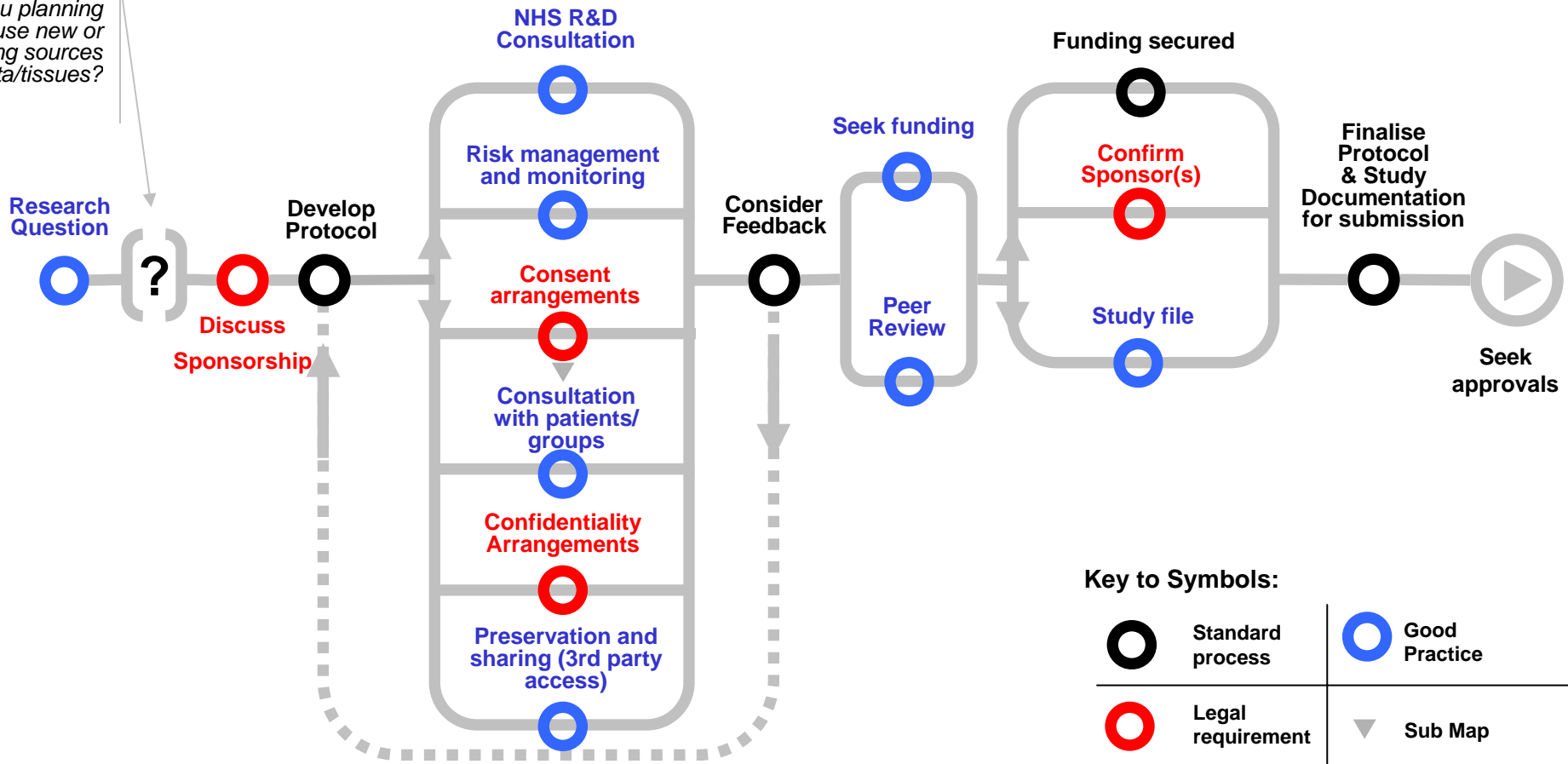
Study initiation & management



Study closure & archiving



re you planning to use new or existing sources of data/tissues?



Public Consultation



Exploration of public attitudes to:

- Risks and Benefits
- “Necessary & Proportionate” use of health information
- Generic consent
- “Consent for Consent”
- Opt-in versus Opt-out
- Role of NHS
- Research Sponsorship

Forward Strategy: Overview



VISION

Long-term objectives for the Initiative underpinned by a set of MRC 'capabilities'

ORGANISATIONAL MODEL

A model to support the capabilities outlined in the Vision

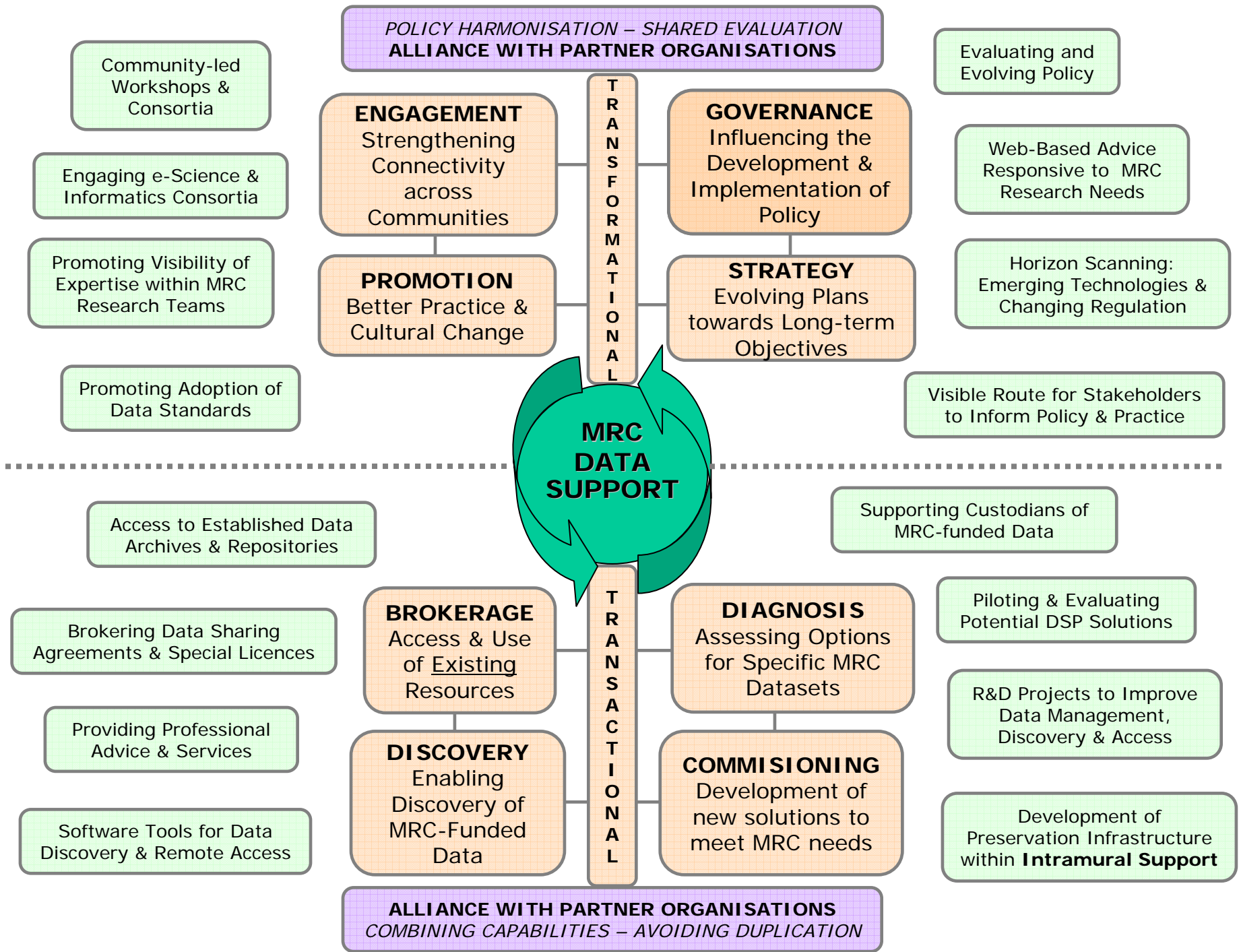
STRATEGIC PLAN

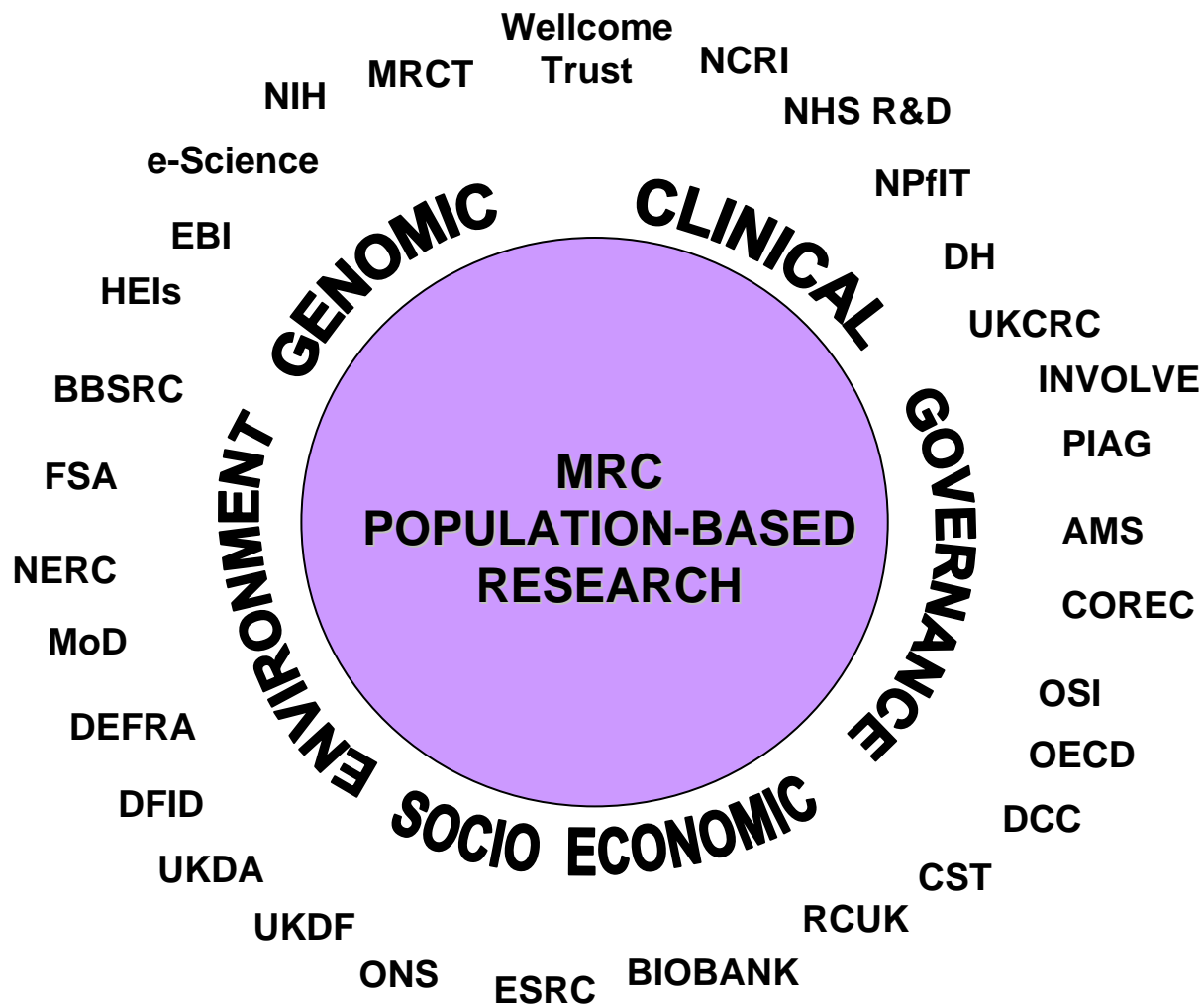
Plan for progressing from current status to achieve model and thereby deliver capabilities

Forward Strategy: **Capabilities**



- 1. Support for discovery of datasets along with metadata & other essential documentation**
- 2. Managed access to, & use of, data for high-quality secondary research**
- 3. Curation to enable informed re-use of preserved data**
- 4. Long-term preservation of high-value MRC-funded datasets**
- 5. Development of tools, standards, guidance & other resources needed to support these activities**





Forward Strategy: Alliance



- **Co-ordinated development and implementation of policy and governance principles**
- **Horizon-scanning informed by awareness of the wider landscape of partner activity**
- **Opportunities for combining capabilities to support practical tools and resources**
- **Shared evaluation of the benefits and costs of prospective data sharing projects.**

Forward Strategy: First Phase Priorities



MRC-funded data resources:

- of unique, long-term value, scientific research benefit been established
- strong research community drive; where added value and costs can reliably be demonstrated

Brokering access to existing preservation infrastructure and data sharing tools where they meet MRC research community needs

Supporting development and maintenance of new infrastructure and tools where strong case but current needs are not met

Building on current Initiative work streams:

- integrated, web-based information resource managed to take account of evolving practices, user feedback, changing regulations
- work on defining principles is supported by resource discovery tools and local governance arrangements

Working with expert partners to:

- support development and adoption of data sharing tools, standards and guidance
- evaluating policies, developing guidance and promoting good practice



QUESTIONS?