# Delivering Effective eResearch Services: The International Situation is Intensifying



International Benchmarking Study for
New Zealand eScience Infrastructure (NeSI)


February 2019

Prepared by Mark Dietrich, in consultation with Nick Jones, NeSI Director, and members of NeSI's Board

BLOODSTONE

# Overview



- Key Research Project Archetypes
- Important eResearch Provider Mindsets
- Mapping the eResearch Value Chain
  - Creating a common framework for discussion
  - Balanced, sustained investment important for resaerch
- Common Organizational Patterns
- Country-by-Country Comparisons
- Key Observations for New Zealand

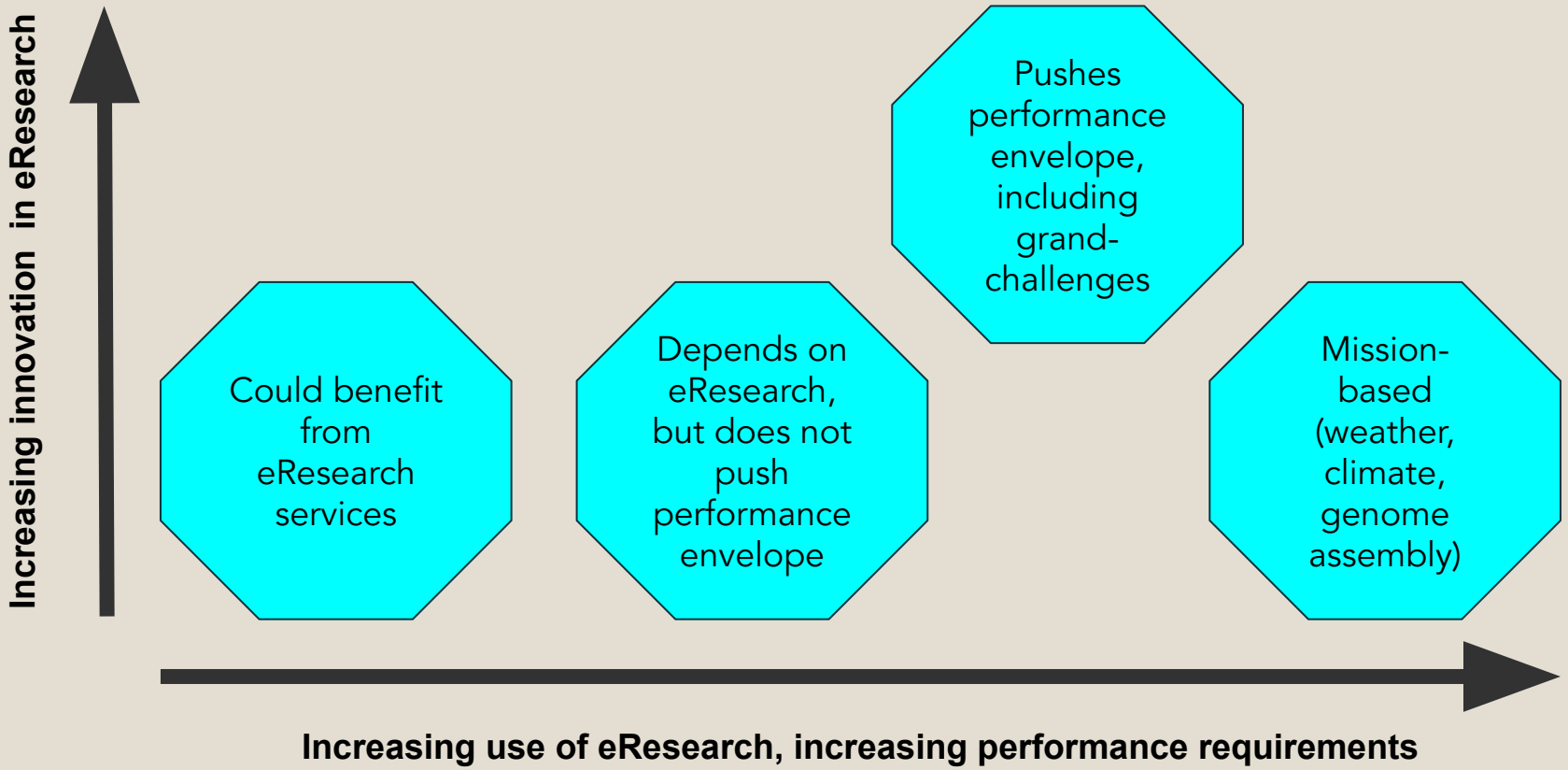
# Spectrum of Research Project Types



BLOODSTONE

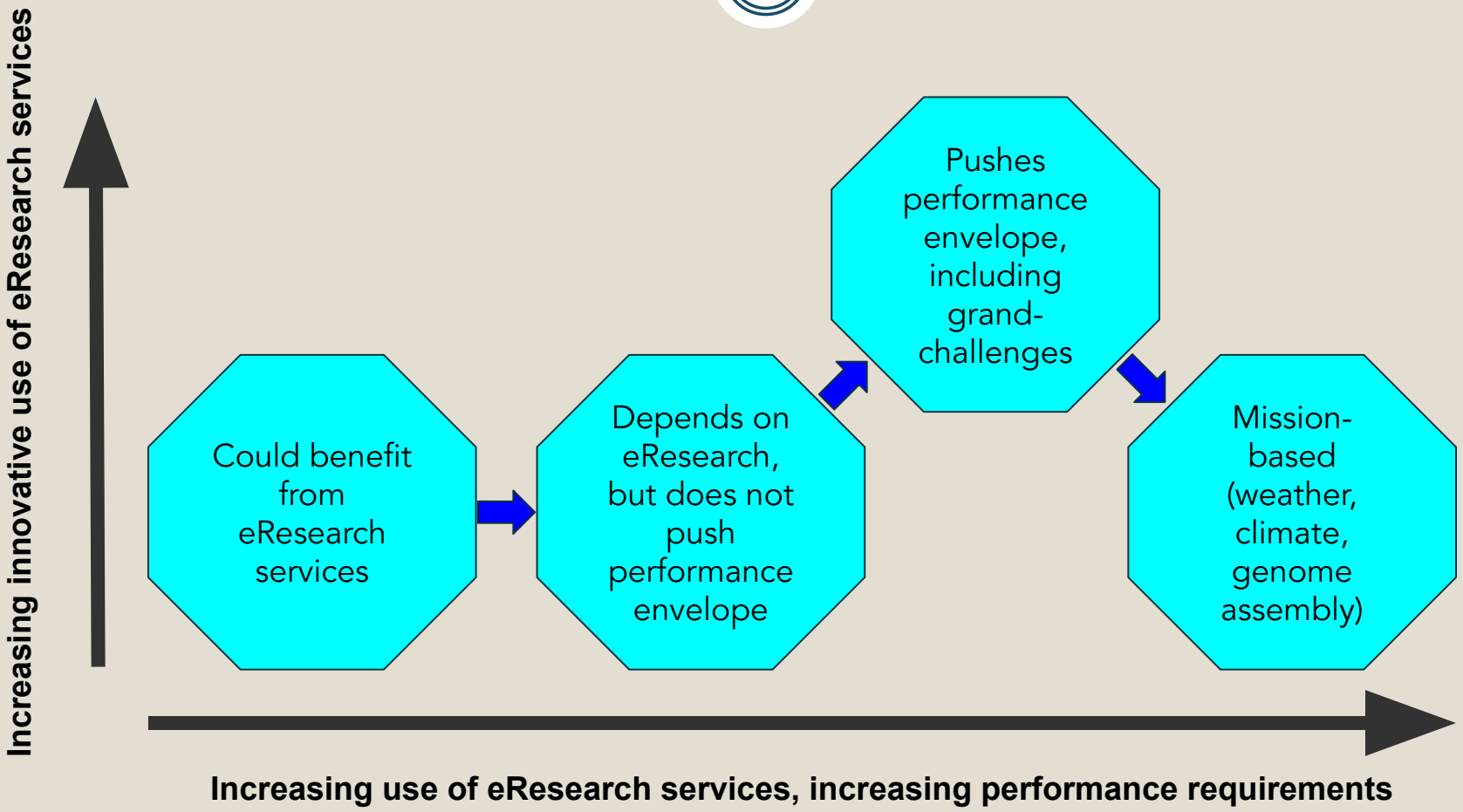
# Understanding (e)Research Project Spectrum illustrates different points where real benefits are possible





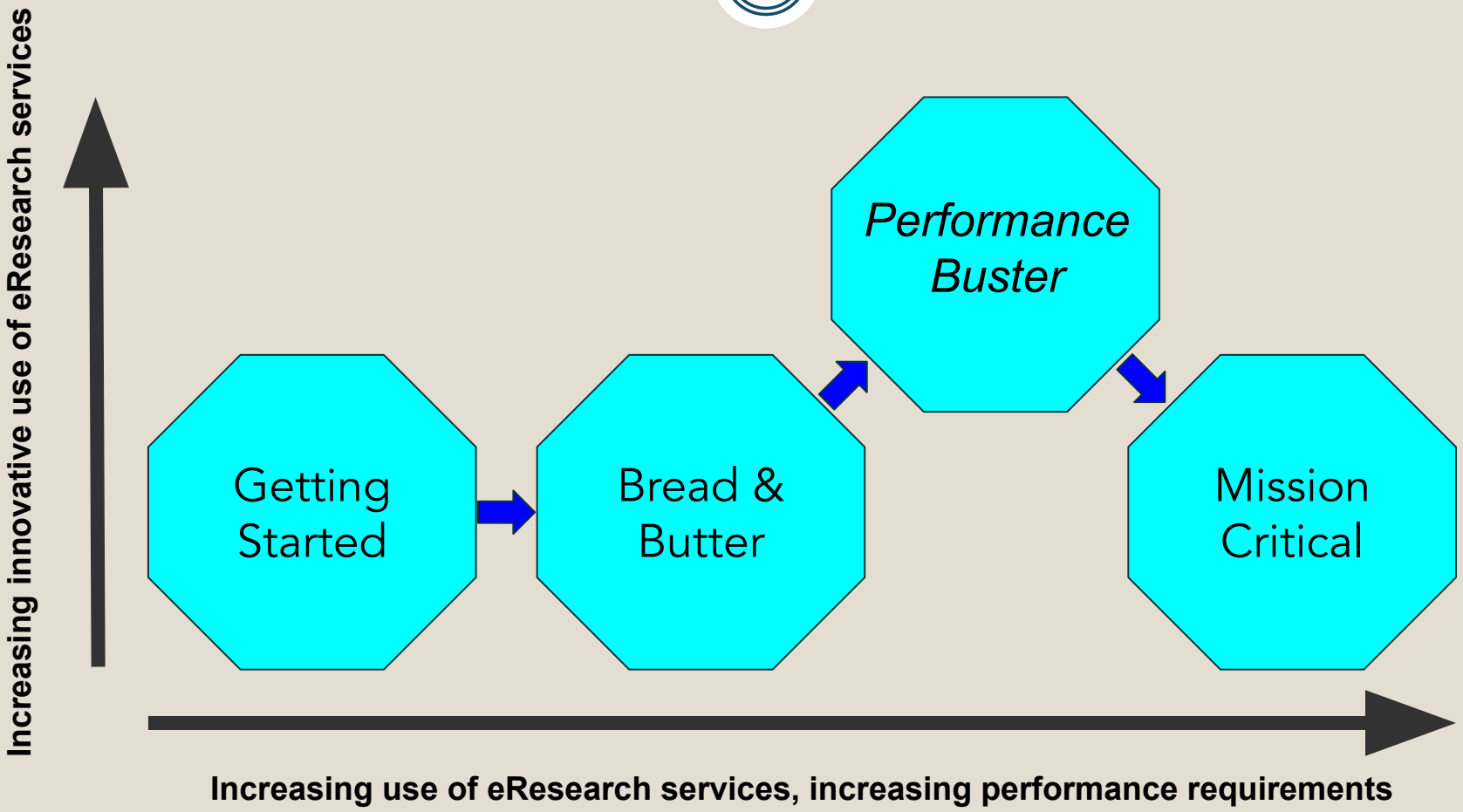
# Over Time, Projects Can Become More Demanding





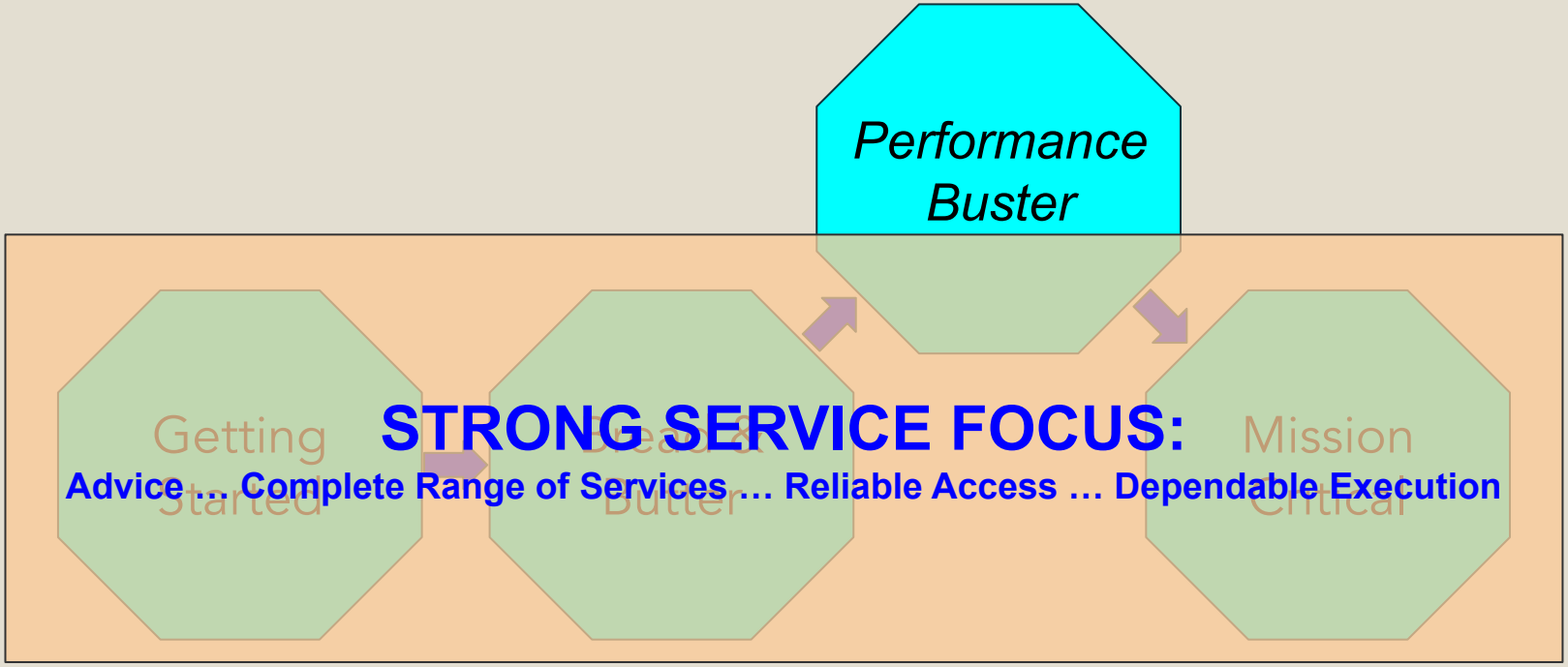
# Shorter Names





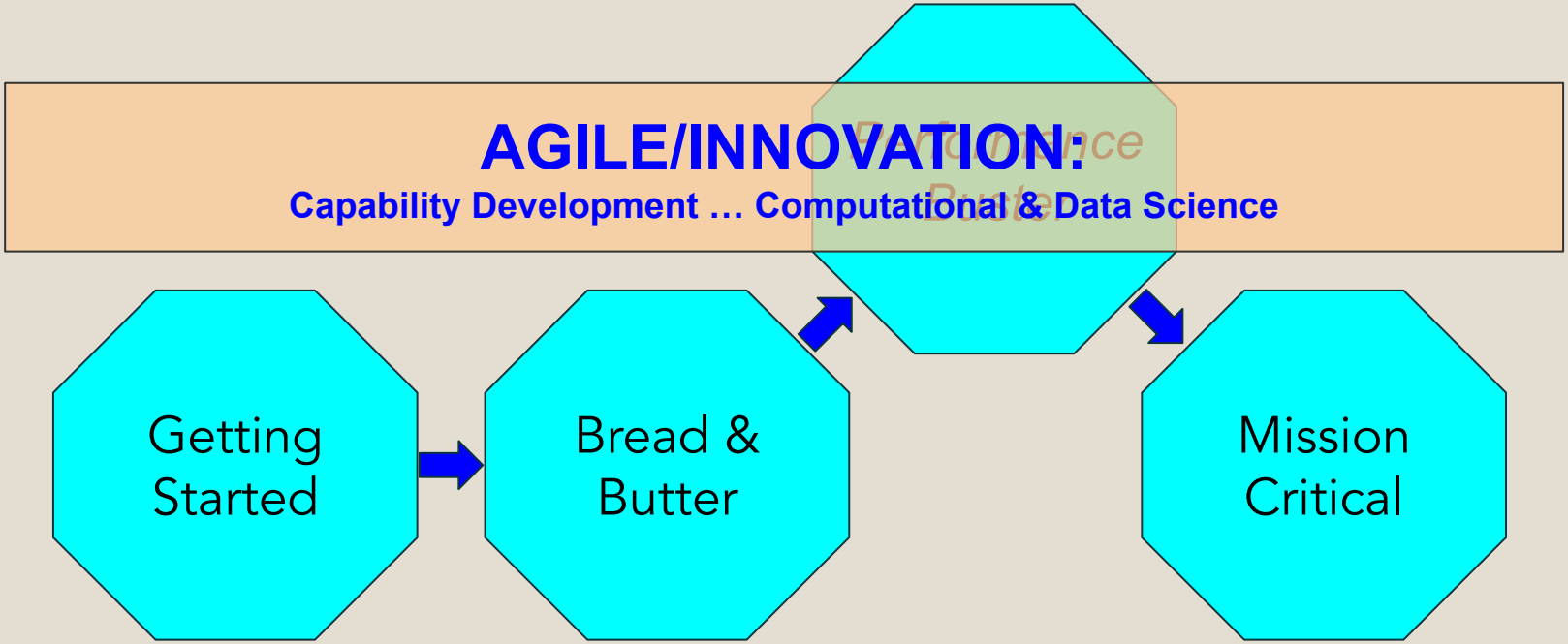
# eResearch Provider Competencies





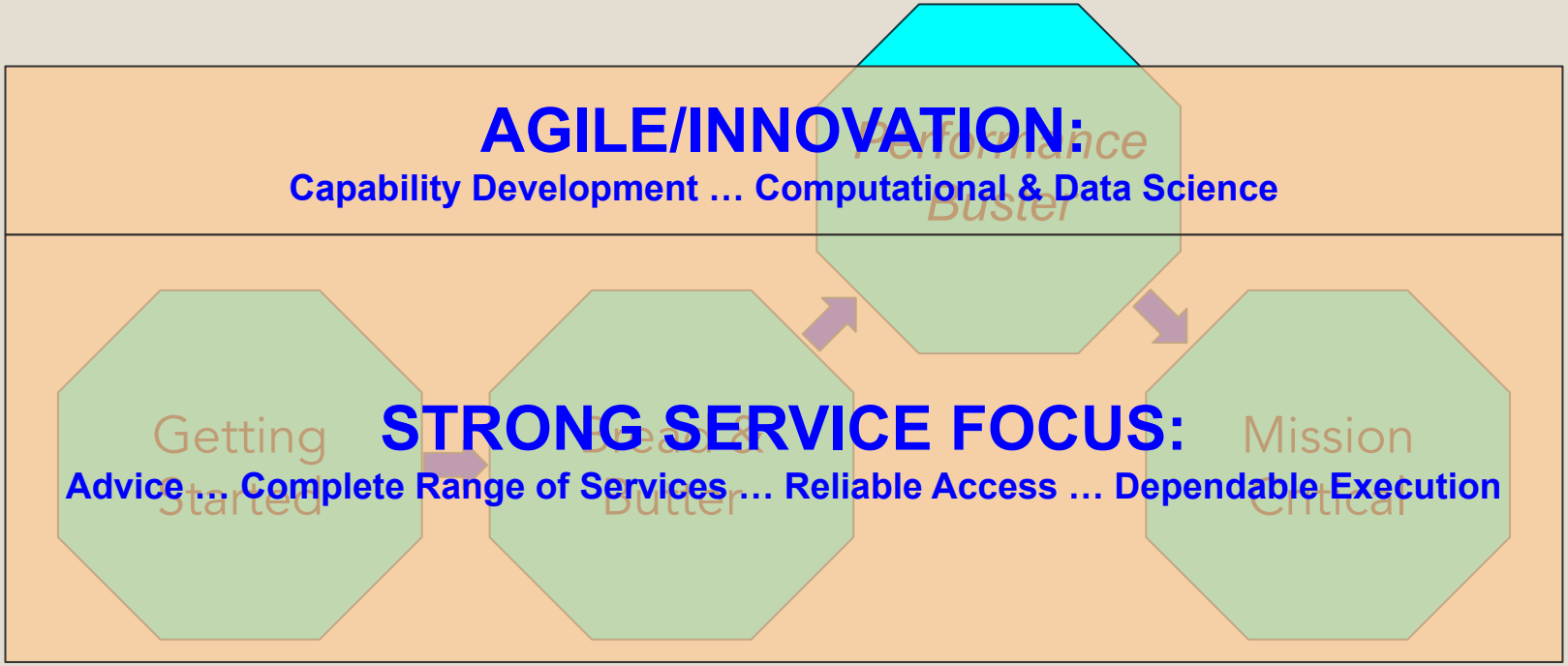
# eResearch Provider Competencies





# eResearch Provider Competencies





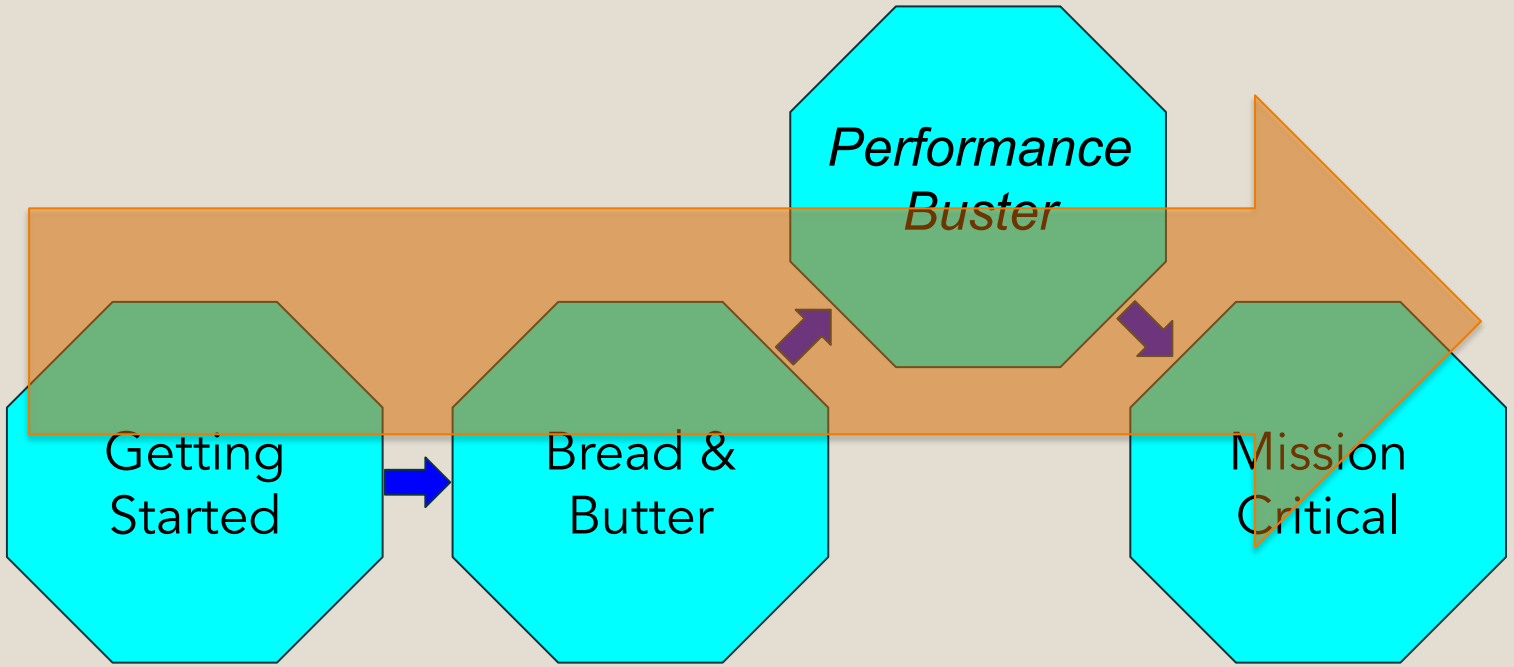
# 2 Possible Dynamics



## *1. Projects Served*

# Overview: Projects Served



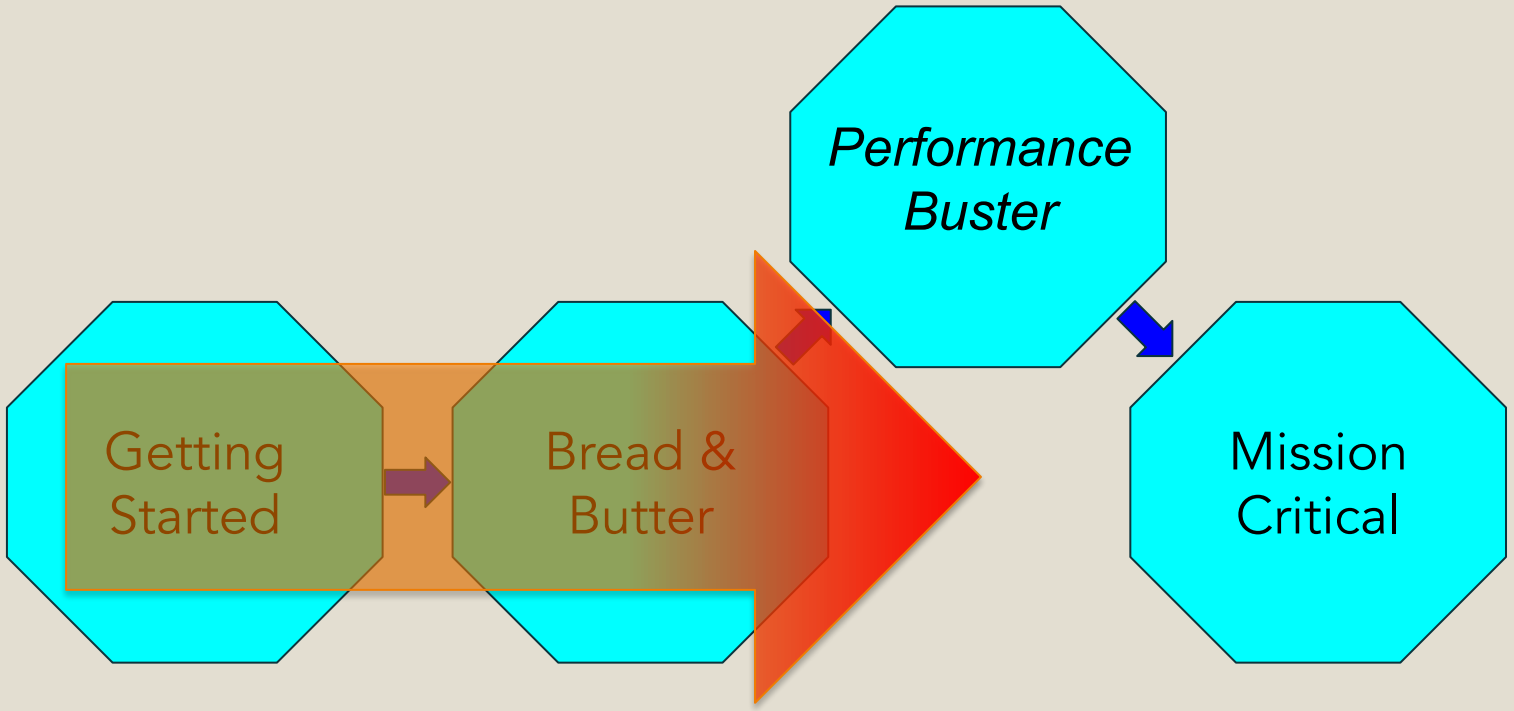


# Project Serving: Objective 1



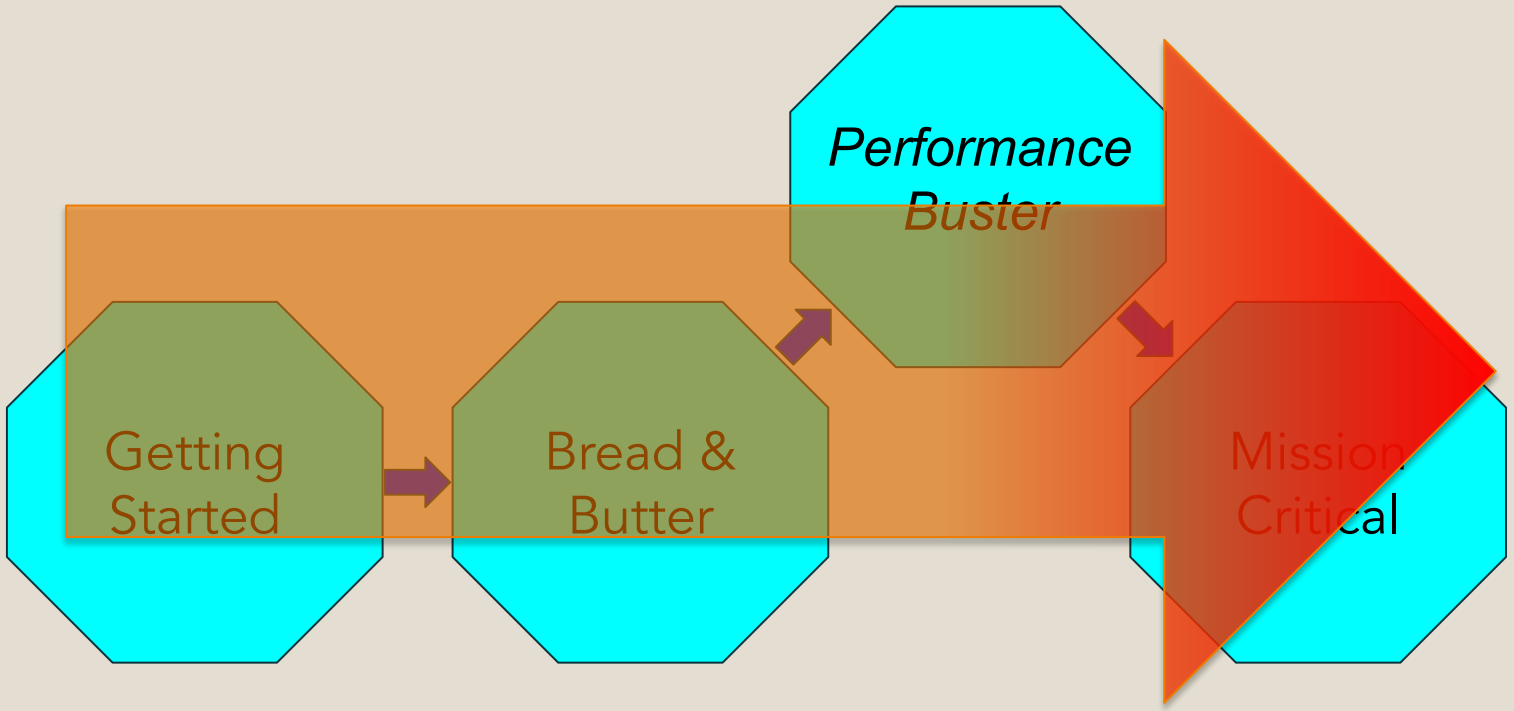
# Problem: Broader Tools do not work well for more performant projects





# Result of Projects Served Dynamic





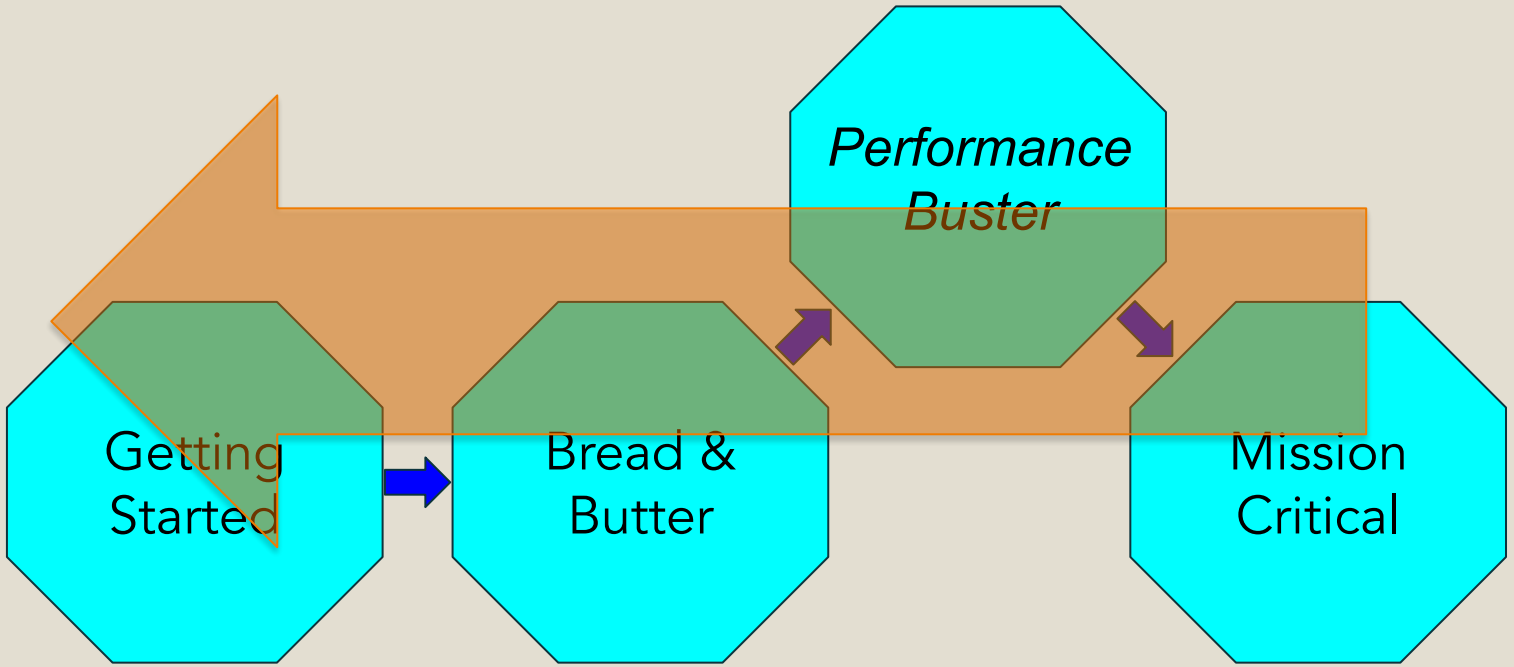
# 2 Possible Dynamics



## *2. Challenge Seeking*

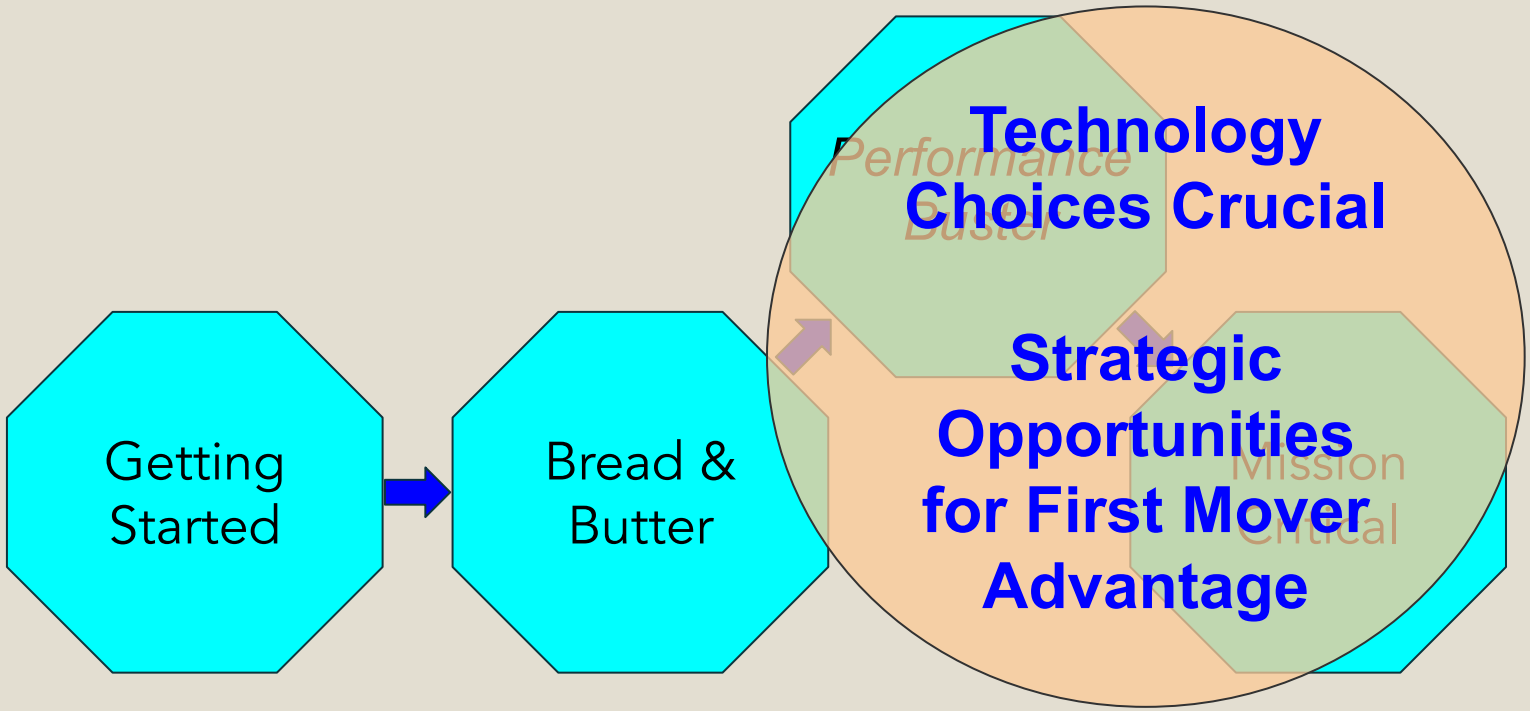
# Overview: Challenge Seeking





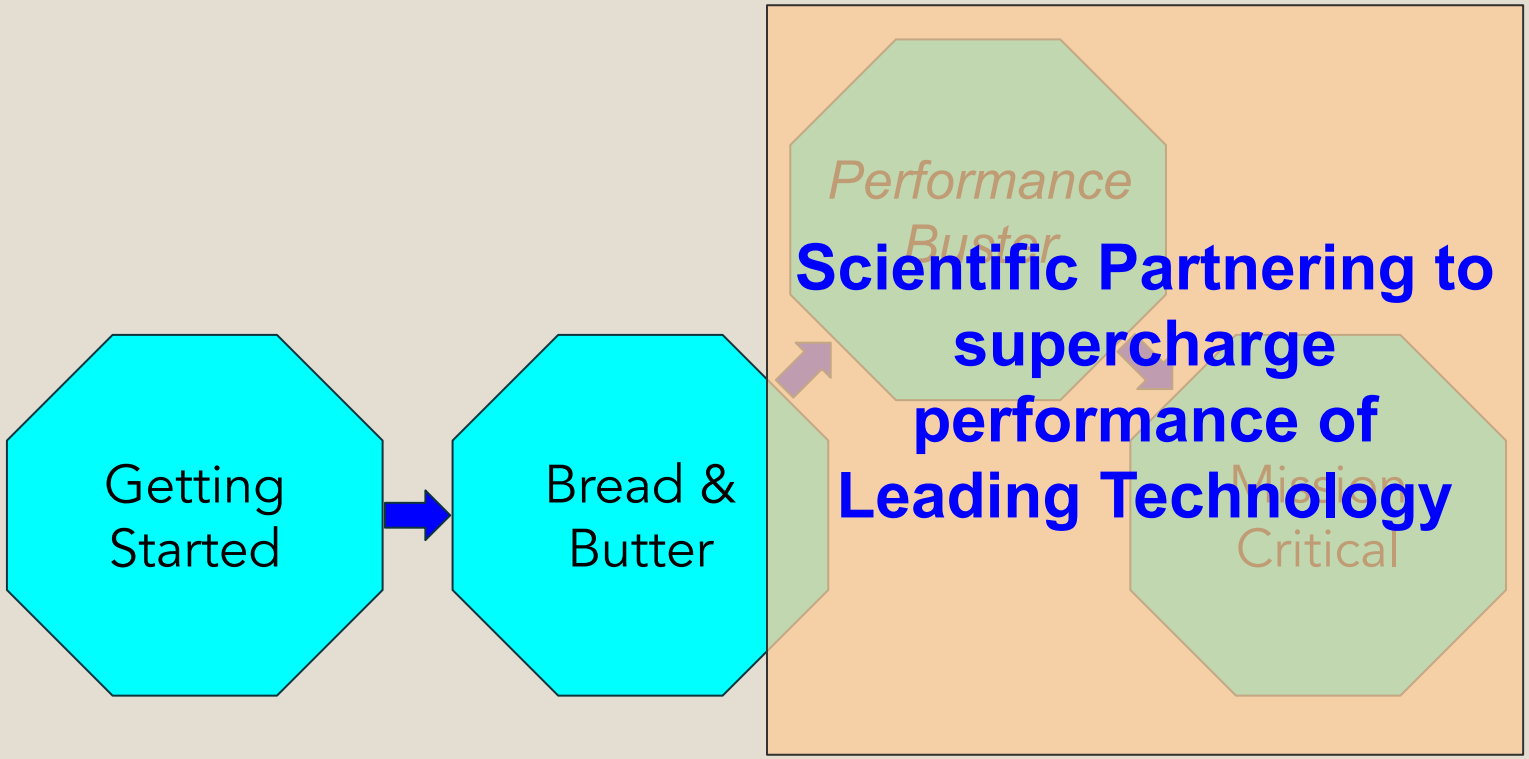
# Challenge Seeking: Objective 1





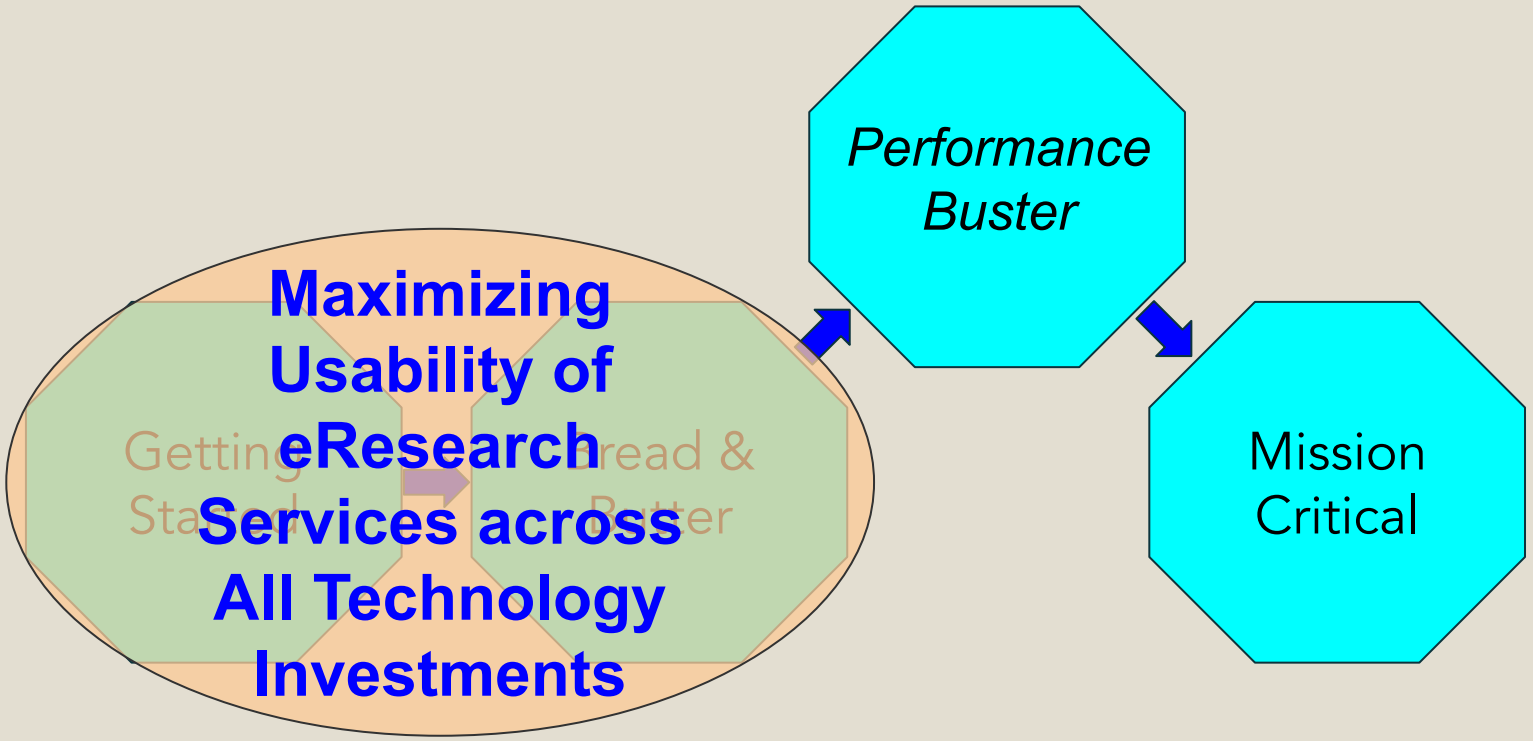
# Competency 1 for "Challenge Seeking"





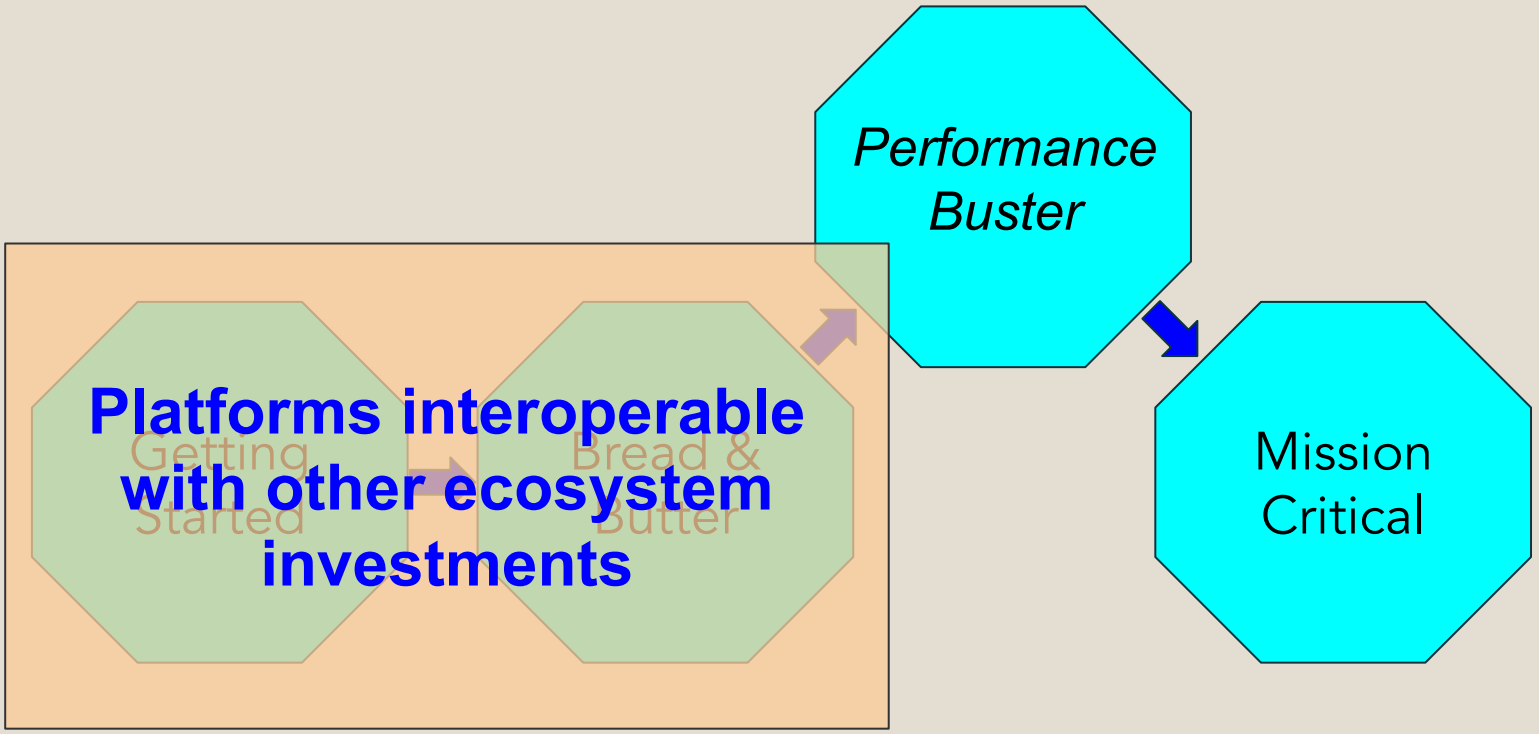
# Challenge Seeking: Objective 2





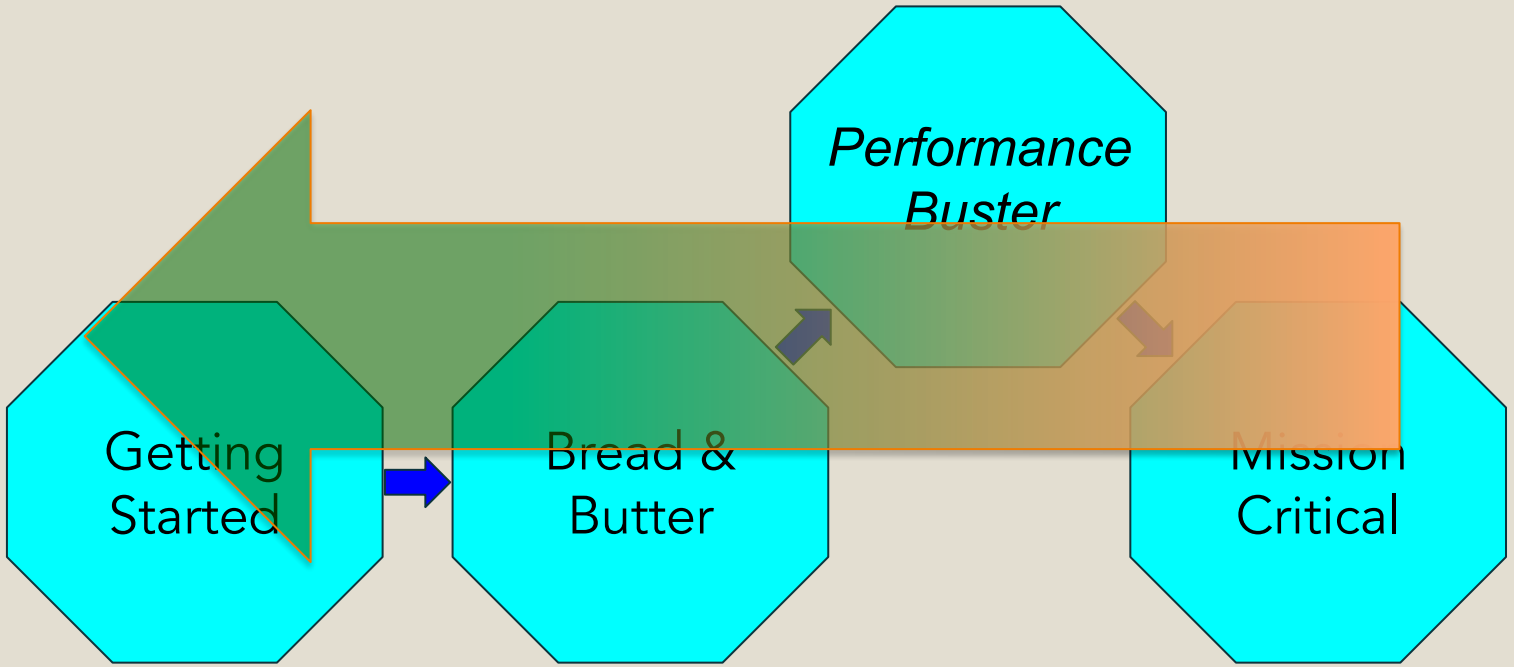
# Competency 2 for "Challenge Seeking"





# Result of Challenge Seeking Dynamic





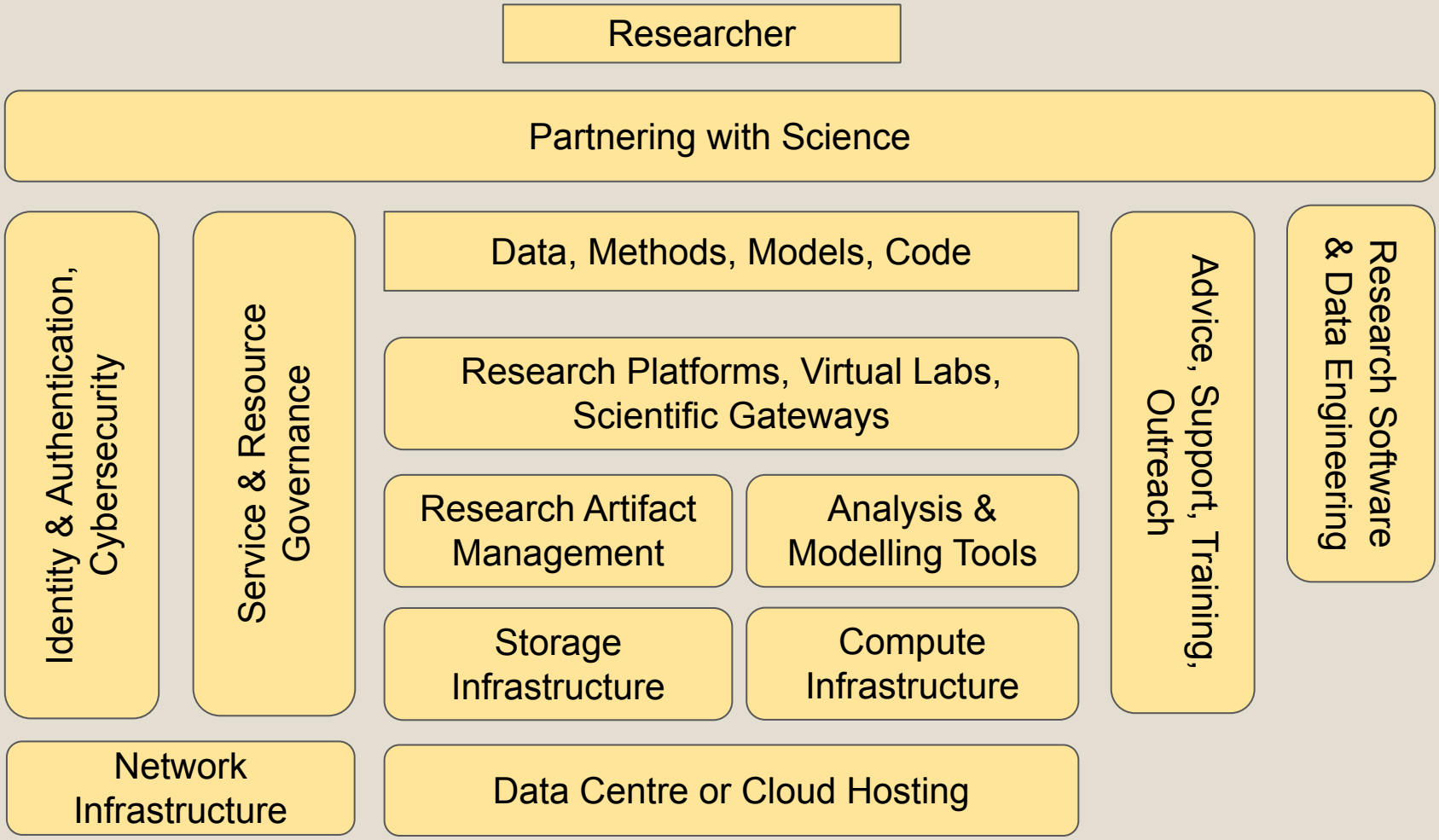
# Mapping the eResearch Value Chain



## Enabling Useful Comparisons Across Jurisdictions

# eResearch Ecosystem Framework (eREF), v4





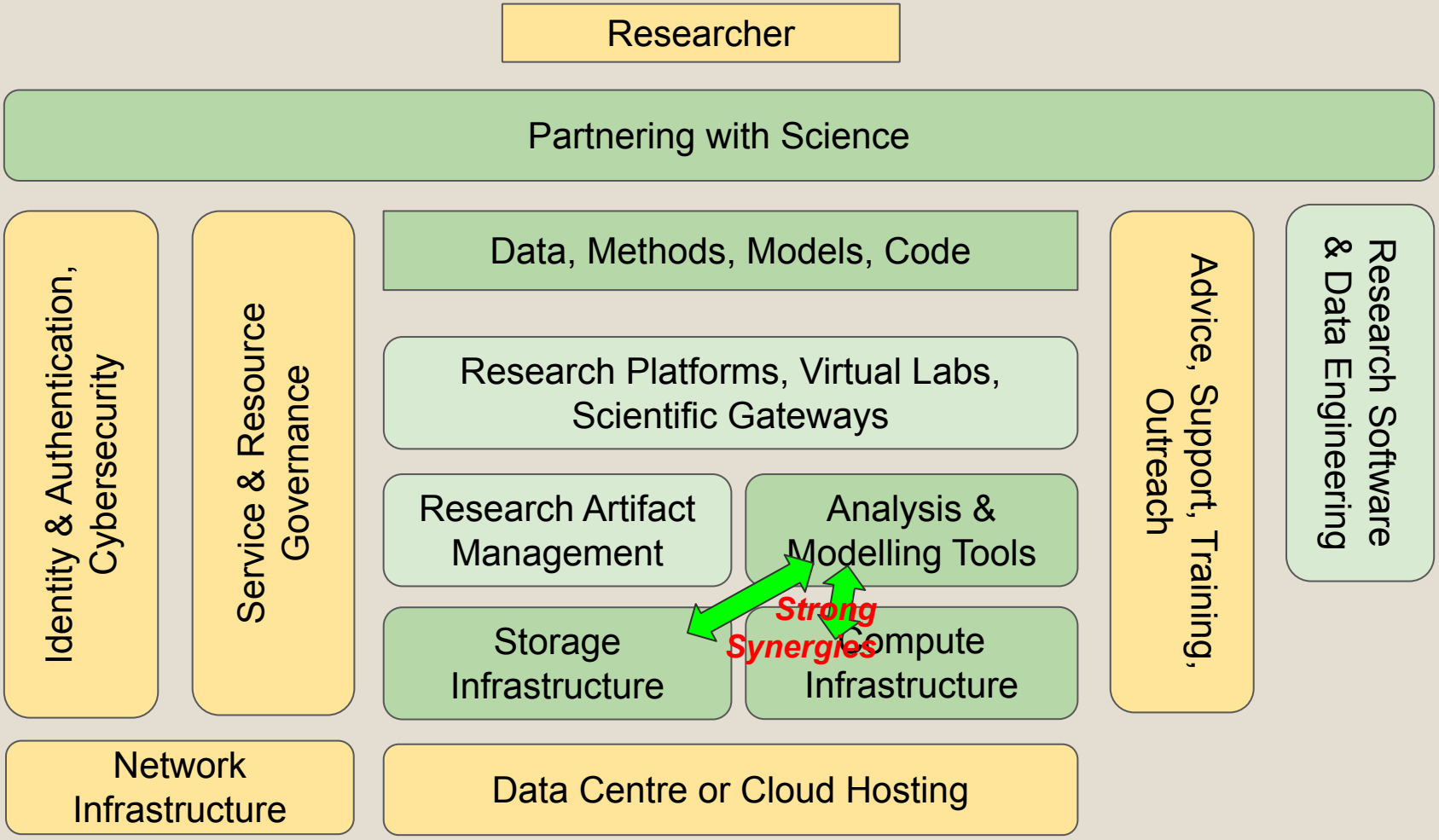
# Key Trends in Ecosystem Components



- Continued Evolution within each Component
  - To be expected since the "customers" (researchers) are themselves always engaged in new activities
- Significant Dependencies between Components
  - E.g. File Replication
    - Provided by Storage Infrastructure or Research Artifact Mgmt...?
    - ...or by a Research Platform, since replication might already be achieved with the copies stored locally by different platform users
- While apparently "trivial" these factors complicate economies of scale (potentially achieved through consolidation) and scope (potentially achieved through integrated service delivery)
- At the same time, there is clear value in scientific/technical competency that spans the full range of components. This competency, if kept current, can ensure exploitation of whatever economies of scale or scope might be available.

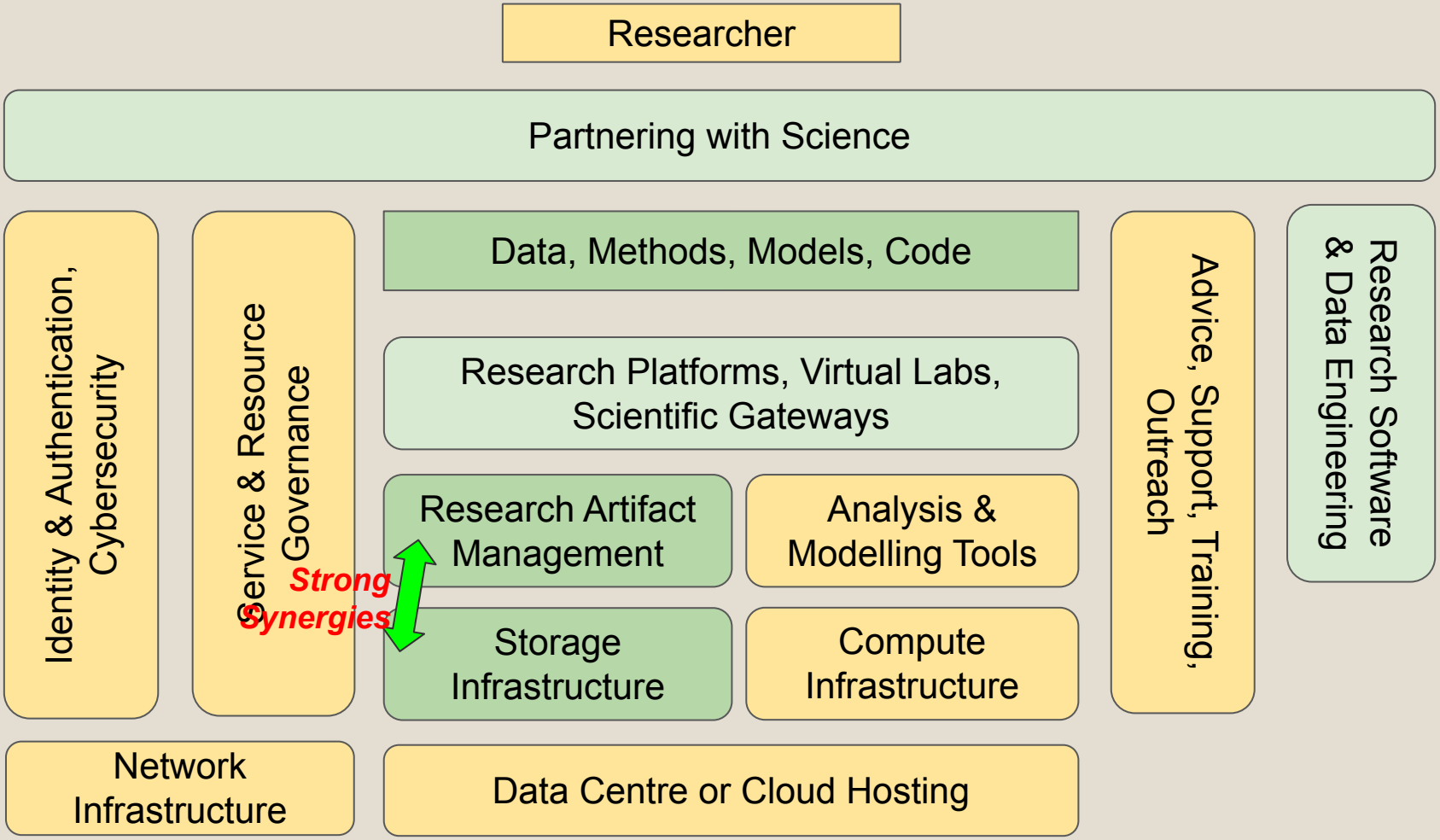
# Computational Science & Data Analytics Build on the Same Foundations





# Research Artifact Mgmt: Synergies with Storage





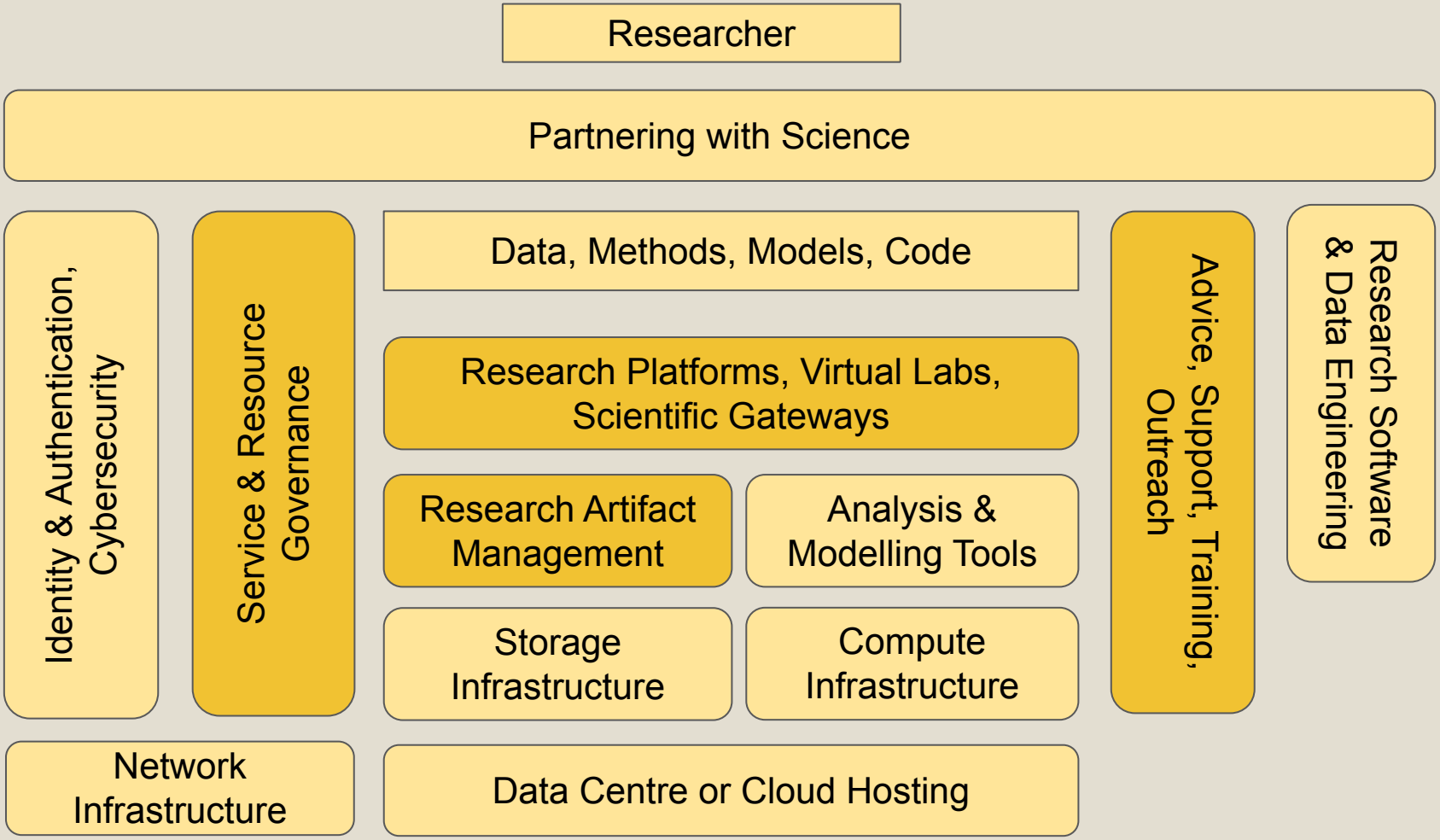
# Components Required by Different Research Project Types



BLOODSTONE

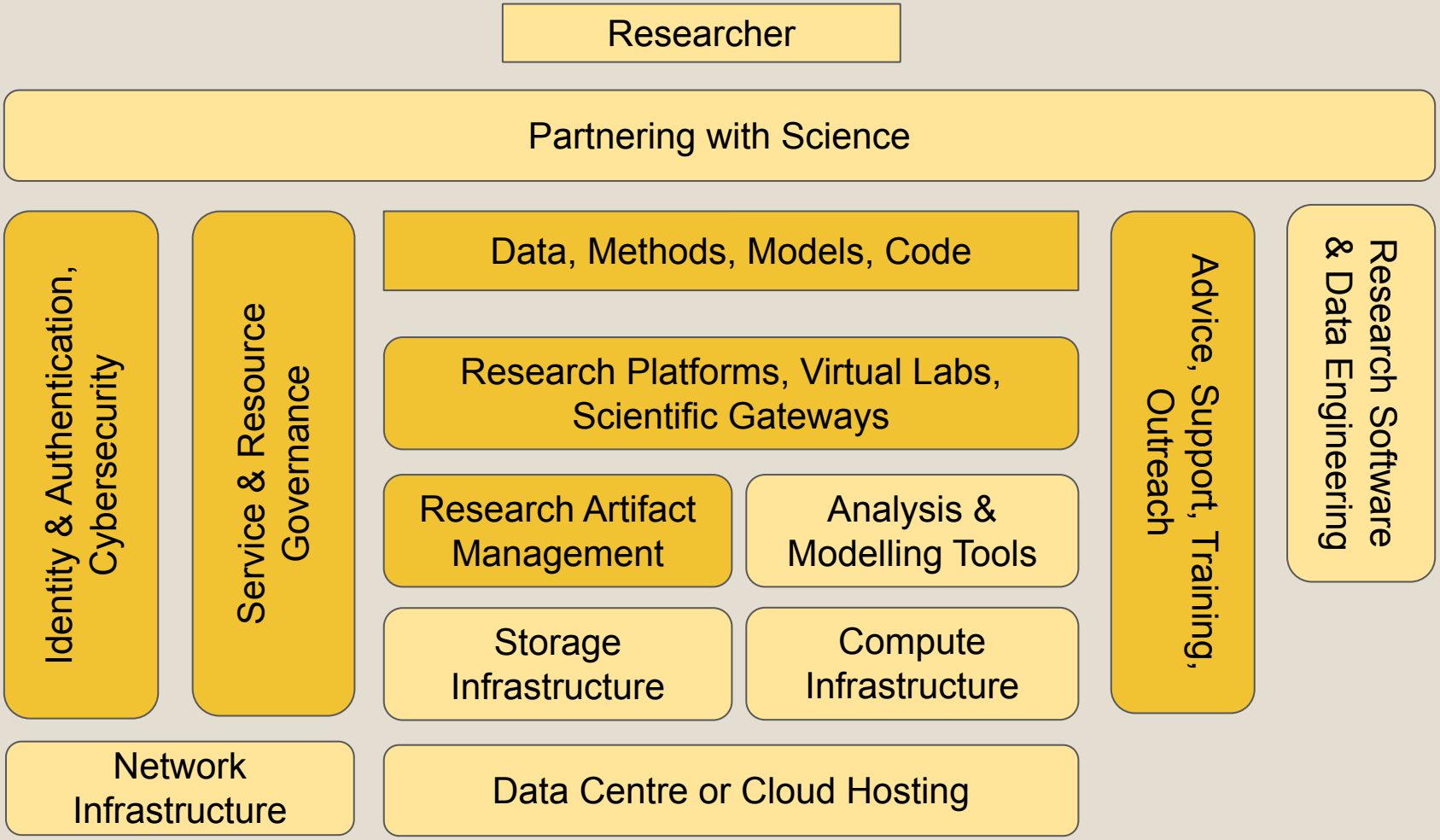
# Key eREF Components: Getting Started





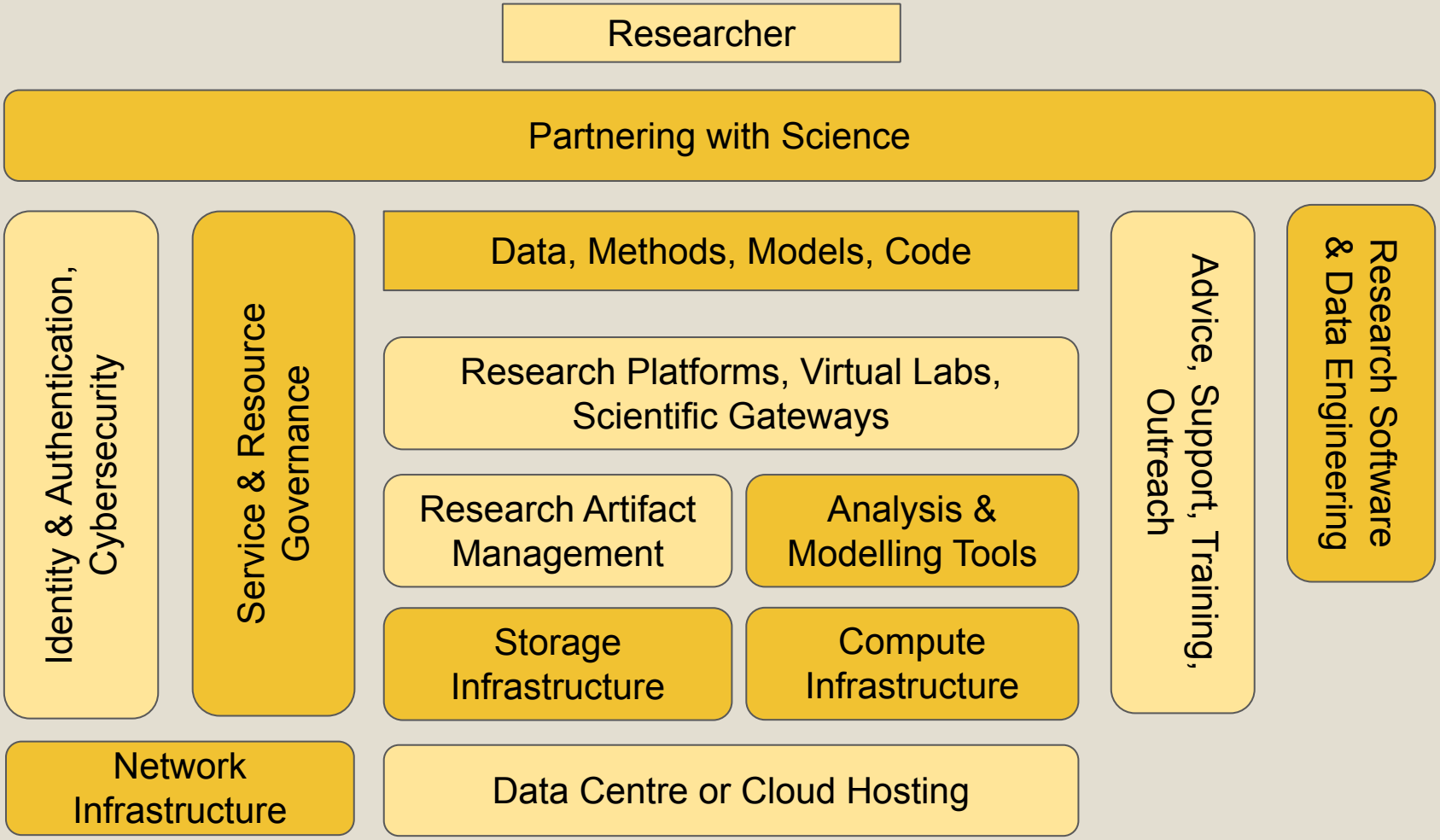
# Key eREF Components: Bread and Butter





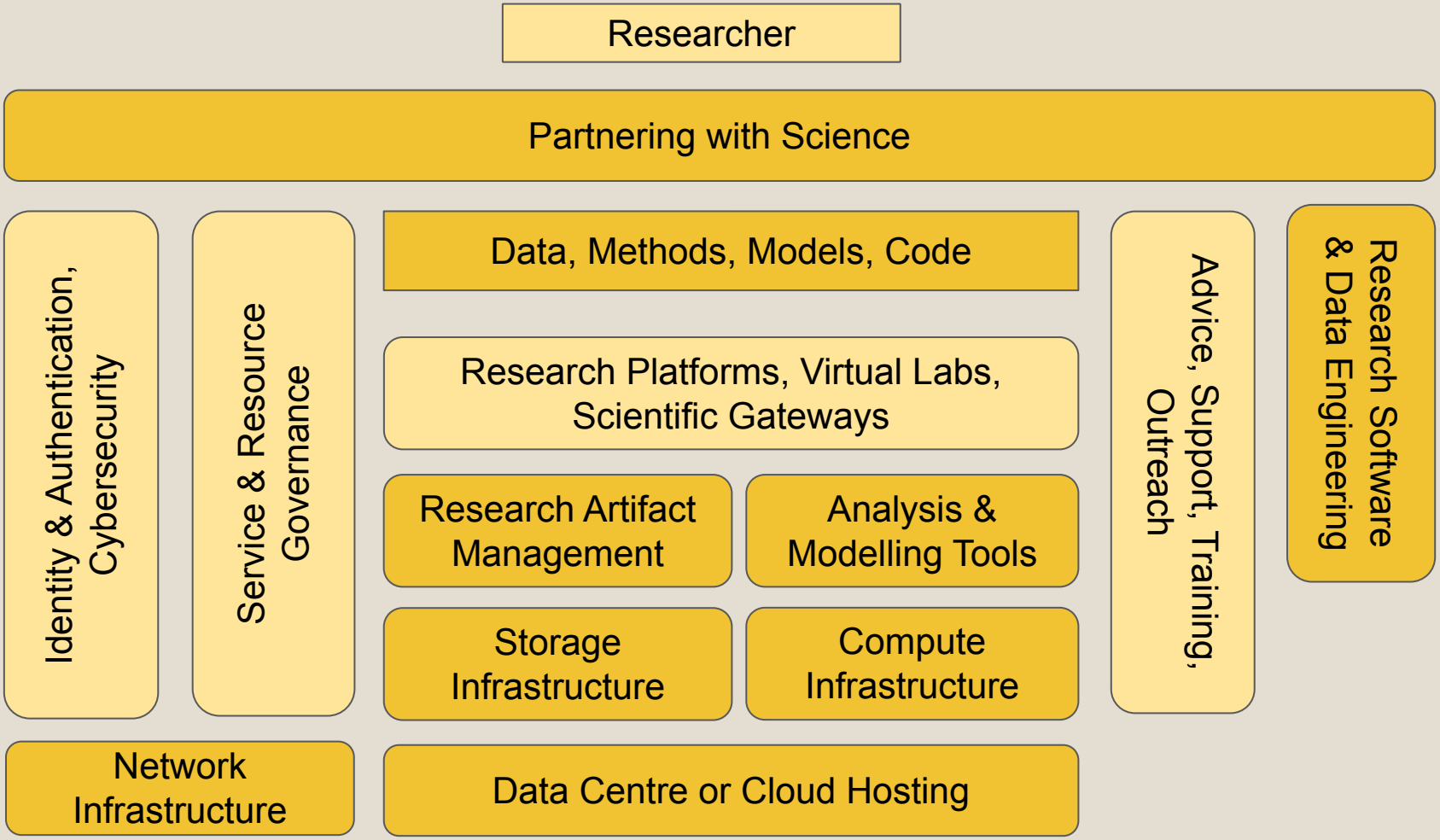
# Key eREF Components: Performance Busters





# Key eREF Components: Mission-based Research





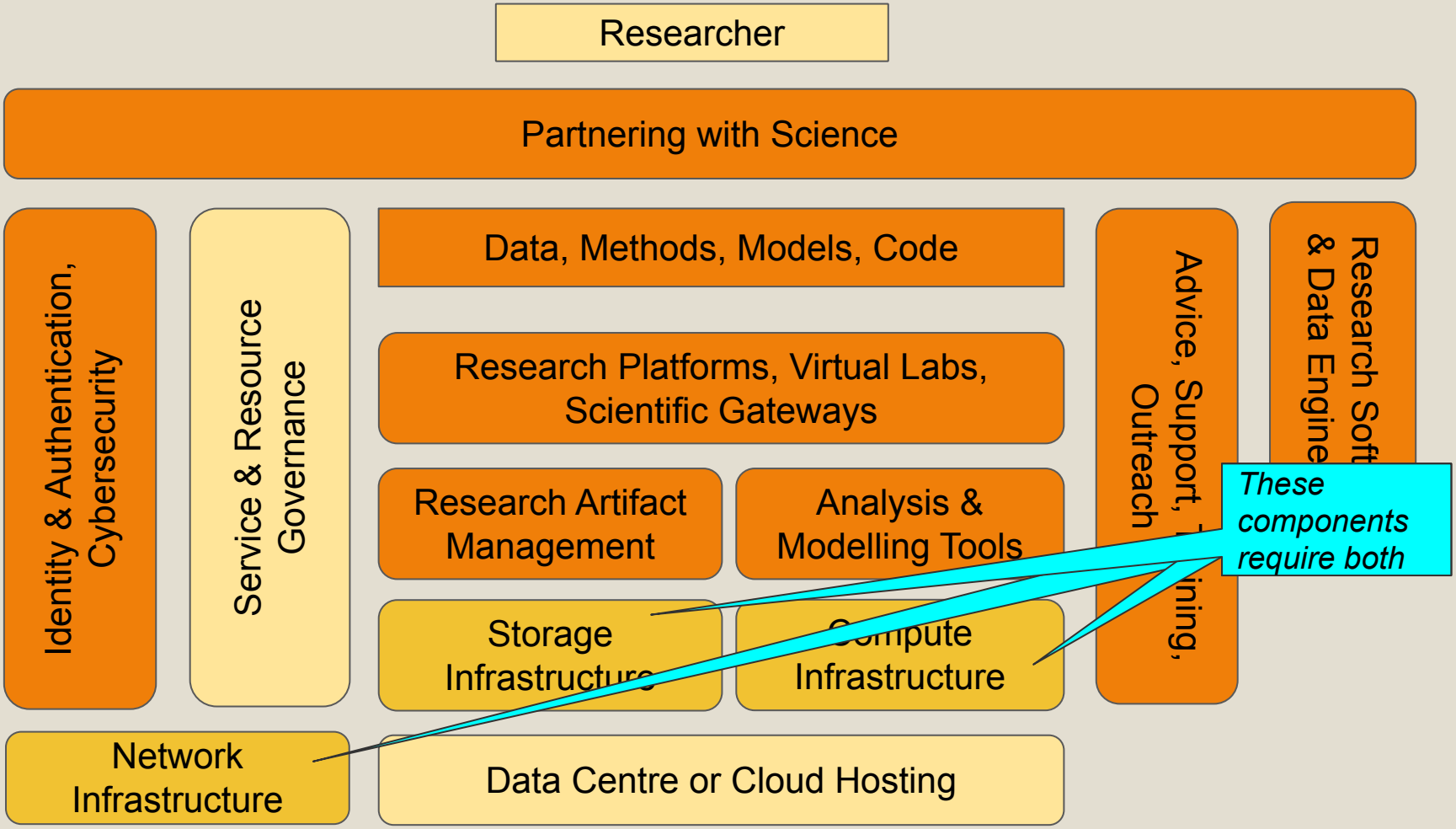
# Strategic Implications



BLOODSTONE

# Expertise More Important Than Technology

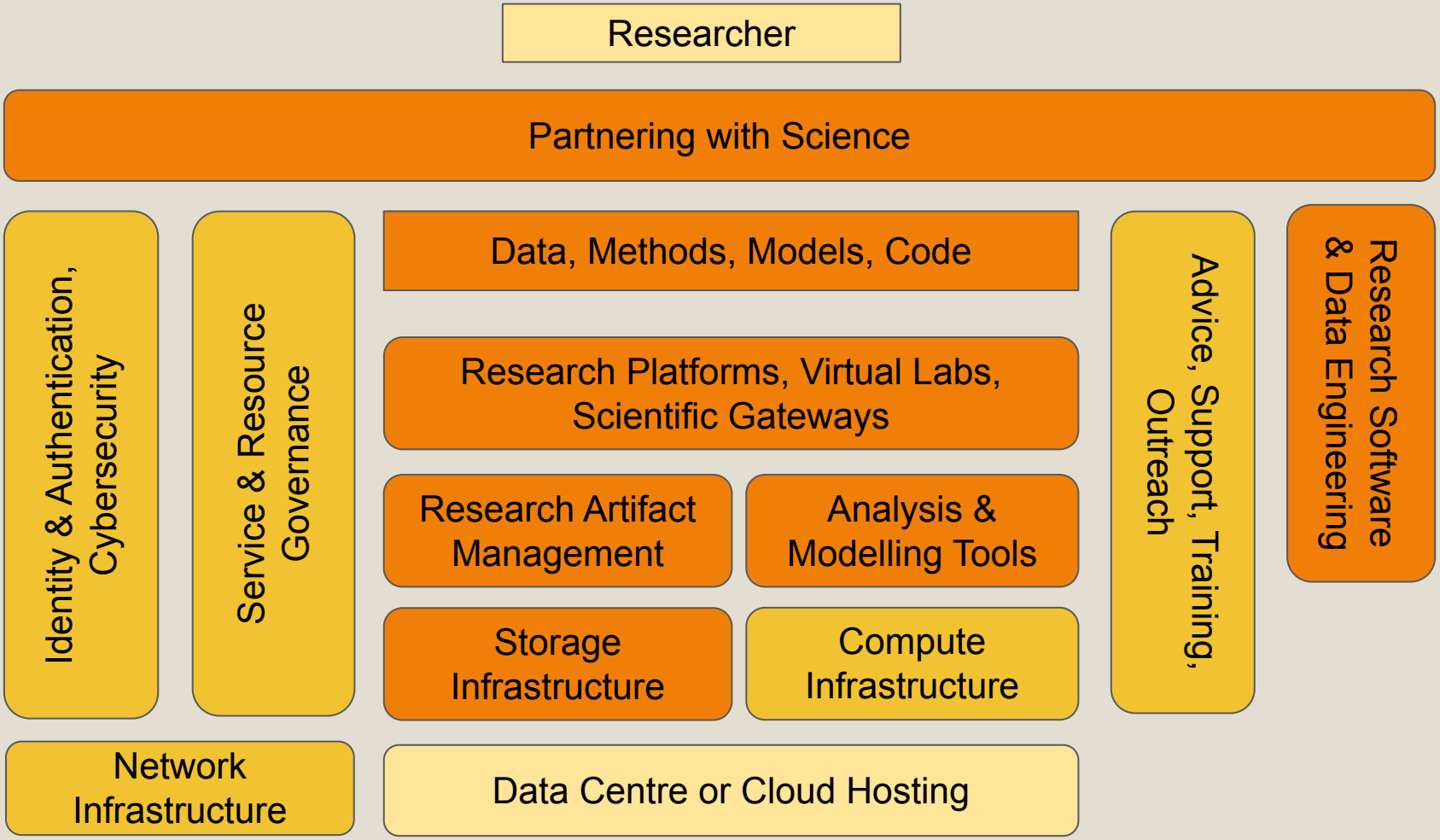




# Enabling International Collaboration

## But possibly requiring support for "different" tools...

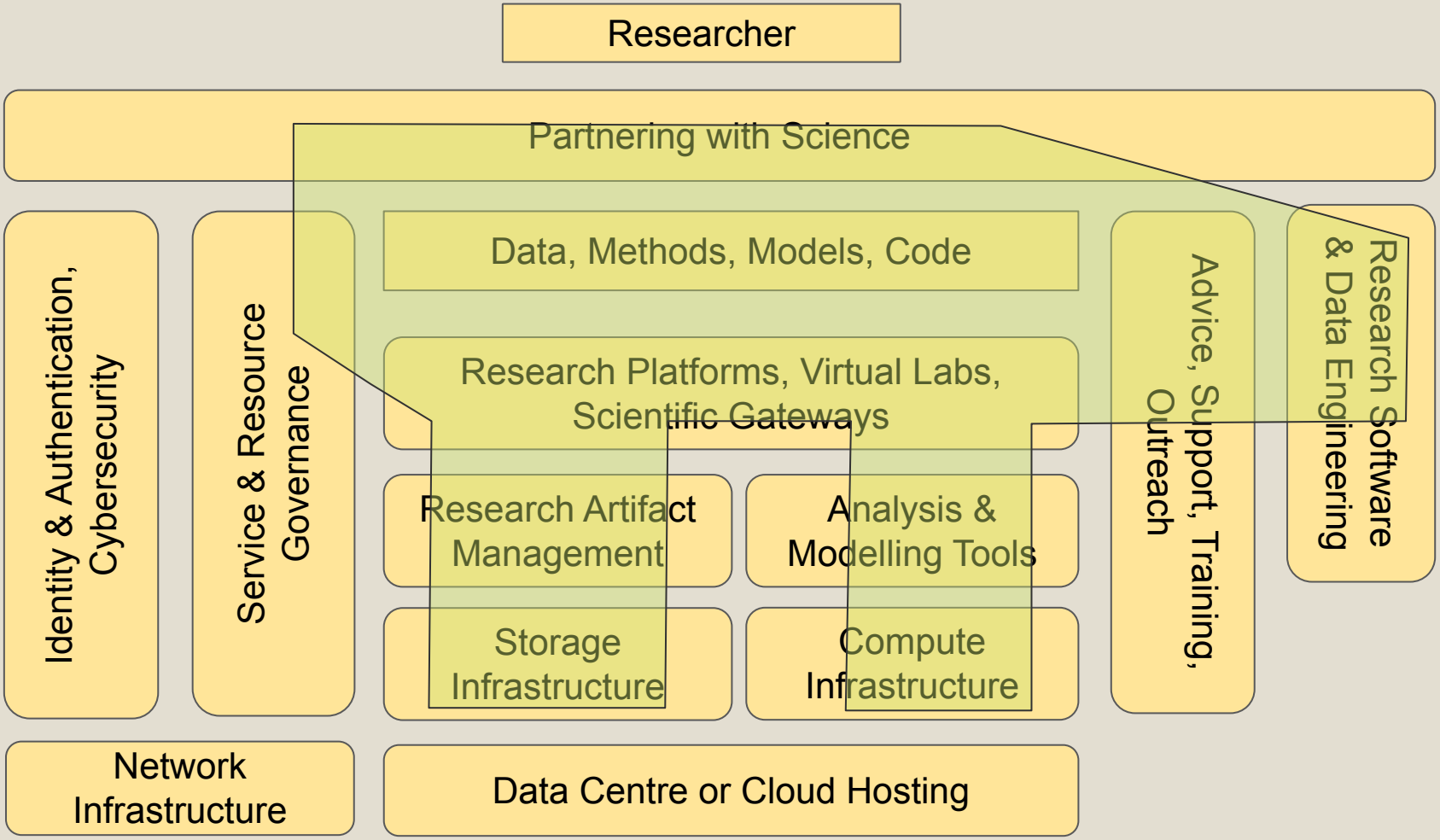




# Where Science Should Lead eResearch Activities

## E.g. Supporting CRIs, "Nationally Significant Databases"





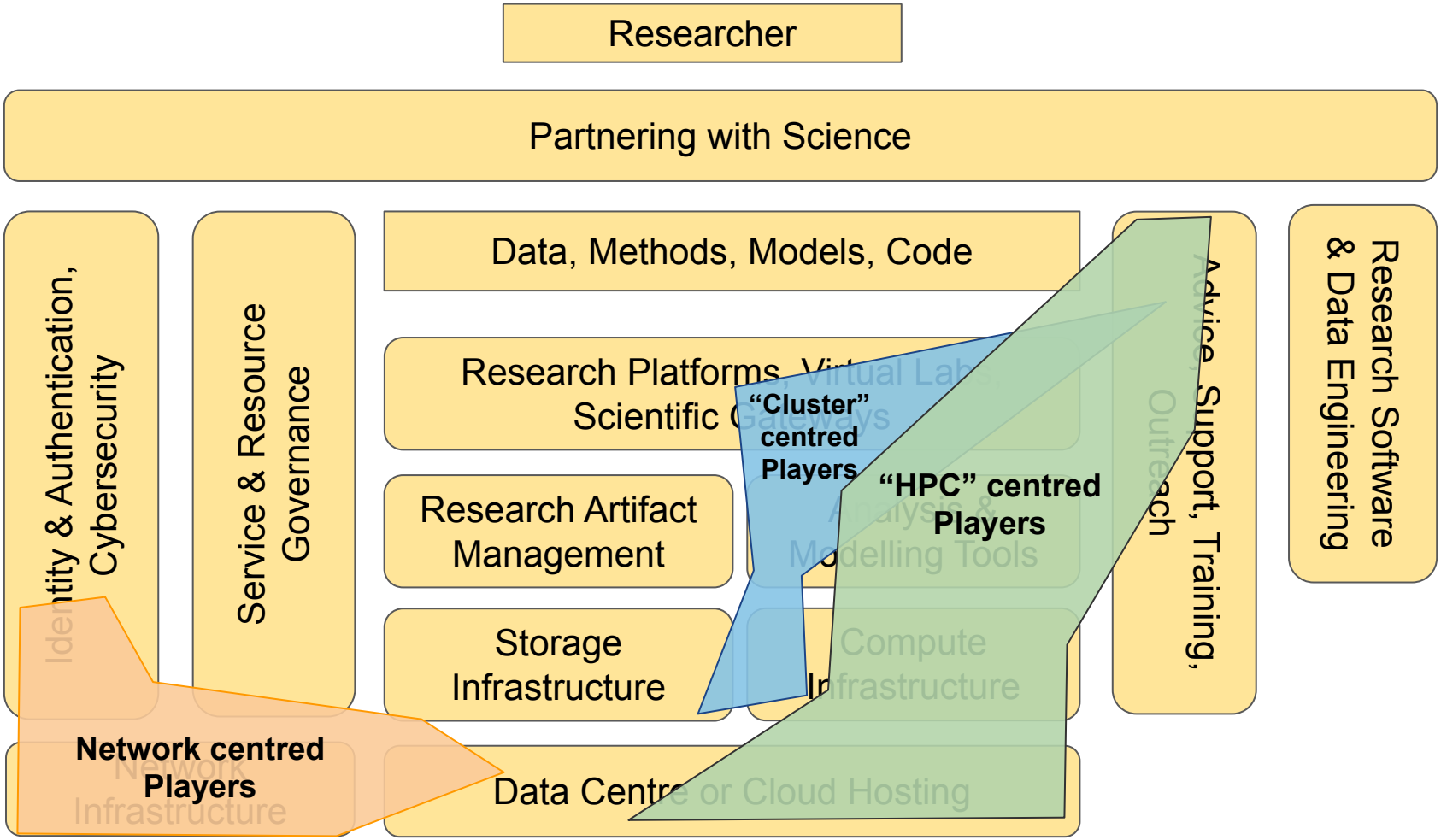
# Organizational Evolution



BLOODSTONE

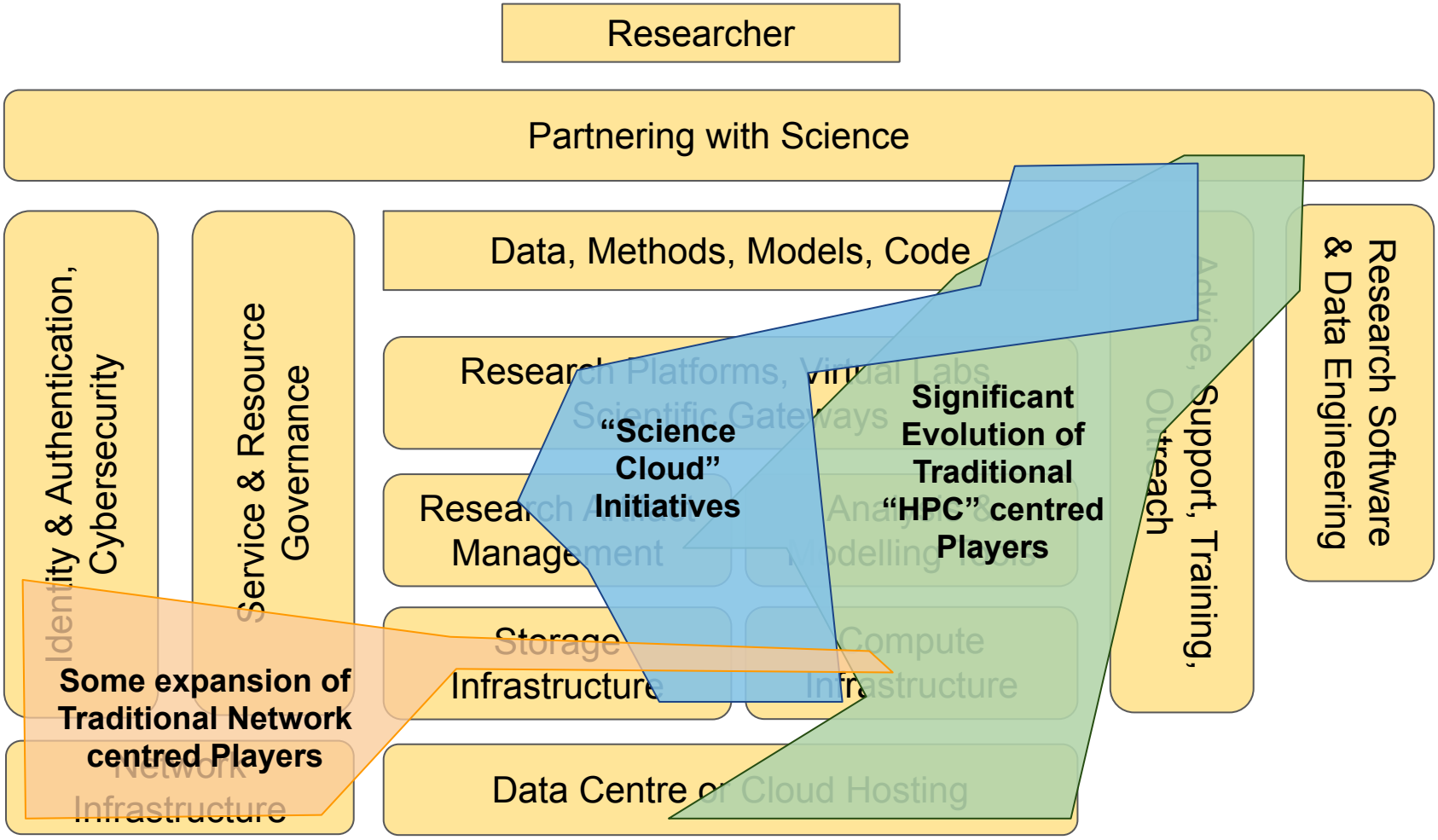
# eREF: Traditional Players





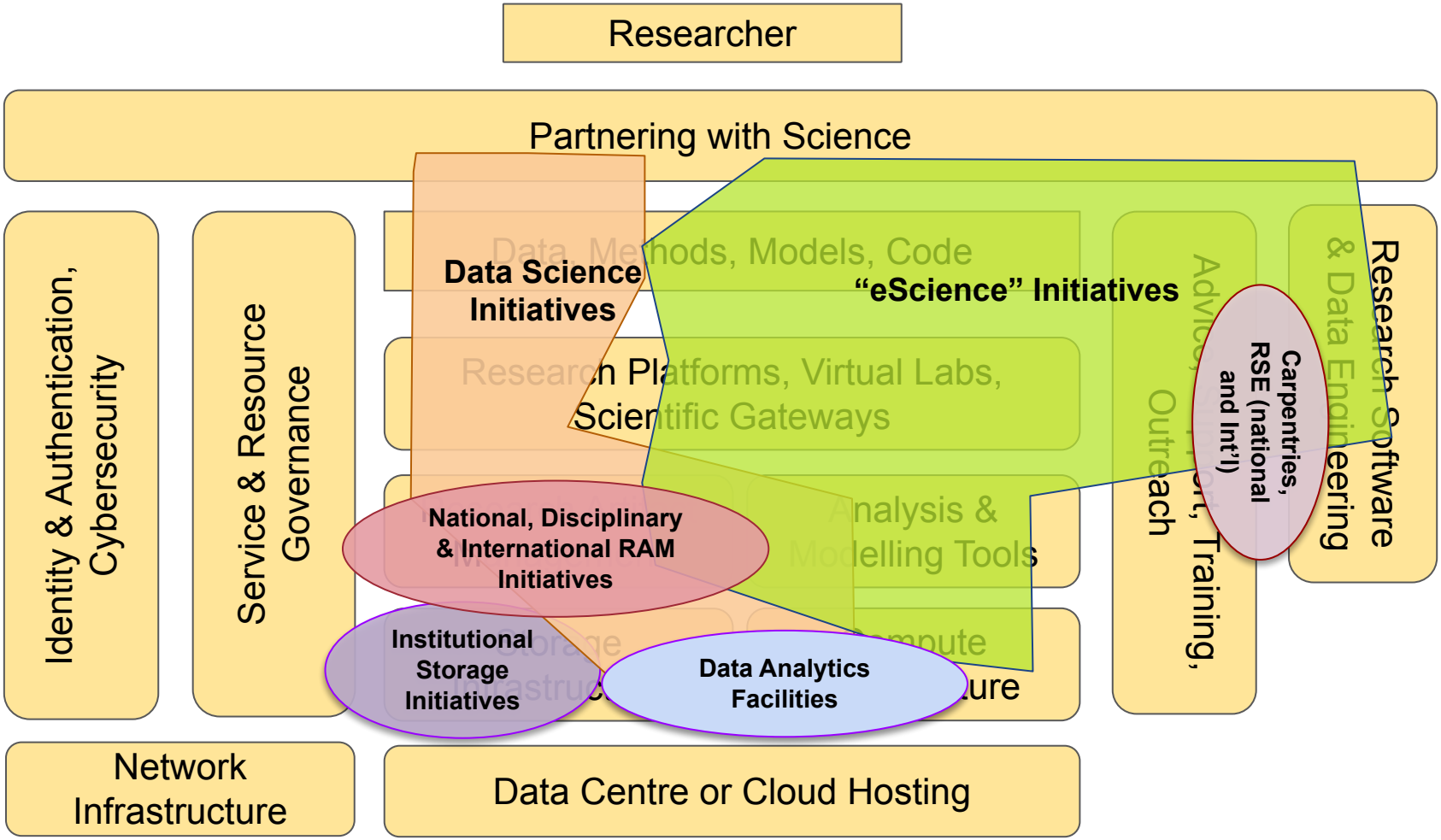
# eREF: Evolution of Traditional Players





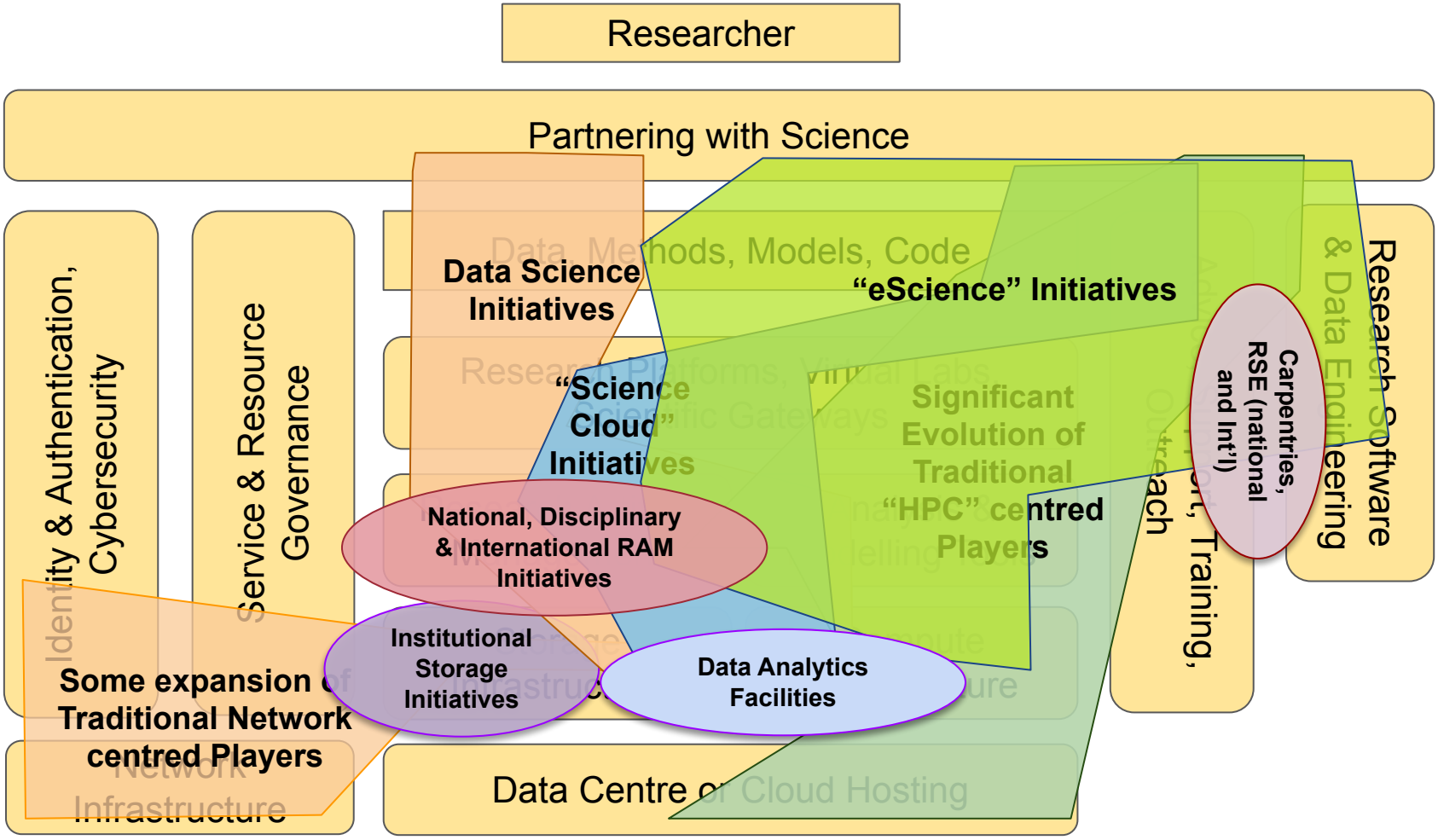
# eREF: New Players





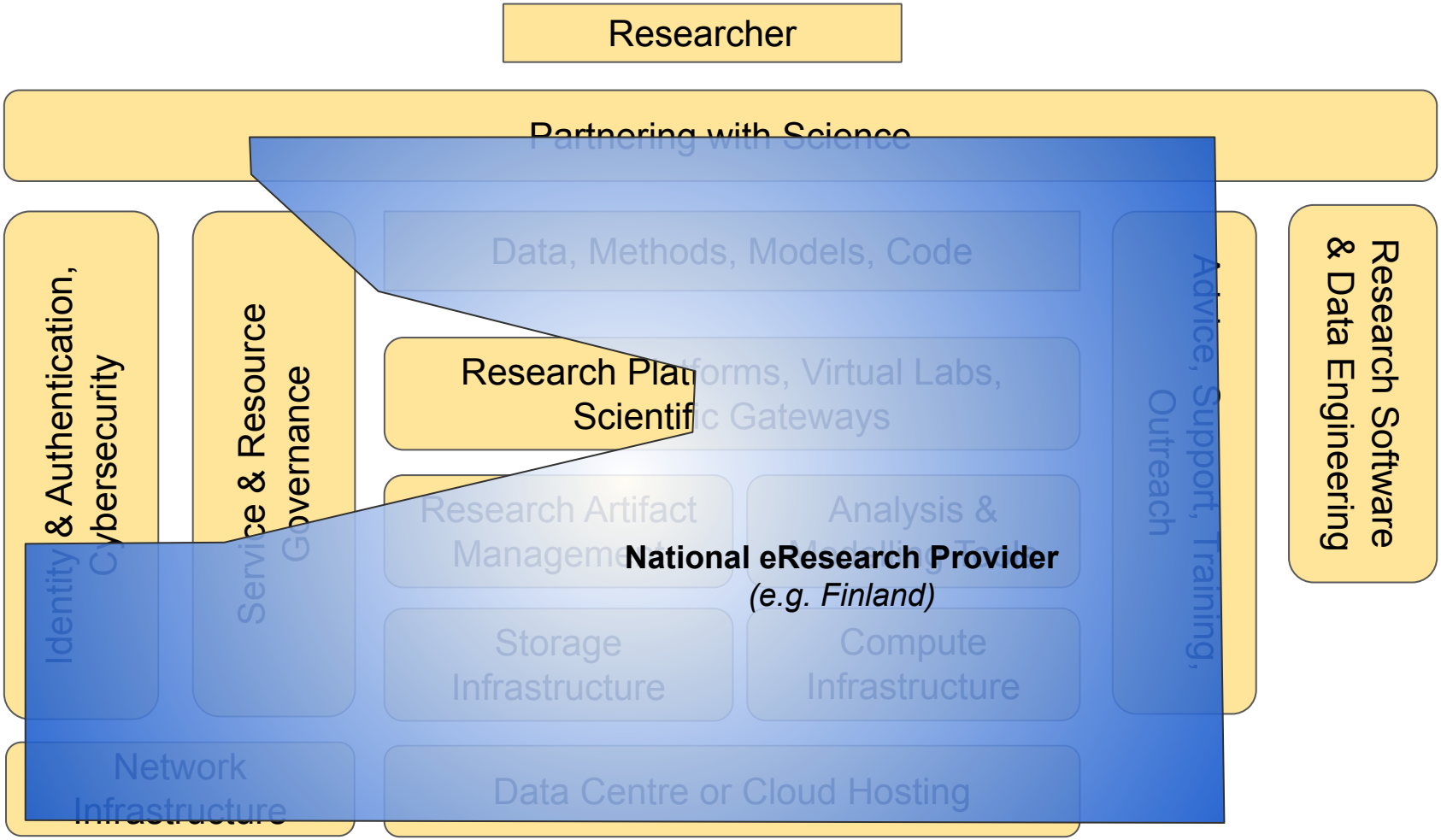
# eREF: Increasing Overlap





# eREF: Single or Federated Providers





# Organizational Learnings



eResearch Services are provided by multiple organizations in most of the jurisdictions examined

- Coordination/complementarity are common issues
  - Researchers must find/integrate the services they need
  - Competition for funding exacerbates this issue
  - Even when funding is clear, roles may still be unclear (e.g. as in Australia) and investment can be unbalanced
- Single national providers with a clear government mandate...
  - Can ensure that the full range of eResearch services are available to all researchers
    - e.g. CSC in Finland has been evaluated as a strong service provider
  - Can ensure that competition and overlap are minimized
  - Have developed in several countries larger than NZ
  - But must ensure they can meet even the most demanding requirements
- Federated structures seem promising but struggle with key issues, which can in turn limit the services offered and effectiveness

# Comparing Ecosystems Across 8 Jurisdictions



BLOODSTONE

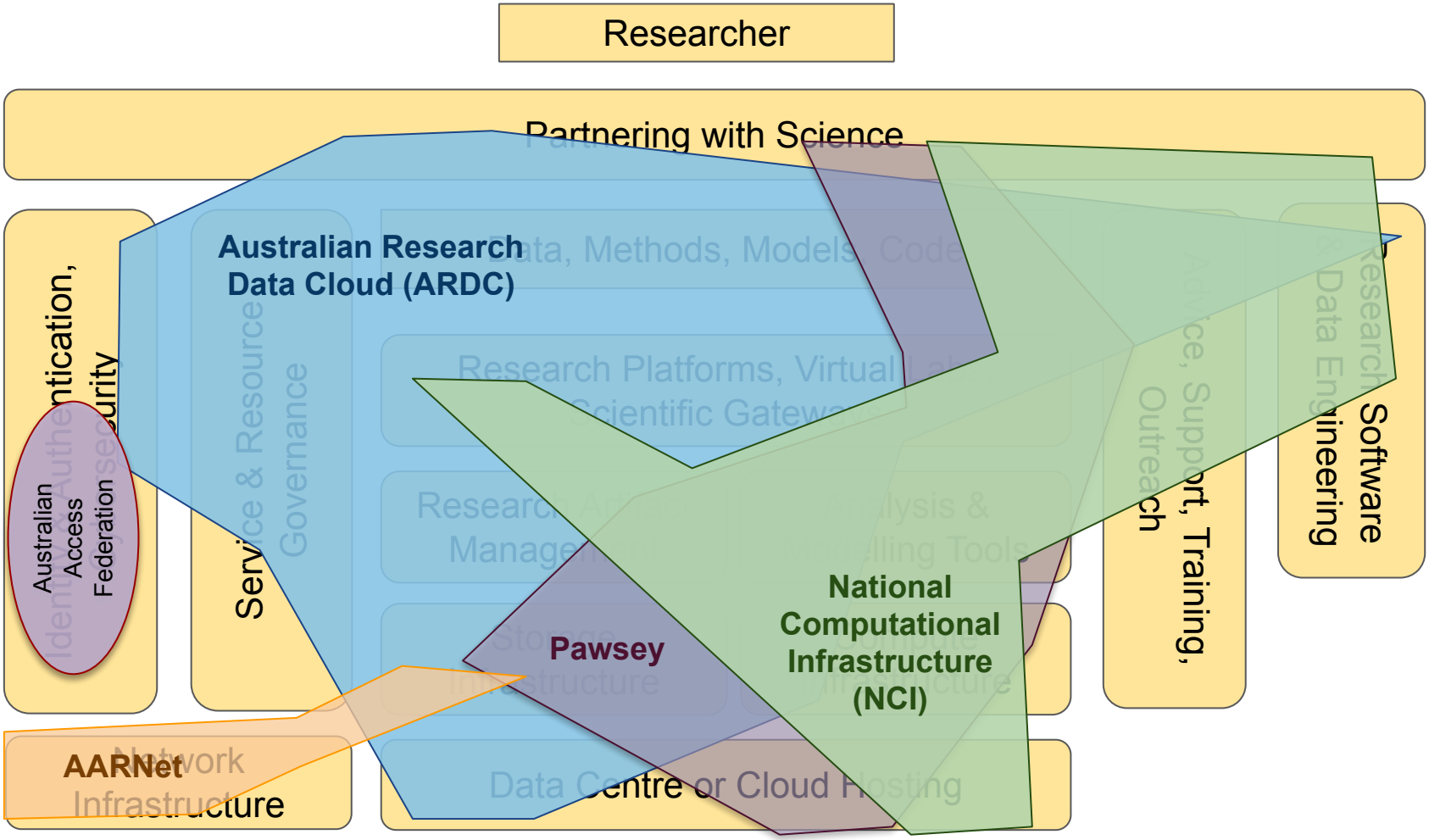
# Country-by-Country Learnings



- Key Players
- End User Experience (heat maps)
- Comments, capturing directions if apparent

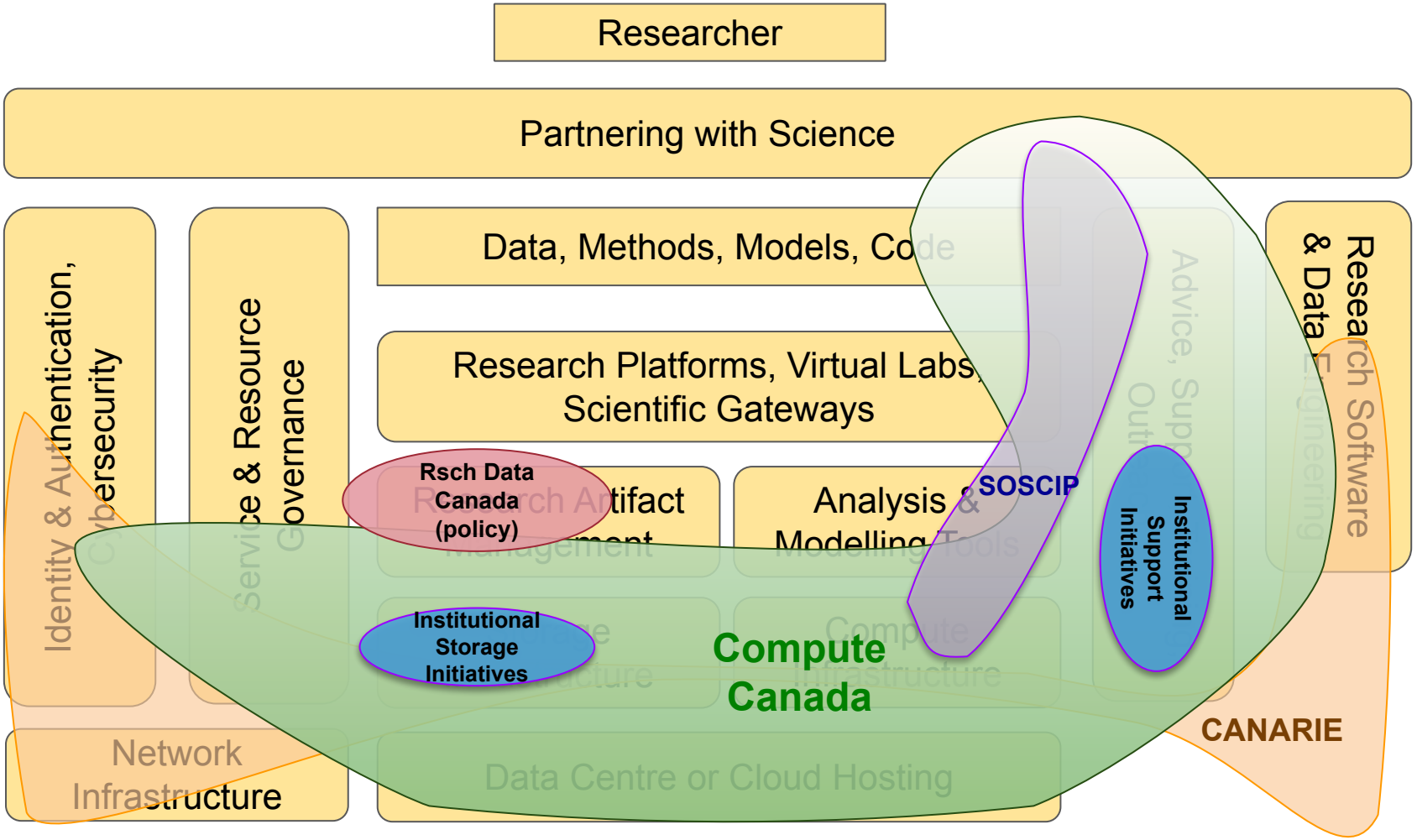
# Australian Players





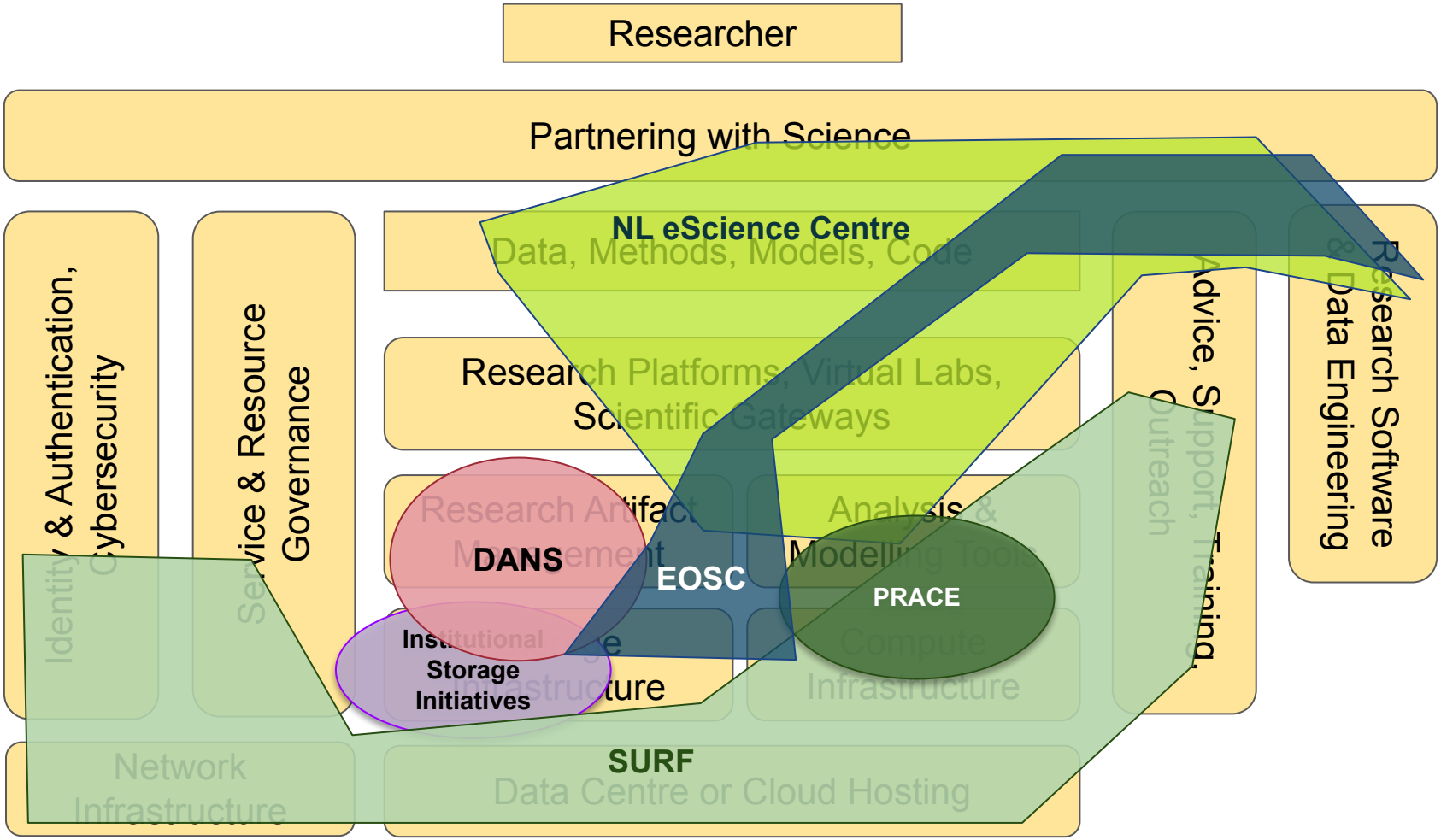
# Canadian Players





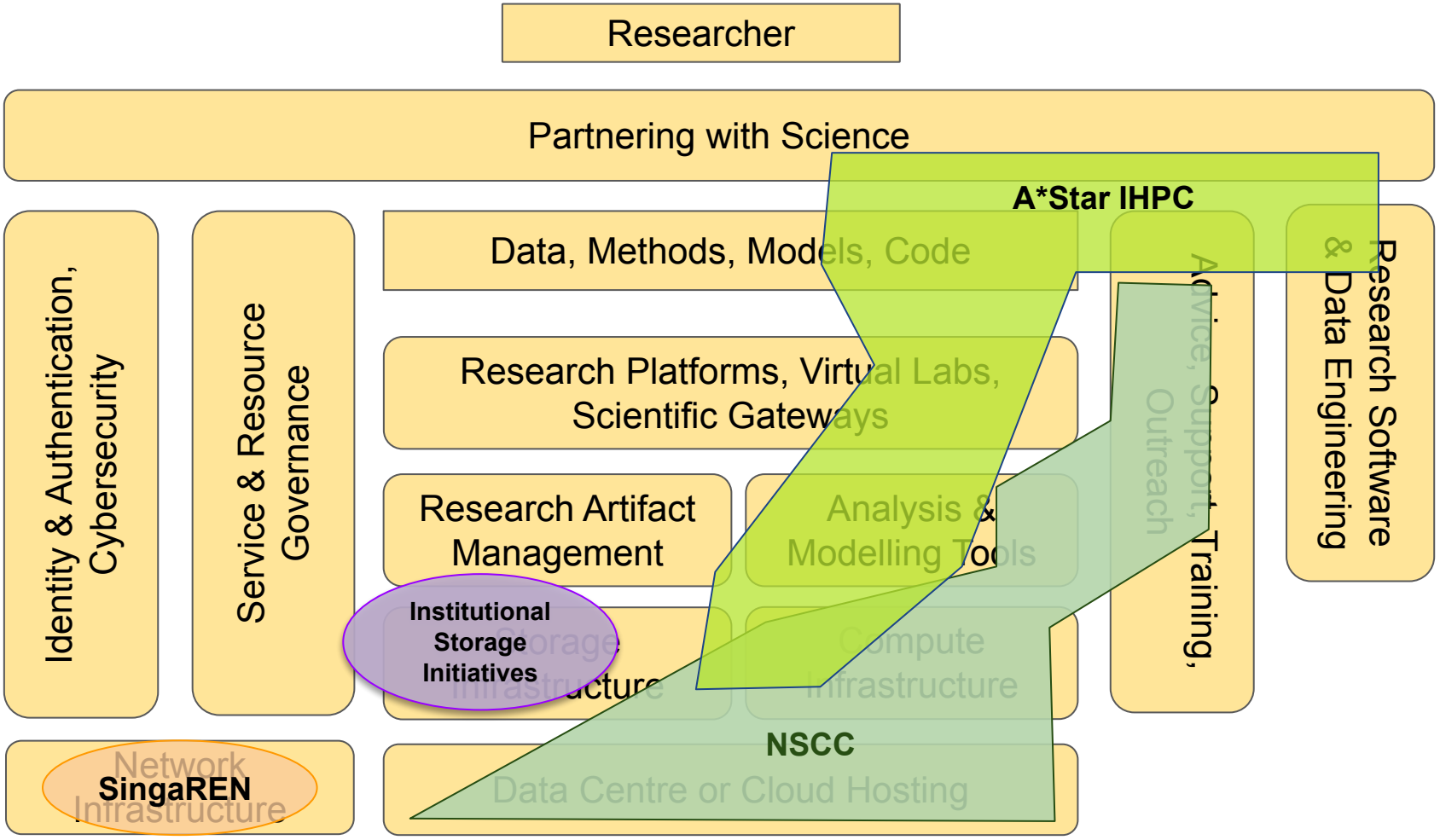
# Dutch Players





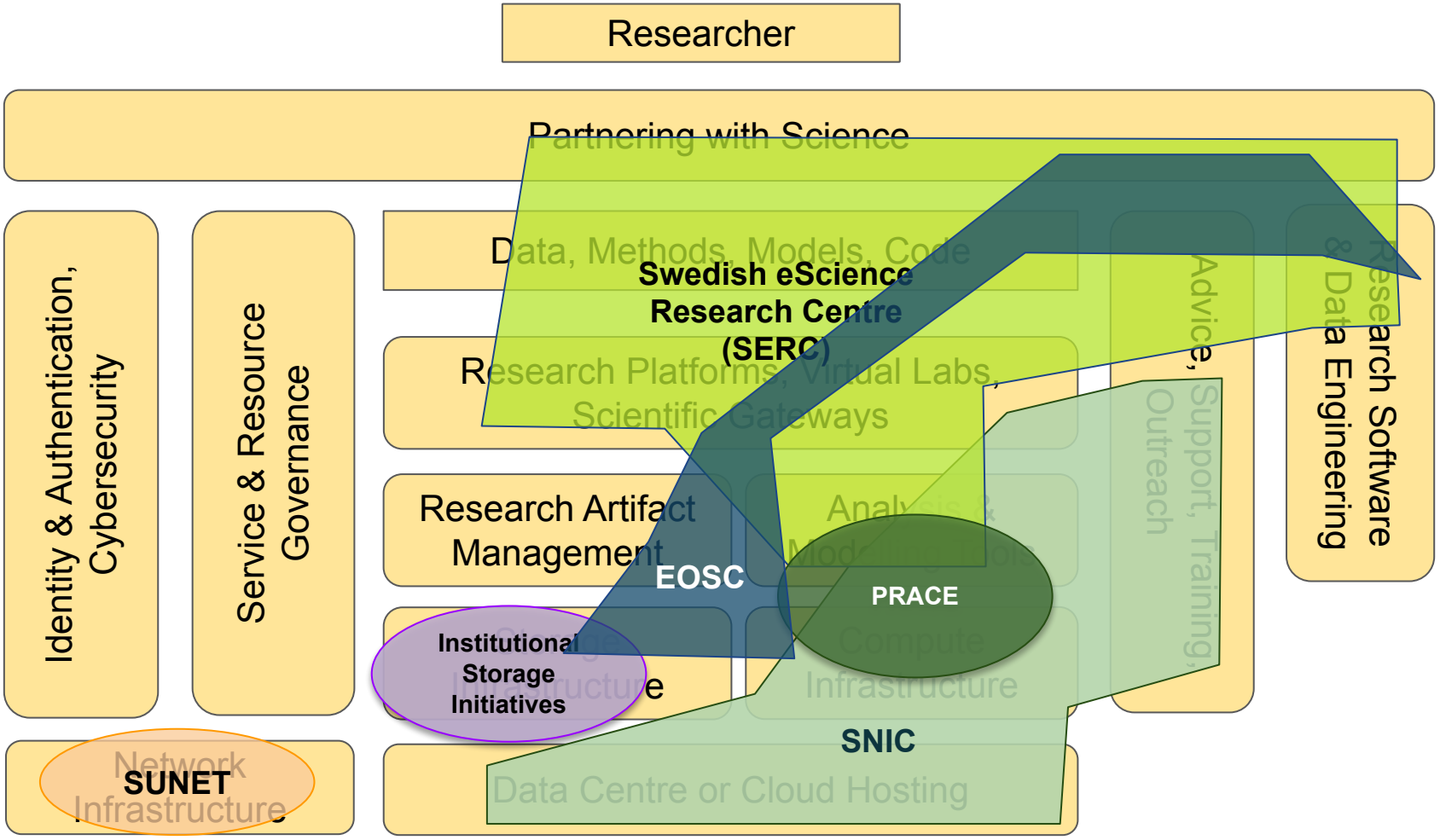
# Singapore Players





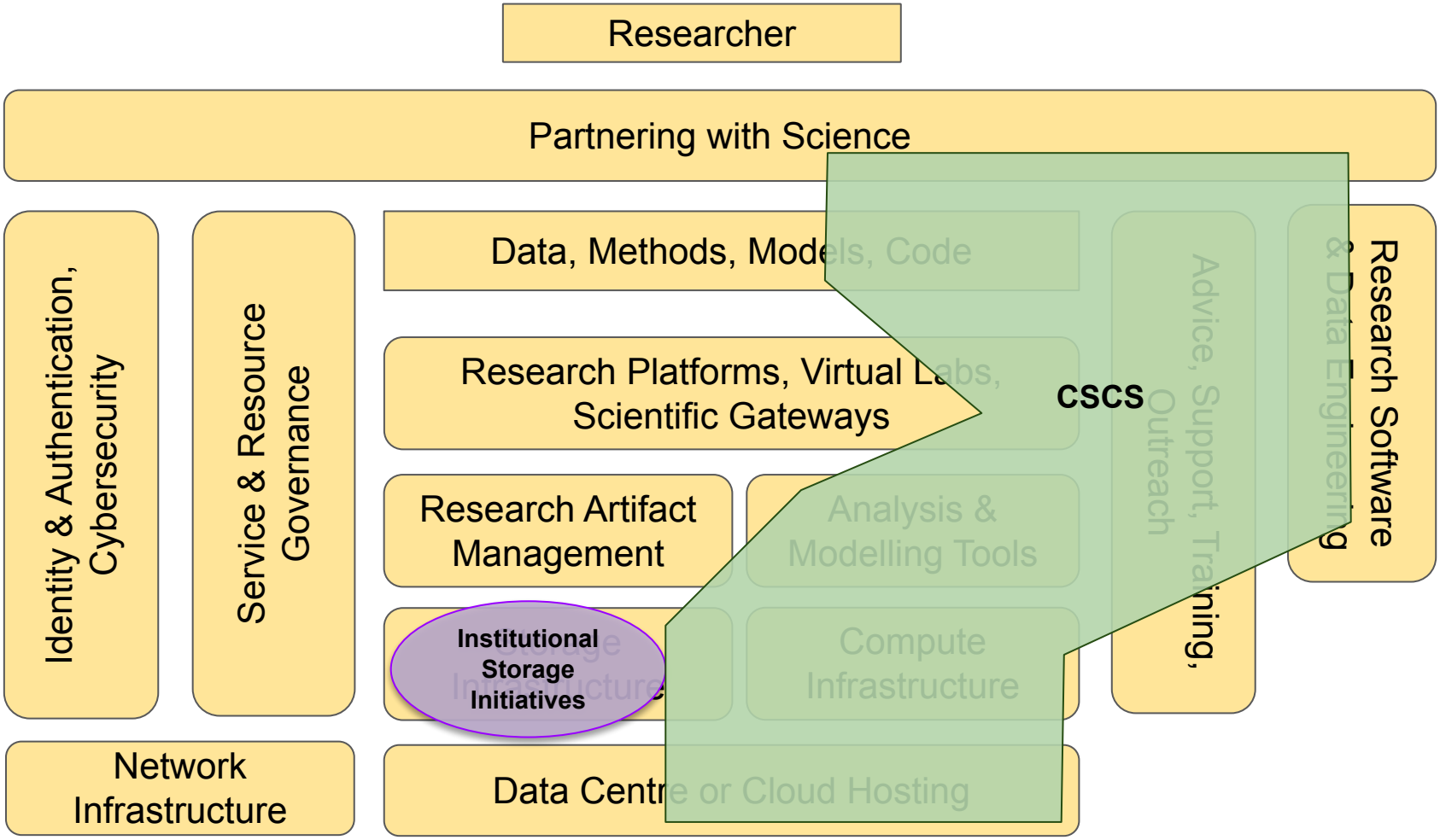
# Swedish Players





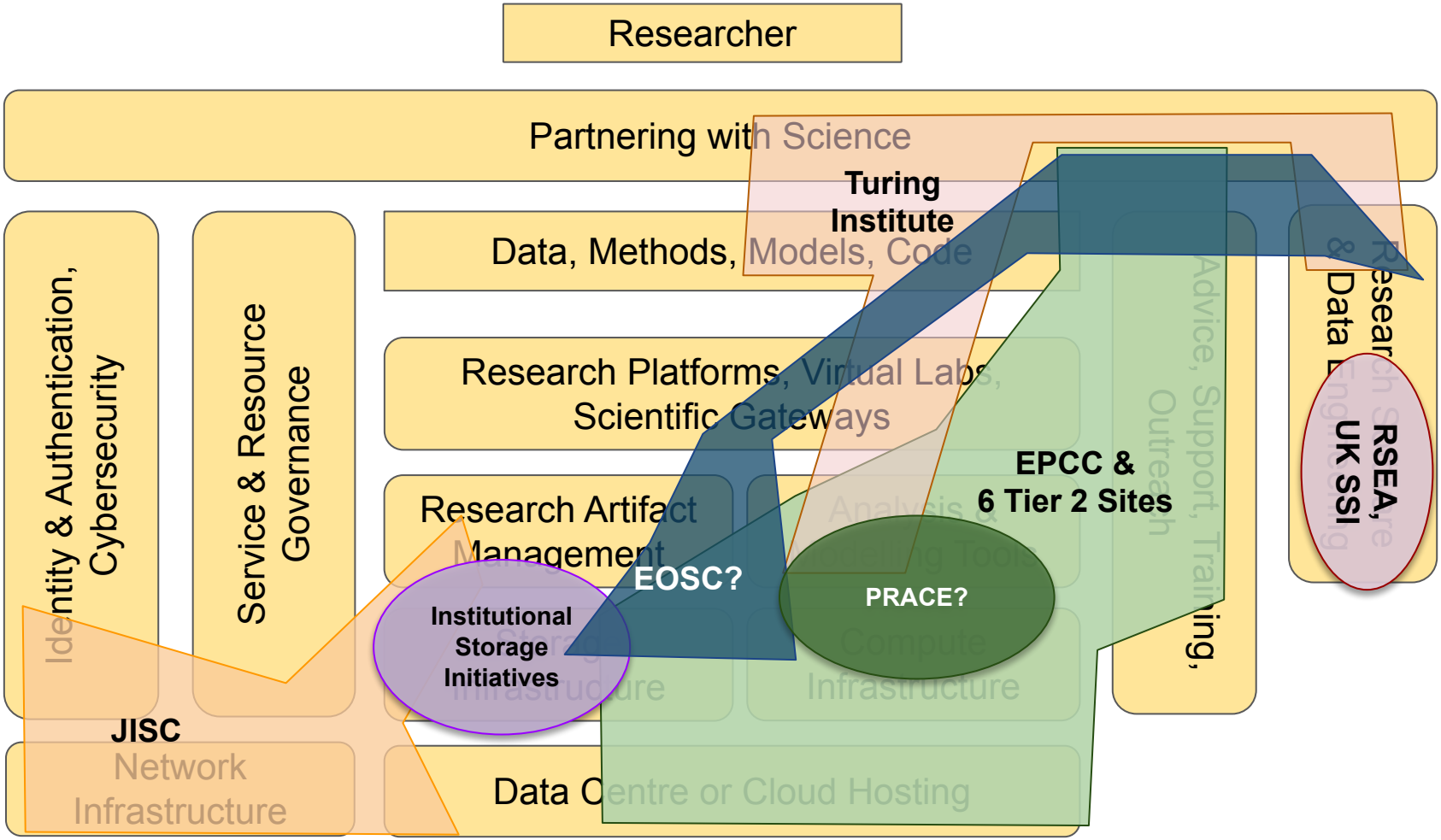
# Swiss Players





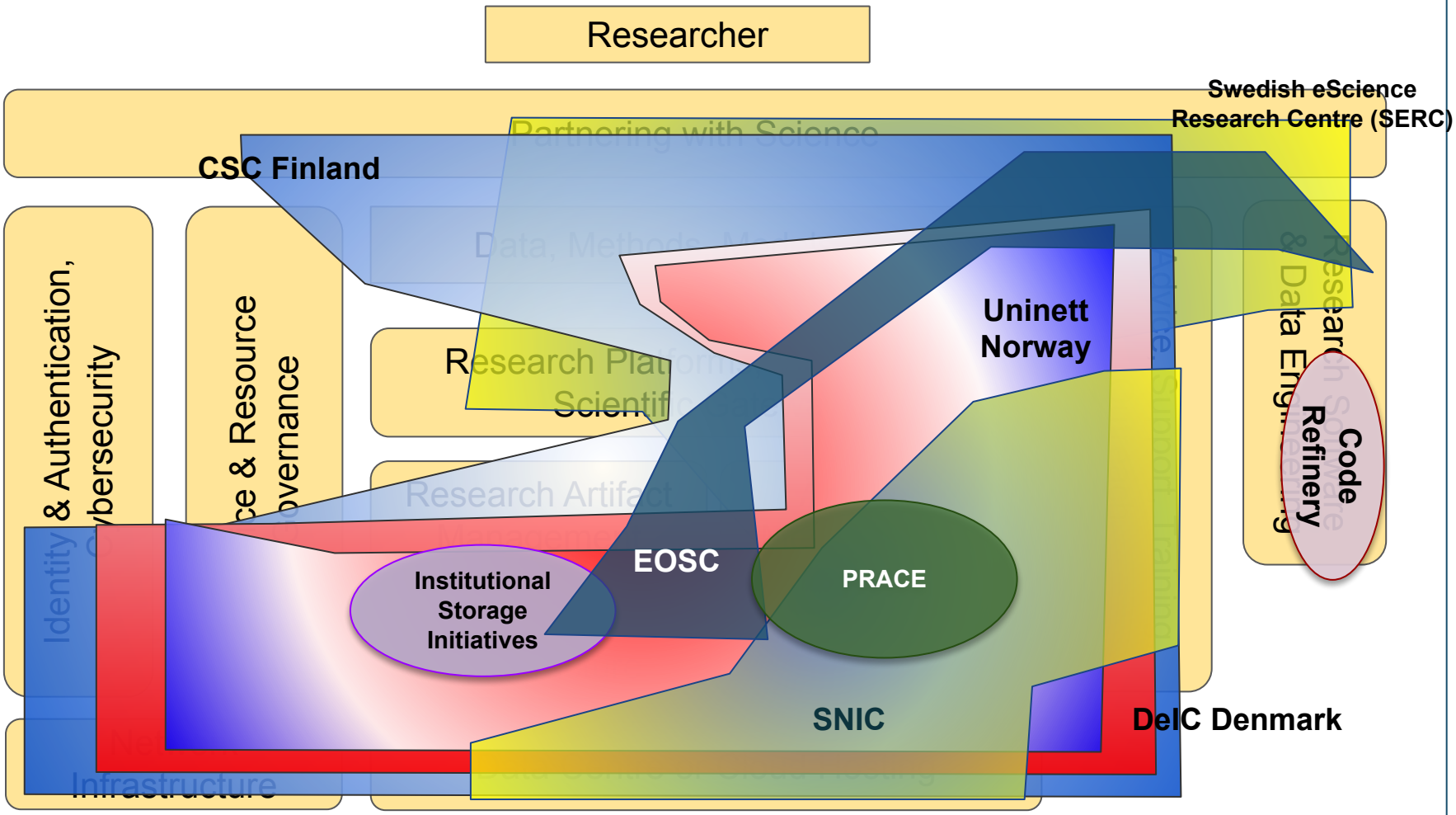
# UK Players





# Nordic Players

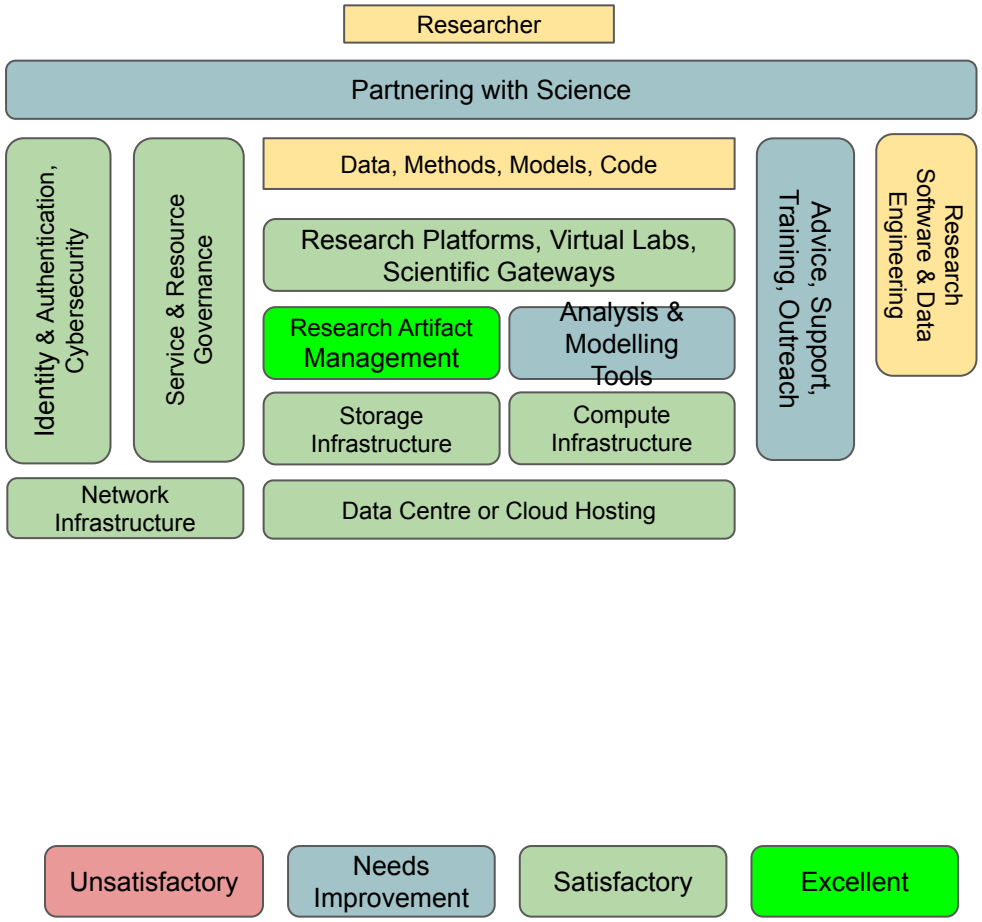




# Australian User Experience



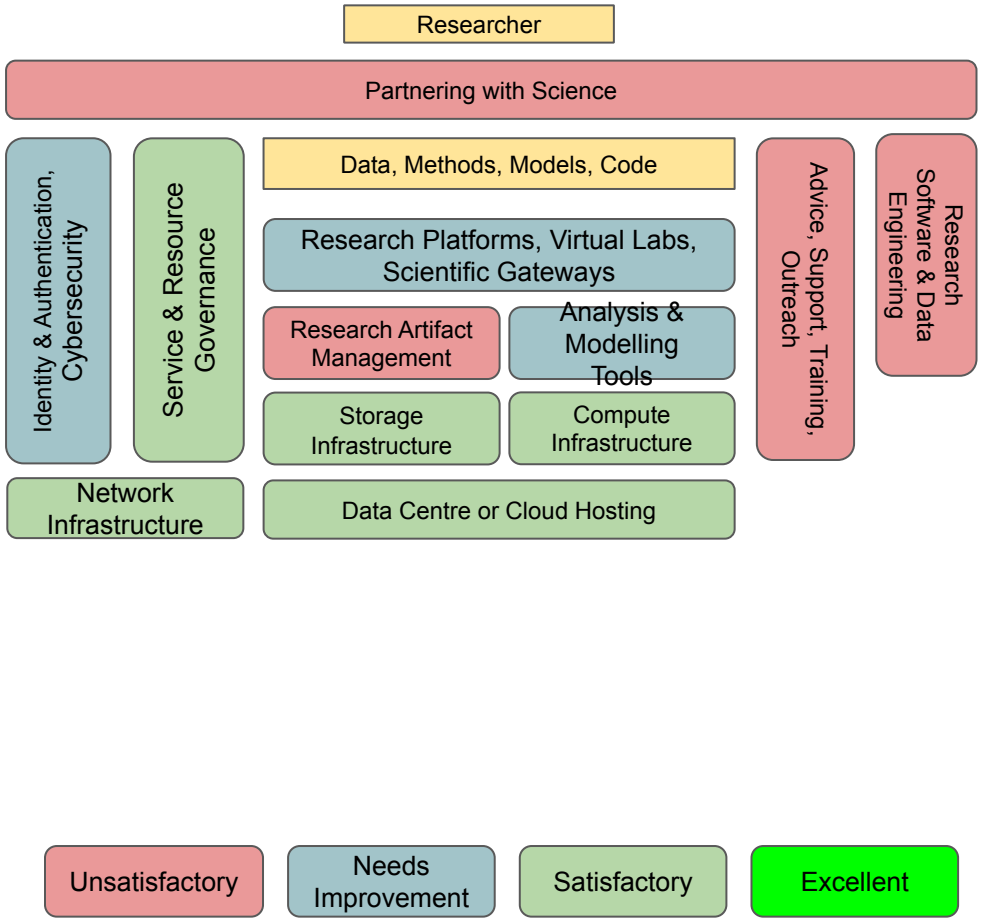
- eInfrastructure needs generally well served today.
  - New investments will improve supply for a period
  - Sustainability still needs to be addressed
- International leader in RDM
- Despite long focus on "upper levels" of ecosystem ("eScience") Australia has fewer explicit supports here. Selected research areas are well served, but main players need to work through conflicting mandates before ecosystem can be broadly effective.



# Canadian User Experience



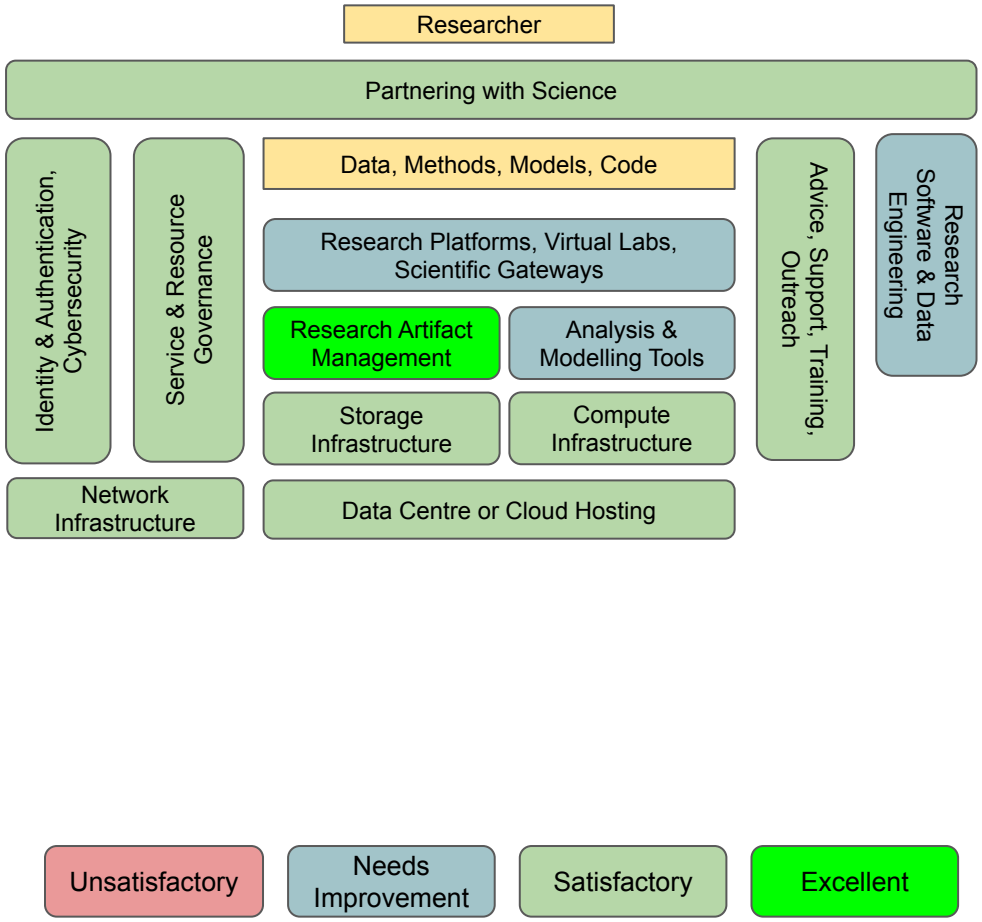
- eInfrastructure needs generally well served today.
  - Recent investments have improved supply and government has committed regular investment
- Aspires to lead in RDM but mandates are unclear and there is a turf war
- Access and security fragmented
- Mixed experience in "eScience" of ecosystem.
  - Ecosystem does not provide appropriate science support (Partnering, Tools, RSE)
  - CC recognizes importance of gateways but there is no funded strategy
  - Support/Training function likely to be devolved to CC's regional partners, but this could exacerbate uneven service levels.



# Dutch User Experience



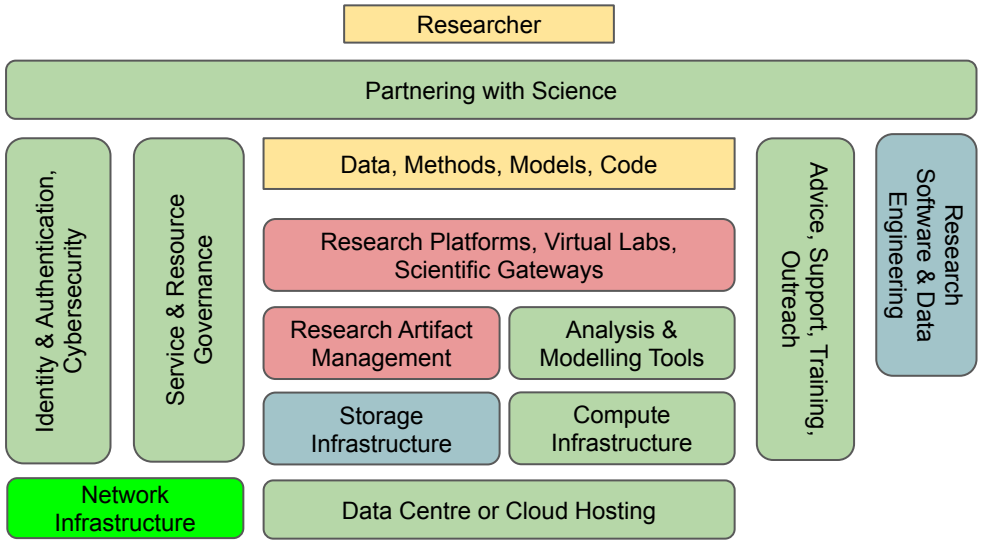
- eInfrastructure needs generally well served today.
  - PRACE relieves pressure on Compute
  - Institutions+DANS address persistent and archival storage
- DANS easy to use for RDM, but larger datasets are more challenging
- NLeSC leads and innovates for "eScience" levels of ecosystem.
  - Institutions moving to integrate eScience supports within each institution.
  - No good mechanisms for sustaining eScience investments (gateways, codes)

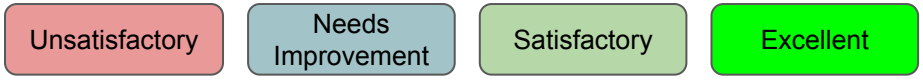


# Singapore User Experience



- eInfrastructure needs generally well served today.
  - NSCC provides significant capability for local research community, as well as industry
  - Institutions have additional resources
- A*Star IHPC collaborates in "eScience" levels of ecosystem.
- Little focus on creating or maintaining higher level enabling infrastructure (gateways, RAM)
- Internal and external networking a strength

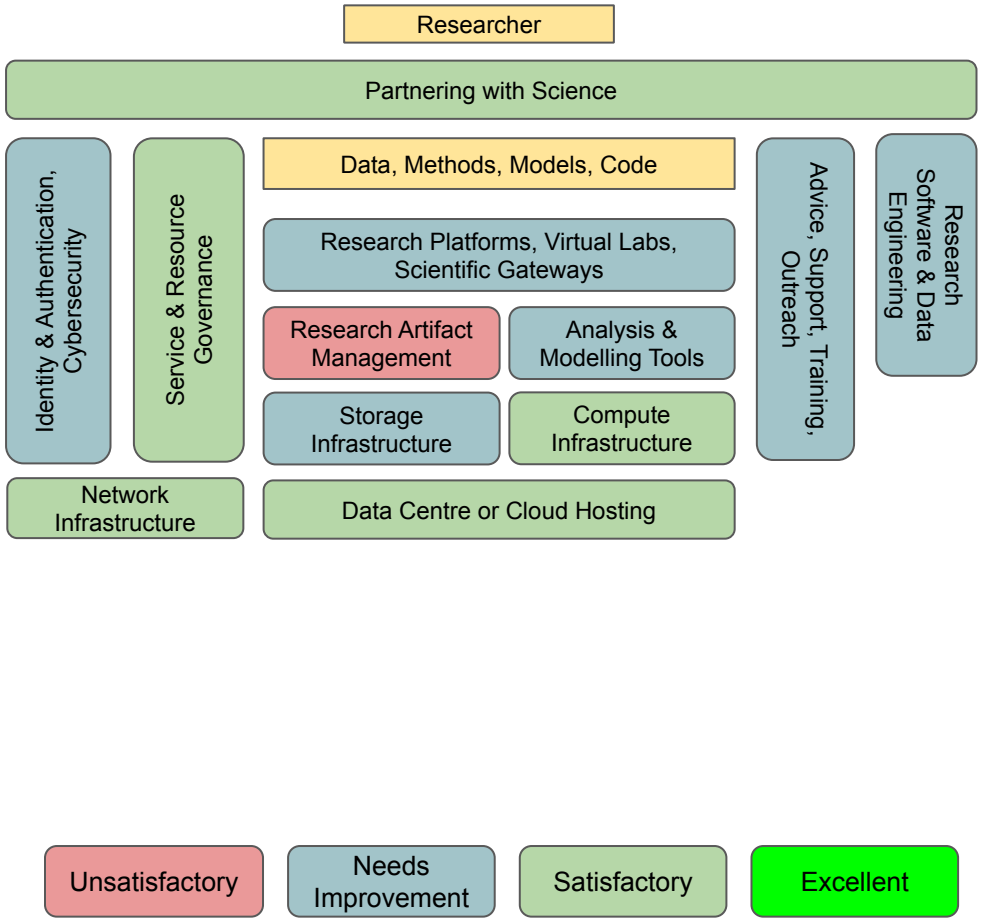


# Swedish User Experience



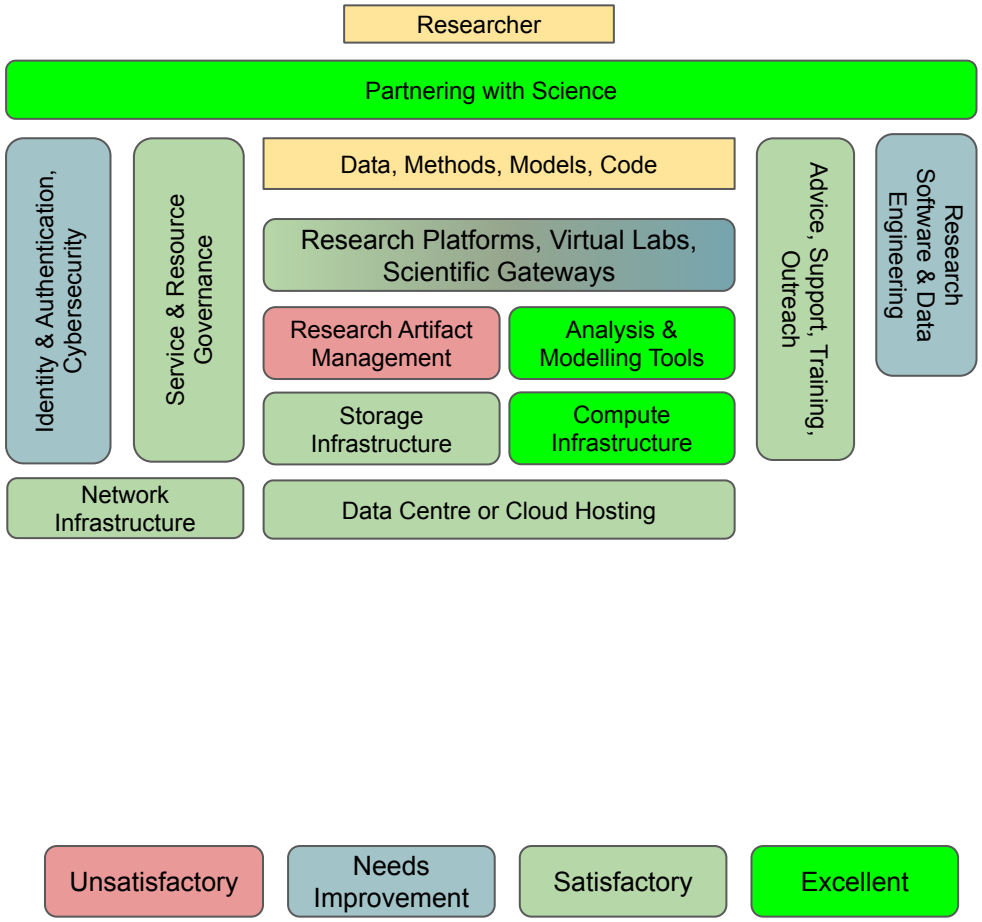
- eInfrastructure needs mostly well served today.
  - Individual SNIC institutions continue to be successful in funding compute systems
  - PRACE relieves pressure on Compute
  - Persistent storage remains challenging -- seen by funding agency as an institutional responsibility
- Robust RAM is a challenge
- Higher level "eScience" activities remain more challenging for users.
  - Swedish eScience Research Centre is a resource but unclear if institutions are stepping up, or if other supports are needed.
  - No good mechanisms for sustaining eScience investments (gateways, codes)



# Swiss User Experience



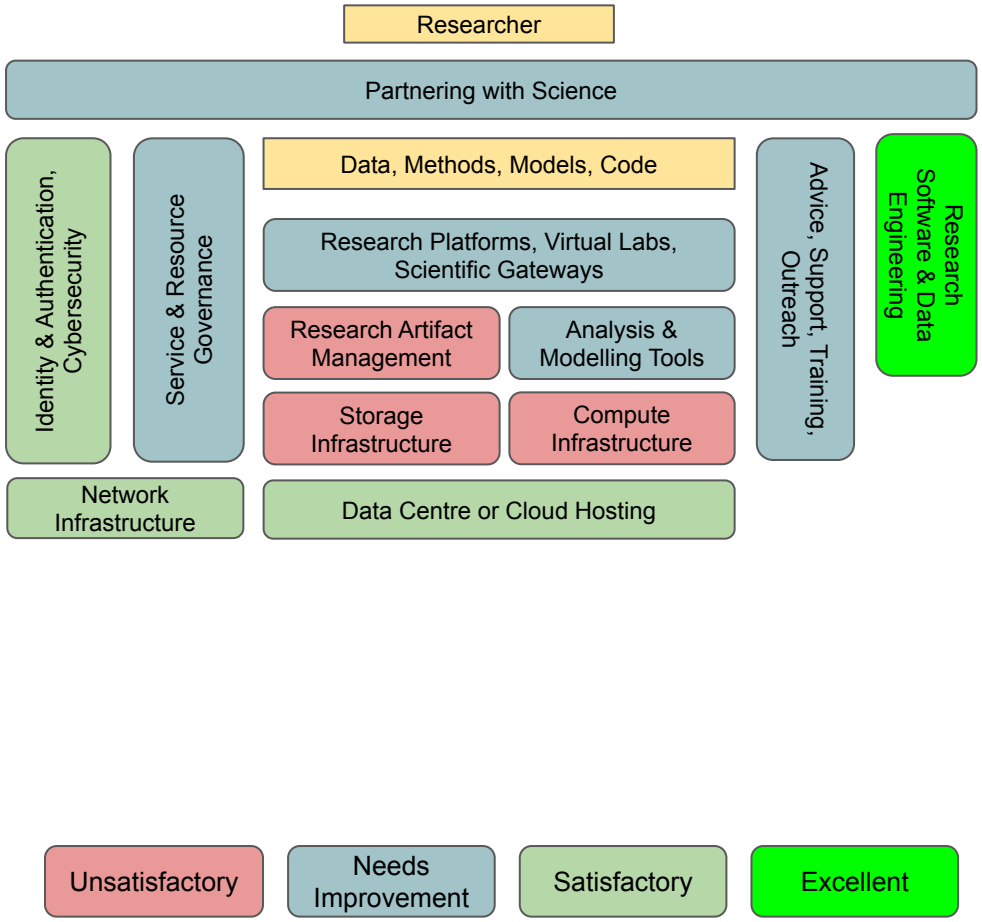
- CSCS focuses on scientifically-led analysis/modelling tools, optimized for its well-endowed HPC system.
  - Some platforms also managed in the same context
- No RAM initiatives beyond those needed by the projects CSCS supports directly (e.g. materials science, neuroscience)
- Other components seem reasonably well supplied, albeit without a clear ecosystem approach.



# UK User Experience



- UK losing competitive eInfrastructure access today.
  - Tier 2 Compute investments will help, but Archer is beyond useful life
- No agreed initiatives for RAM, although JISC attempted.
- Storage is up to institutions.
- Enviable support for software initiatives and professionalization through RSEA
  - Otherwise no strategic supports for eScience tools (gateways, codes, RAM)



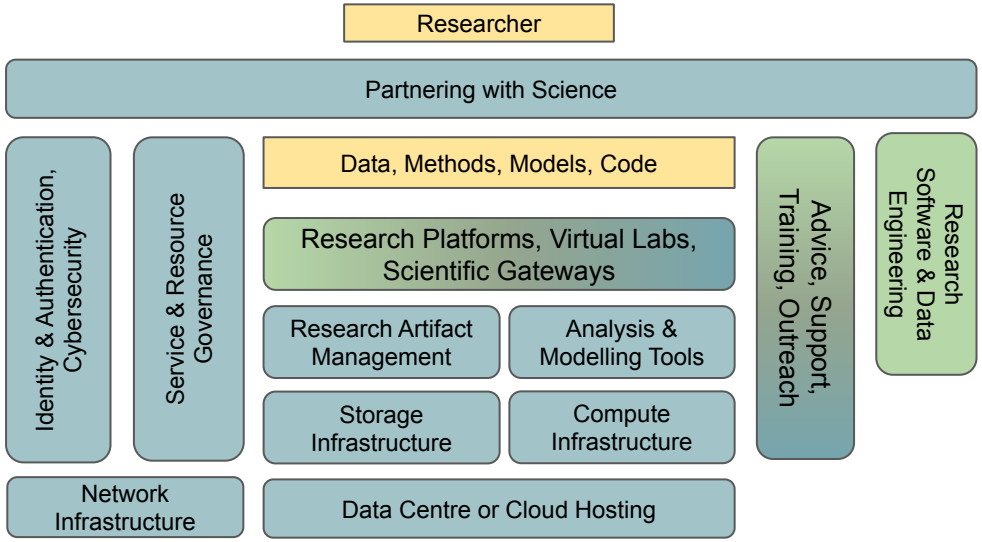
# Nordic User Experience

60

Generalizing the Nordic user experience is difficult -- there is really no "common" Nordic user experience.

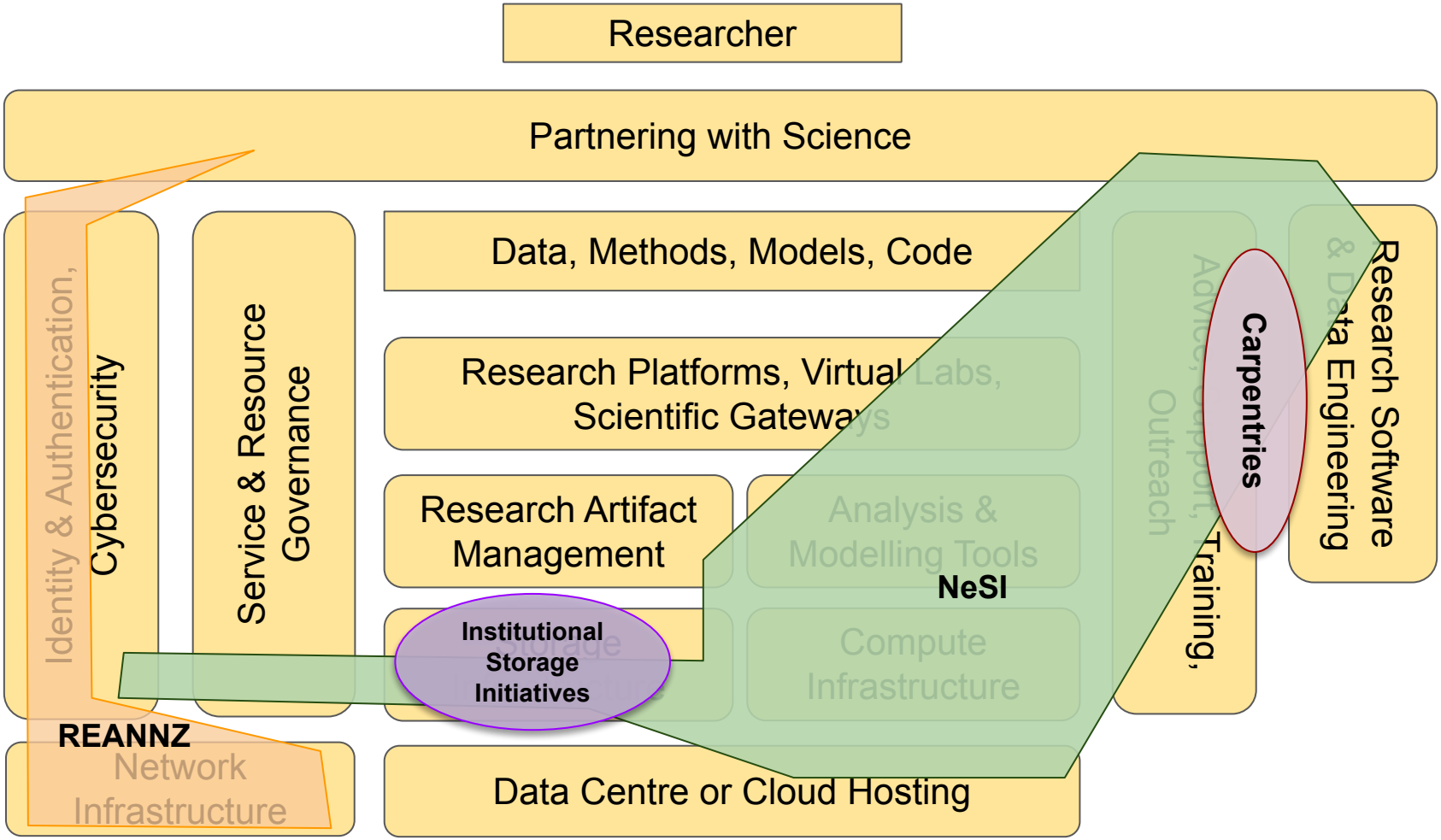
- ➢ Sensitive data platform (Tryggve) has been successful
- ➢ Nordic "Code Refinery" initiative also very promising

Services are available from several providers but this can create coordination and interoperability challenges for the researcher.



# NZ Players

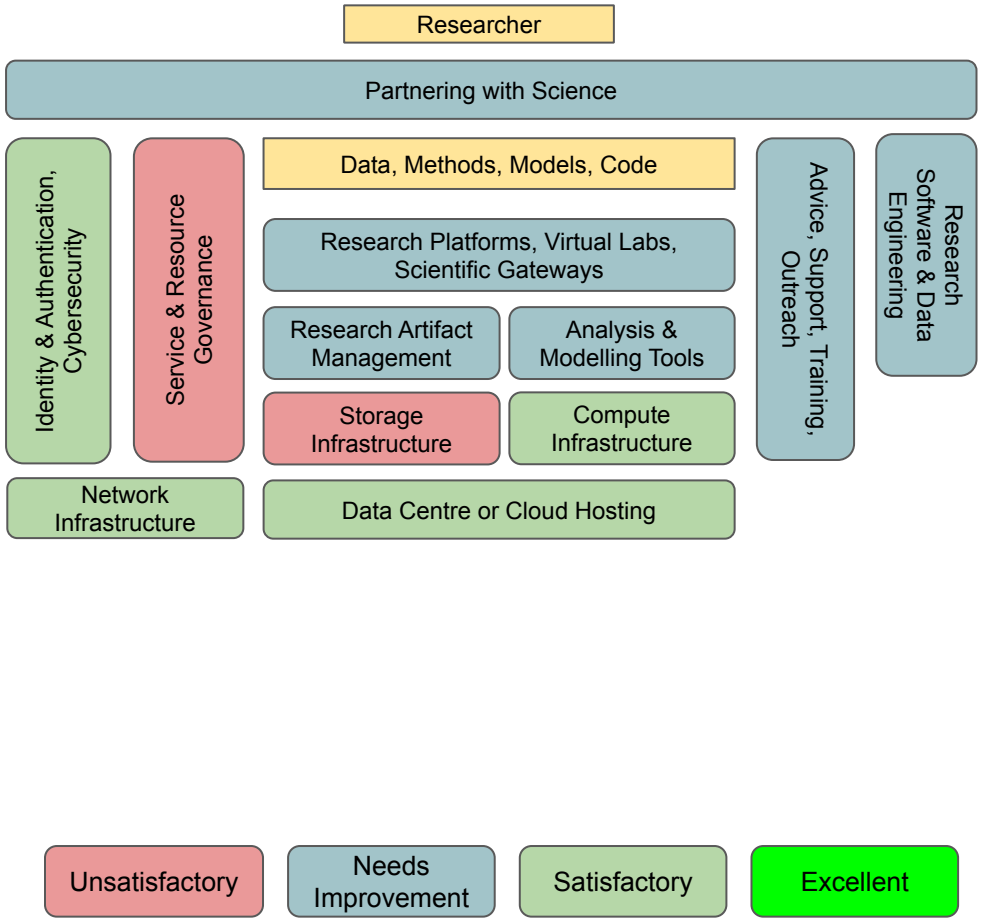




# NZ User Experience



- NZ operating under conditions of long periods of uncertainty and instability
  - New SSIF portfolio approach implementation early in life
  - New Genomics investment
  - Ongoing reviews of NREN
  - Challenges attaining broad membership and access
  - Unclear collaboration drivers
- No strategic supports for eScience tools (gateways, codes, RDM)
  - No agreed initiatives for RAM, although NRDP case made eResearch 2020
- MBIE resource allocation priorities are unclear, and allocation schemes are unsatisfactory to key stakeholders in investments



# International Exemplars



BLOODSTONE

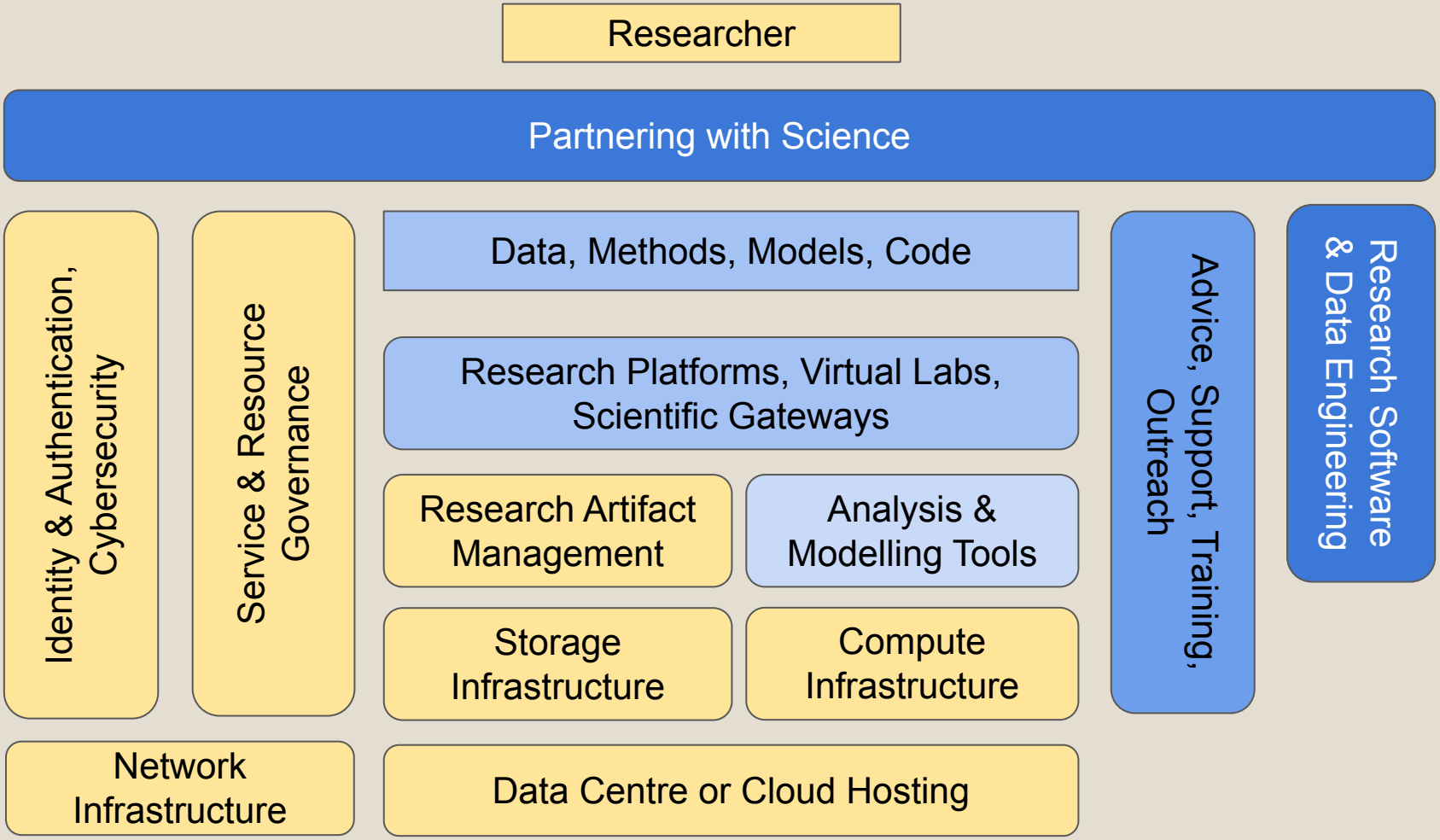
# Exemplar Jurisdictions & Providers



- Australia -- segmenting to some extent by user base
  - National Computational Infrastructure (NCI)
  - Australian Research Data Commons (ARDC)
- Netherlands -- segmenting by function
  - Netherlands eScience Centre (NLeSC)
  - SURF
  - DANS
- Finland -- a single provider
  - CSC
- Switzerland -- illustrating strong partnership with science
  - CSCS

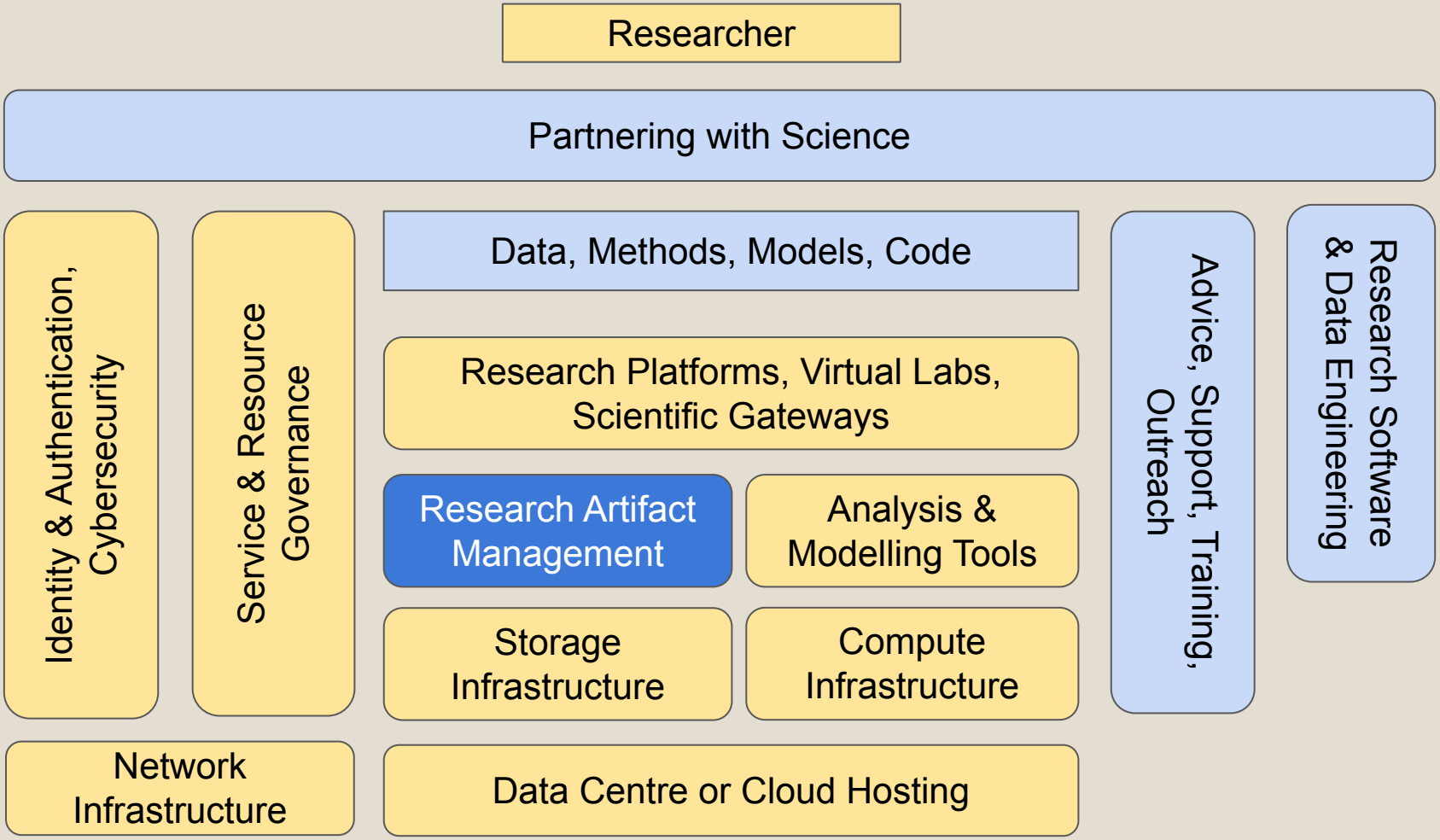
# Netherlands eScience Centre





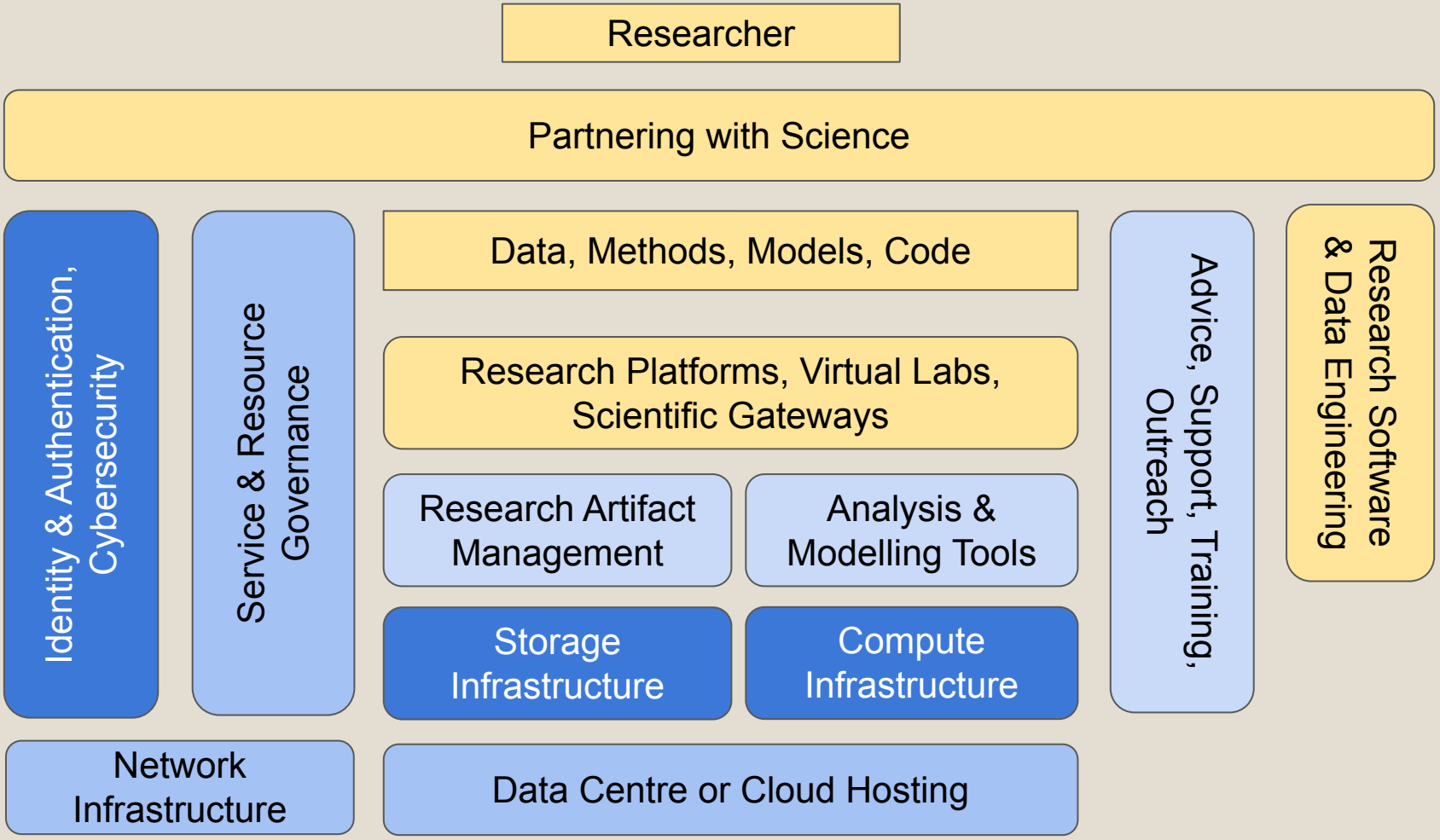
# Data Archiving & Networking Service (DANS)





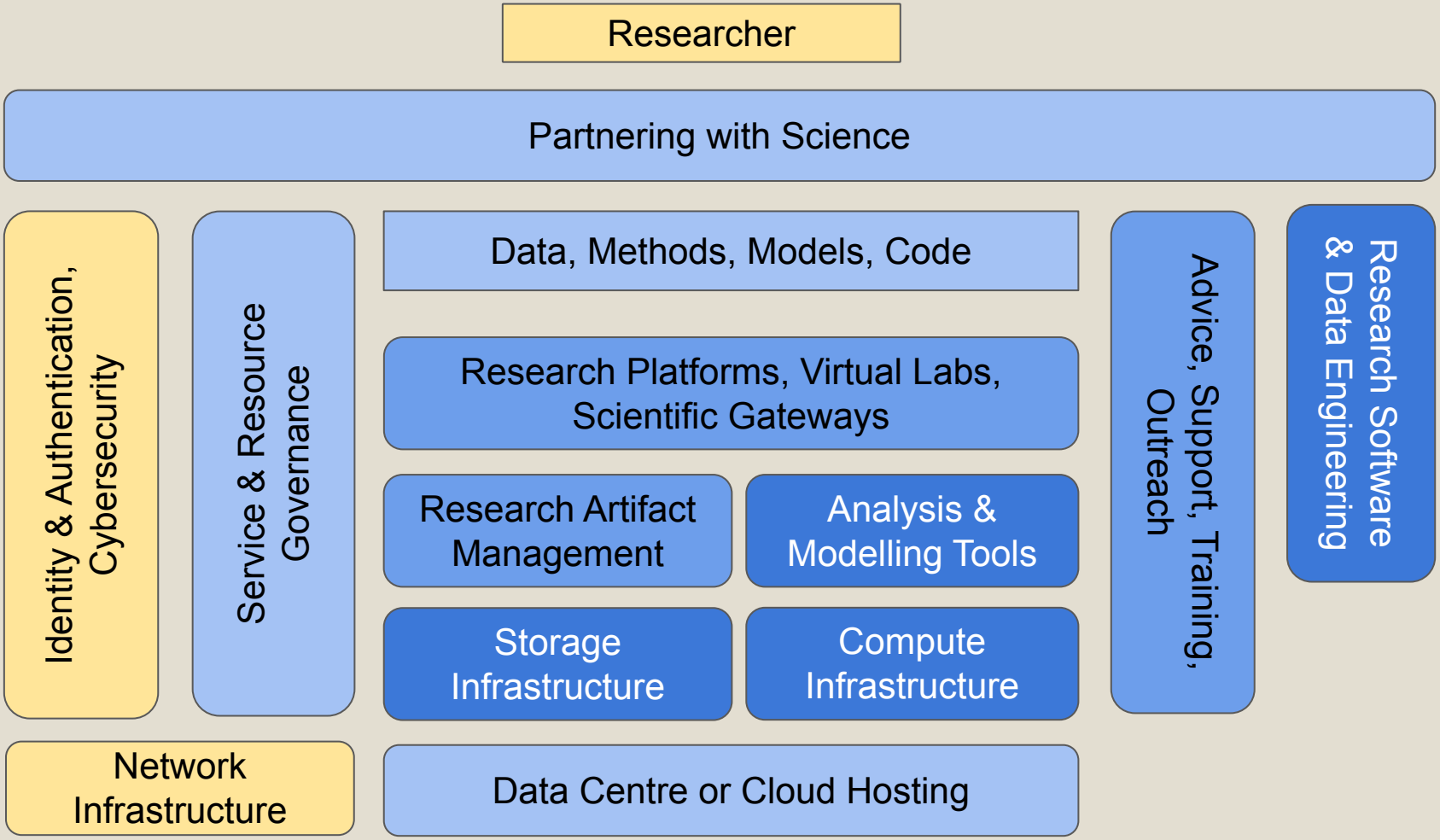
# SURF




Researcher
Partnering with Science
Identity & Authentication, Cybersecurity
Service & Resource Governance
Data, Methods, Models, Code
Research Platforms, Virtual Labs, Scientific Gateways
Research Artifact Management
Analysis & Modelling Tools
Storage Infrastructure
Compute Infrastructure
Network Infrastructure
Data Centre or Cloud Hosting
Advice, Support, Training, Outreach
Research Software & Data Engineering

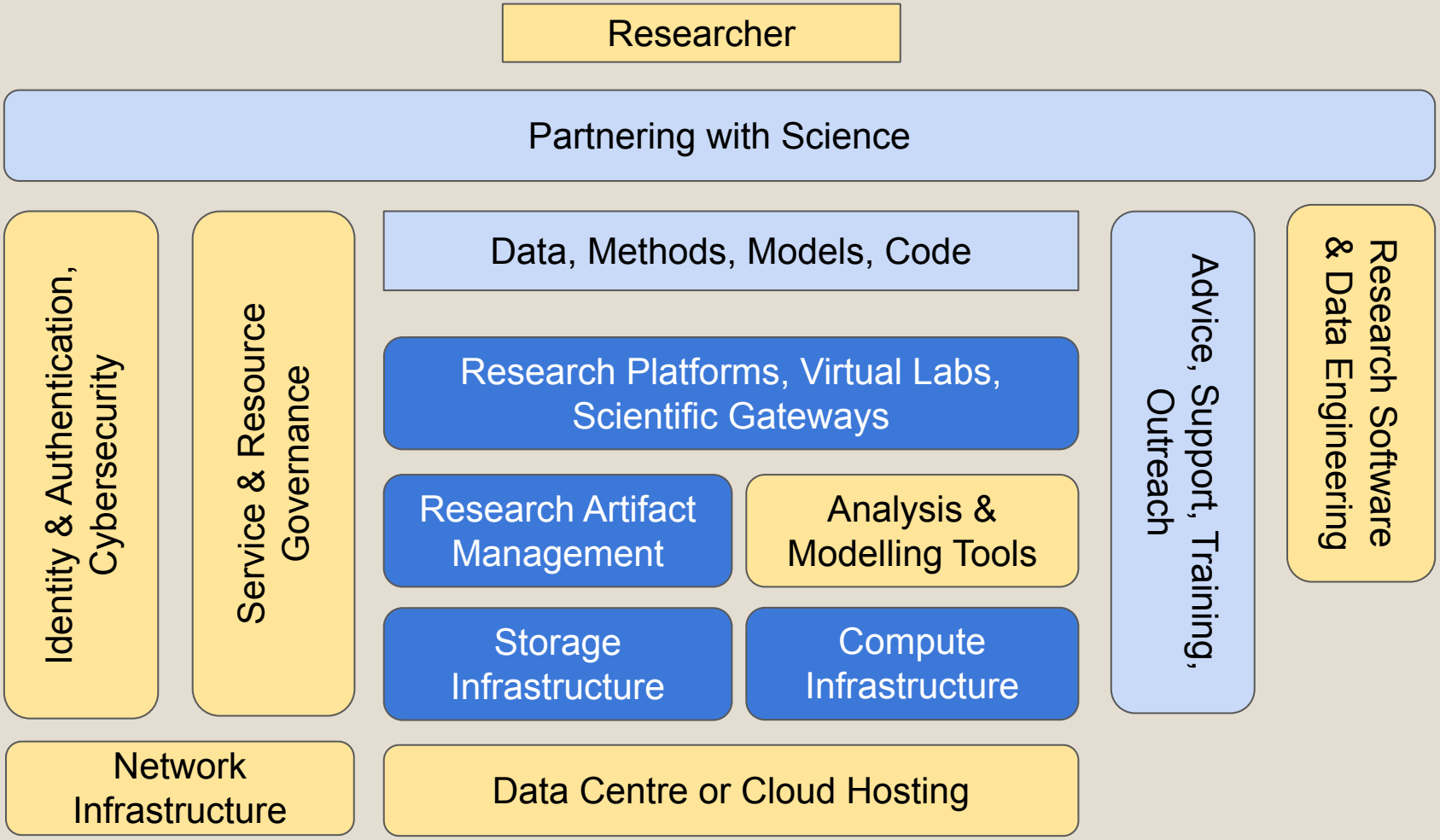
# National Computational Infrastructure (NCI)





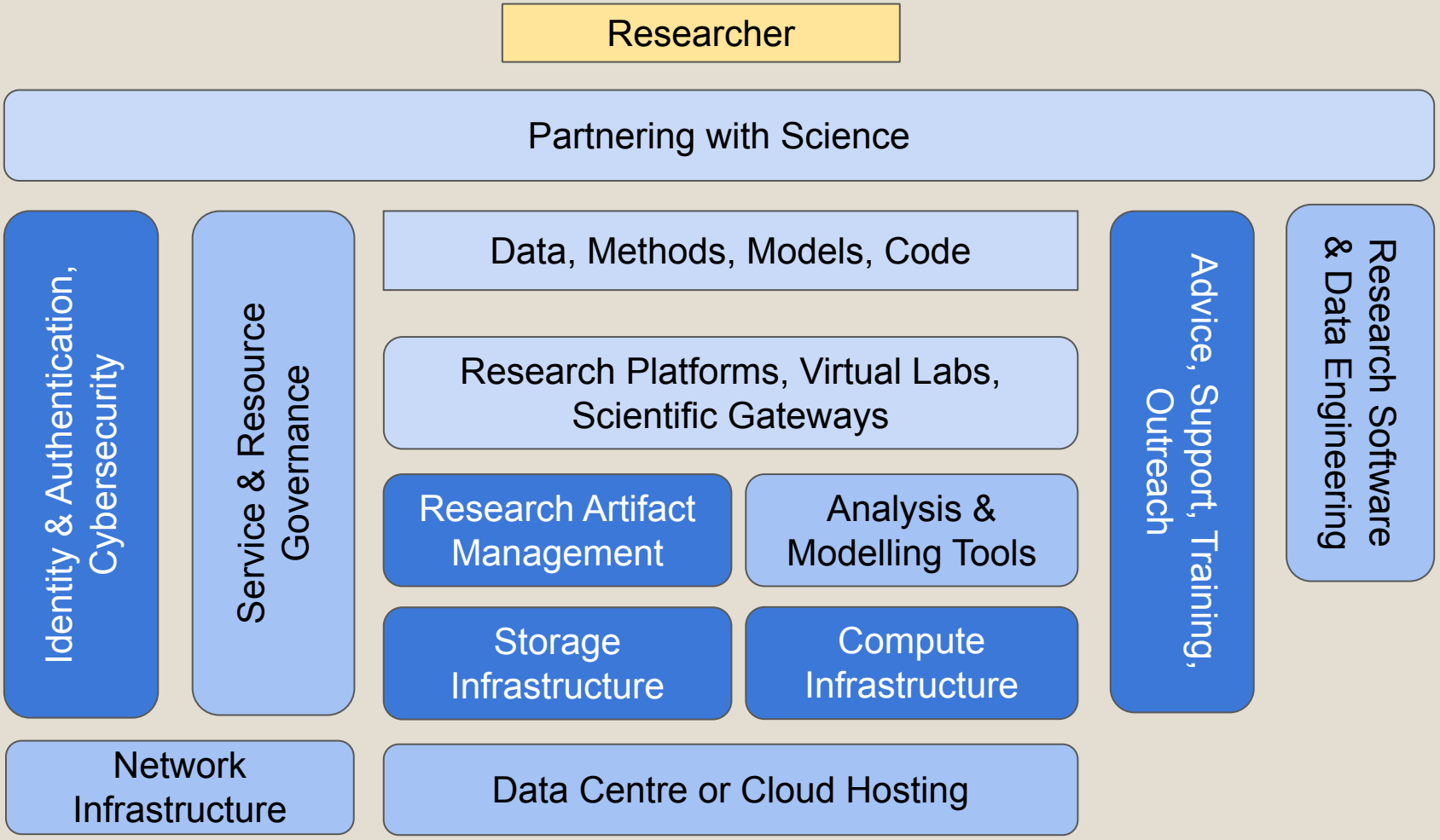
# Australian Research Data Commons (ARDC)





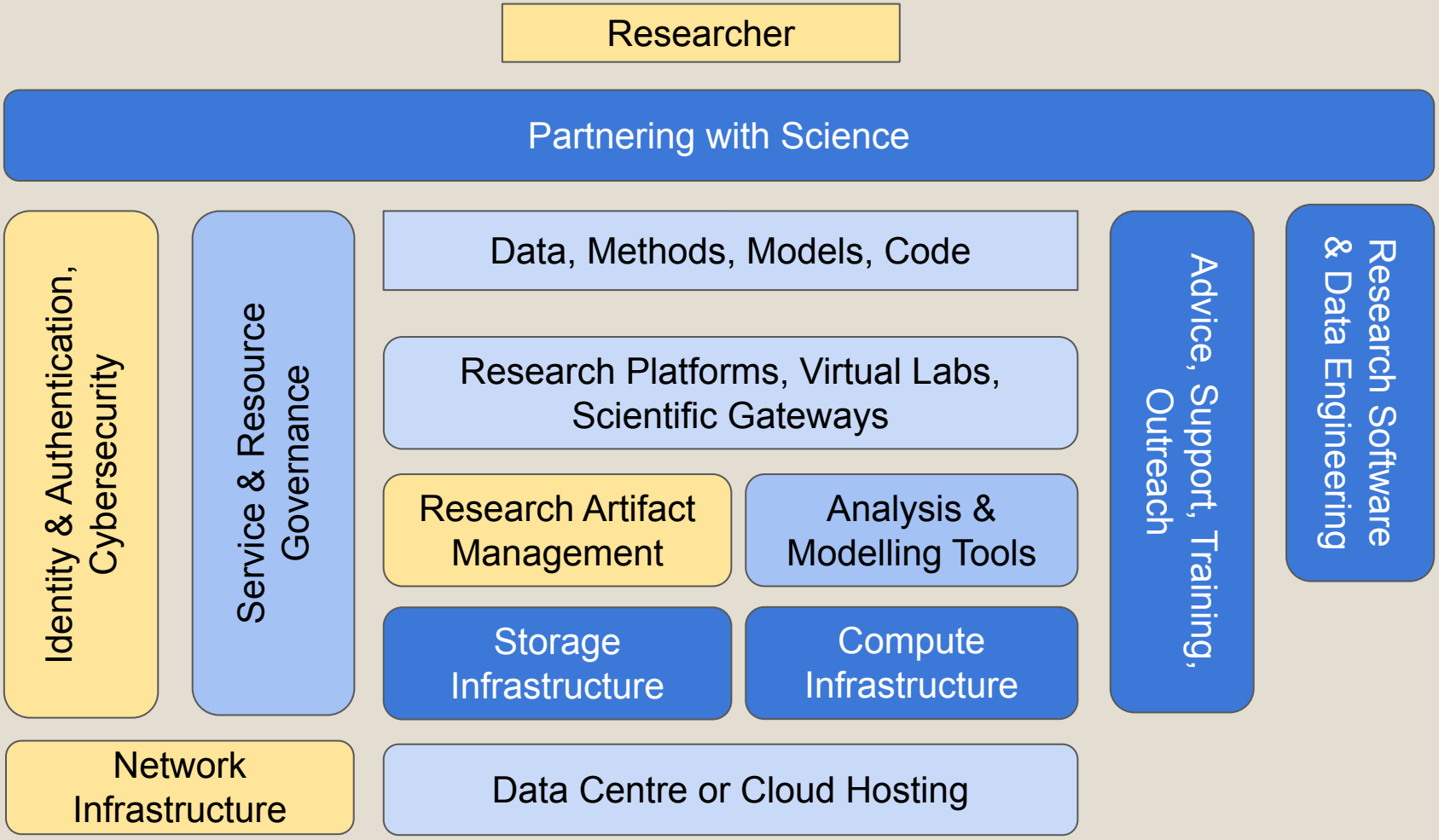
# Finland's CSC
## Effective Single Provider





# Switzerland's CSCS
## Strong partnership with science





# Main Observations for NZ



BLOODSTONE

# Main Observations for NZ eResearch



- Adopt a "Challenge Seeking" Posture
  - Partner with the most demanding science
  - Leverage the Full Value Chain to Supercharge Performance where needed
  - Maximize Utilization of all Technology using Platforms, Labs & Gateways.
- Balanced, Sustained Investment across the Ecosystem Components
- Organizational Design Should Allow Synergies among Components to be Optimized
  - Particularly in support of Performance Busting projects

# Thank You!



Mark Dietrich
Bloodstone Solutions Inc.
1-416-505-5079
mjdietrich@bloodstonesolutions.org



BLOODSTONE

# Stakeholder Perceptions & Concerns

# Institutions



- Meeting the most demanding eResearch challenges straining internal resources
- External resources could include commercial services, as well as "shared service" providers.
- Relying on external eResearch providers for services that are increasingly "mission critical" to the institution:
  - Raises risk when the institution does not fully control the external provider
  - Higher concerns when the provider is controlled by a competing institution, even if only superficially (e.g. through a lead institution).
- Institutions also promote their own eResearch capabilities as a competitive advantage in attracting faculty, industrial collaborators, government funding mandates, etc.  adding reputation management concerns as well.

# Government



- **Wants to be sure researchers have the eResearch tools and resources they need to be productive and reach global excellence**
- **Wonders about its role – as compared to the role of Institutions – in meeting these needs.**
- **Wants to enable important benefits of eResearch Investment**
  - First mover advantage/Strategic Opportunity
  - Economies of scale.
  - Economies of scope/Spillover effects
  - Networking effects/Spillover effects:
  - Reputational impacts

# Researchers



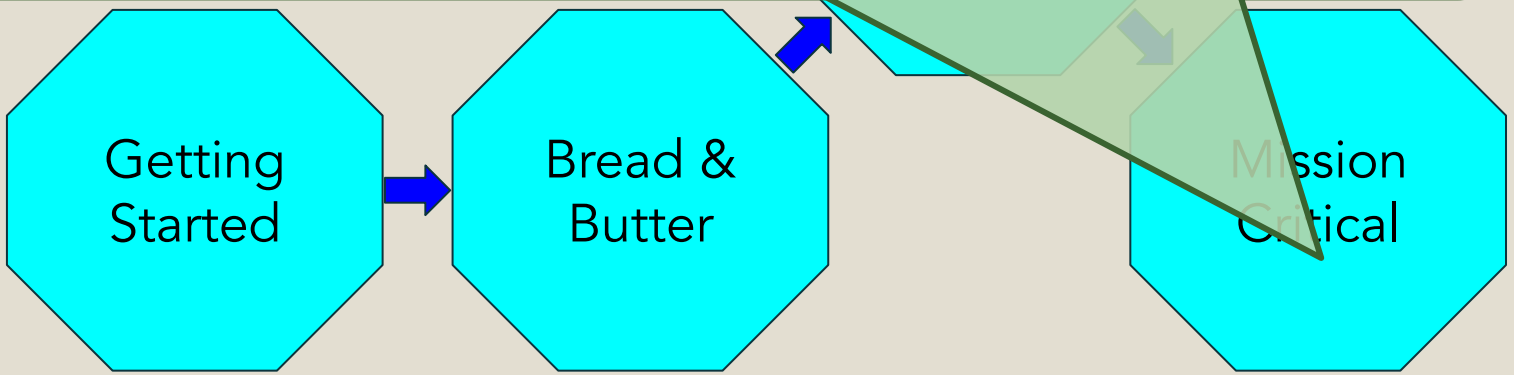
Researchers are properly focused on their own projects -- if a shared services eResearch provider does not provide appropriate levels of service, researchers can complain that the investment in a shared service was wasteful.

Leading researchers argue that a separate eResearch service, by definition, will never be able to meet their high expectations and that resources should flow directly to the researcher.

# Benefit Opportunities: Mission-based Research

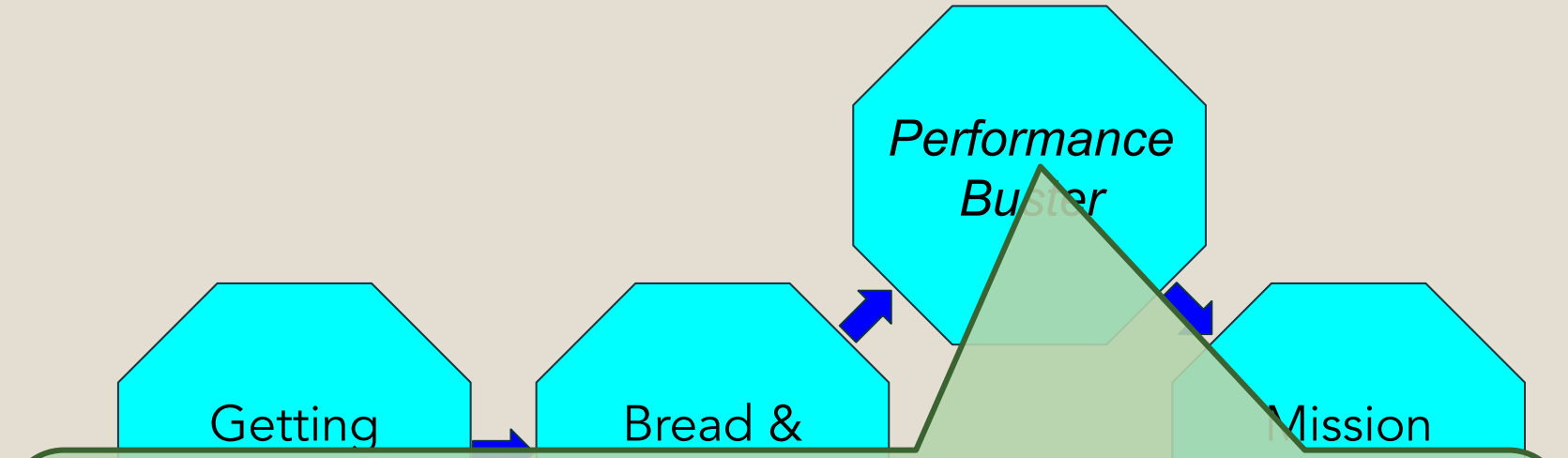


- First Mover Advantage/Strategic Opportunity
- Cannot themselves take advantage of economies of scale in infrastructure (i.e. they have to be fully funded), but can seed such economies for other segments.
- Economies of scope may be possible, as long as the "demands" of the production environment can be respected when using more widely used tools. Introduce best-in-breed tools as a spur to innovation.
- Collaborate with computational/data scientists to supercharge performance
- Strong "science led" use cases help make responsive eResearch services highly relevant and improve adoption (economies of scale) elsewhere
- Service-orientation critical -- robust processes and accountability



# Benefit Opportunities: Pushing the Envelope



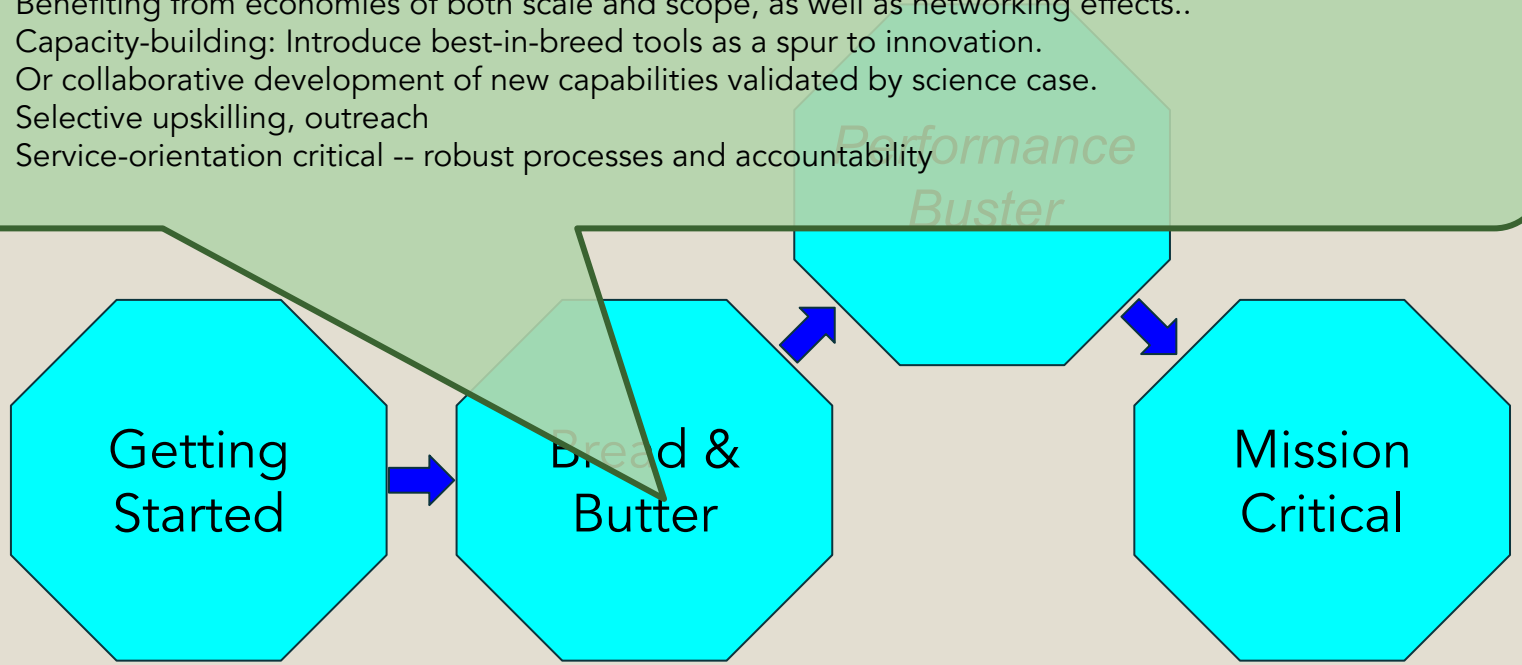



- First Mover Advantage/Strategic Opportunity
- Benefit from economies of scale seeded from other segments (e.g. mission-based science)
- Economies of scope may be possible: eResearch providers can deliver best-in-breed tools that also serve a demanding research environment.
- Or collaborative development of new capabilities validated by science case.
- Either way, strong "science led" use cases help make responsive eResearch services highly relevant
- Collaborate with computational/data scientists to supercharge performance
- Success here will improve adoption (economies of scale) elsewhere...

# Benefit Opportunities: Bread and Butter



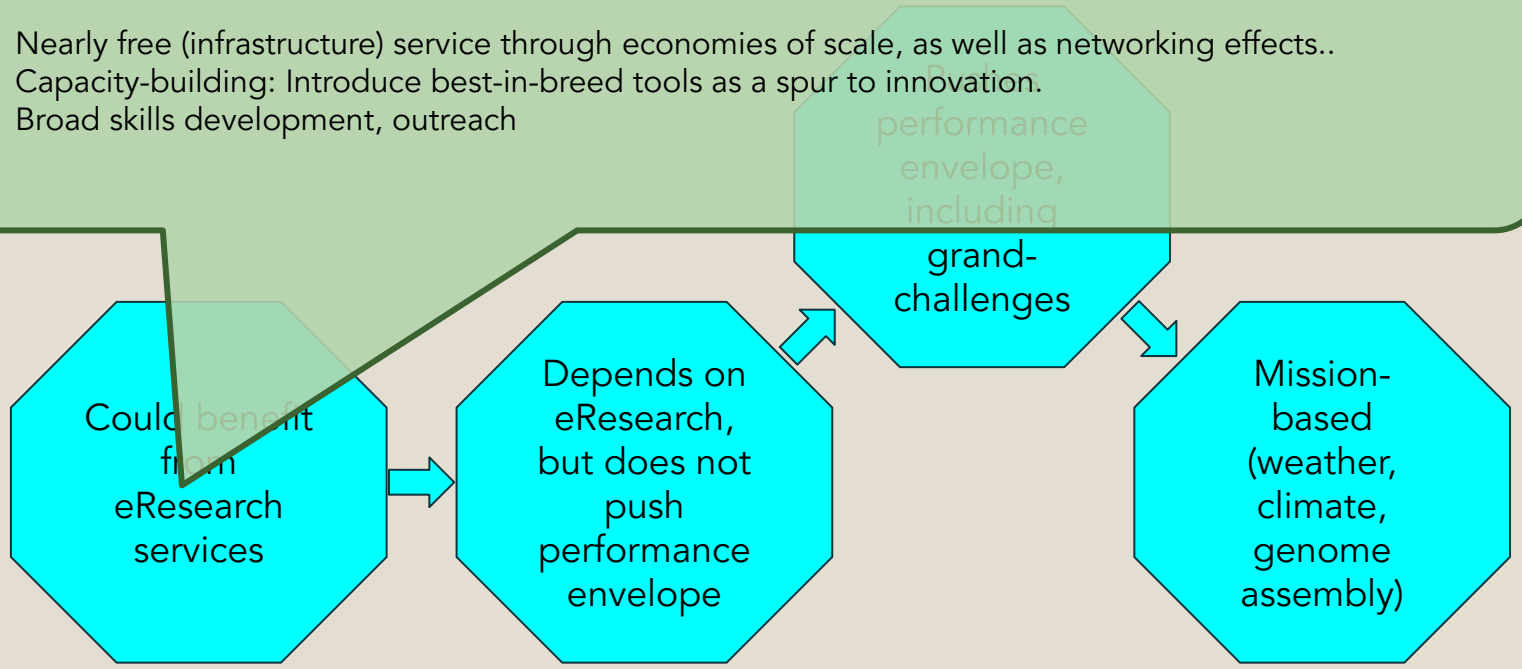
- Benefiting from economies of both scale and scope, as well as networking effects..
- Capacity-building: Introduce best-in-breed tools as a spur to innovation.
- Or collaborative development of new capabilities validated by science case.
- Selective upskilling, outreach
- Service-orientation critical -- robust processes and accountability



# Benefit Opportunities: Getting Started

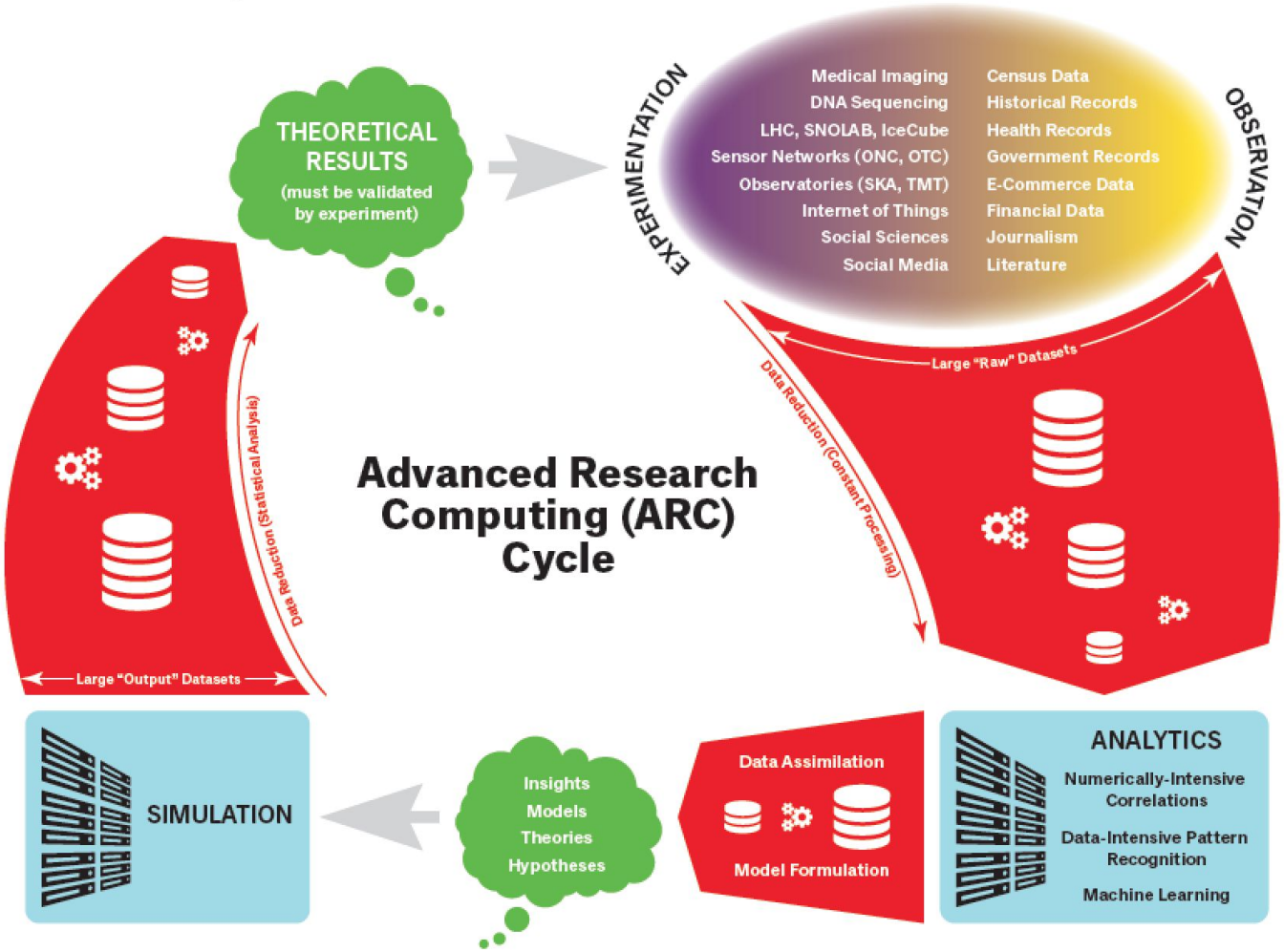


- Nearly free (infrastructure) service through economies of scale, as well as networking effects..
- Capacity-building: Introduce best-in-breed tools as a spur to innovation.
- Broad skills development, outreach



# Convergence of HPC and Data Analytics





# eREF: Recognizing the Role of Software

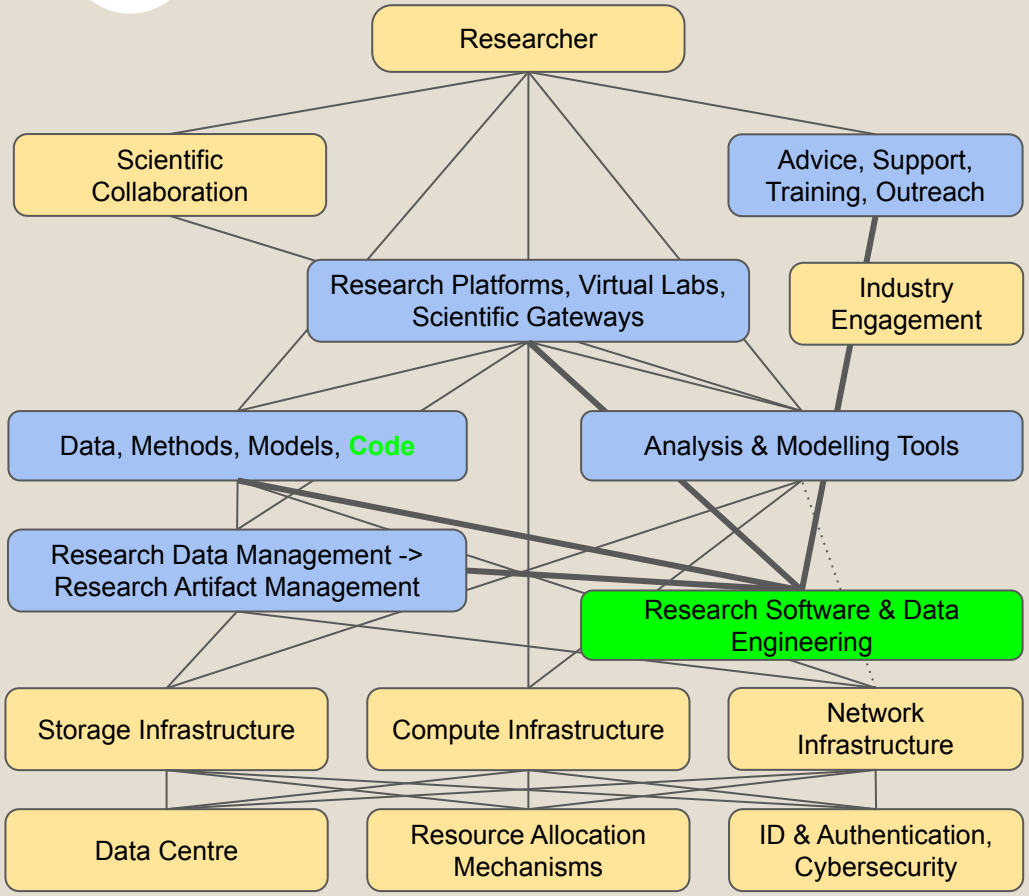


Software is integral to many of Components in the eResearch Ecosystem Framework (see highlights at right)

- Analysis & Modelling Tools ARE software, which needs to be developed, tested, validated and maintained in a professional way.
- Data, Methods, Models INCLUDE code to ensure sharing and reproducibility
- Gateways ARE software systems
- Code is one type of artifact that needs to be MANAGED by Research Artifact Management, with repositories and version control.
- Best practice methods in software development, testing, validation, bug management, etc., need to be TRAINED.

Software Sustainability Institutes have been established in the US, UK and elsewhere to promote best practices in research software development and maintenance.

The Research Software Engineering Association was established in the UK to recognize the contribution that professional research software developers make to the eResearch Ecosystem. Similar initiatives are underway in several other countries.



# Other tweaks to the Framework



Scientific Collaboration → really about "partnering" with science, rather than expecting everything that "collaboration" implies (e.g. article authorship)

Industry Engagement → Outreach and support to a particular audience; merged.

Research Data Management → Covers more than data: Research Artifact Management

Resource Allocation Mechanisms → Covers services (not just tangible resources), policy-alignment and measurement of utilization: Service & Resource Governance

Data Centre → addresses questions of on-premise, vs. cloud, vs. hybrid hosting and provisioning.

Note: Researchers are the users, and they bring their "Data, Methods, Models and Code" to the ecosystem to make discoveries, so these "components" of the ecosystem now have slightly different symbols in the diagram.

# Providing eResearch Services



Informed observers agree the framework is useful:

- They confirm that much more is required than just compute or network infrastructure – the value chain must be complete before value can be created.

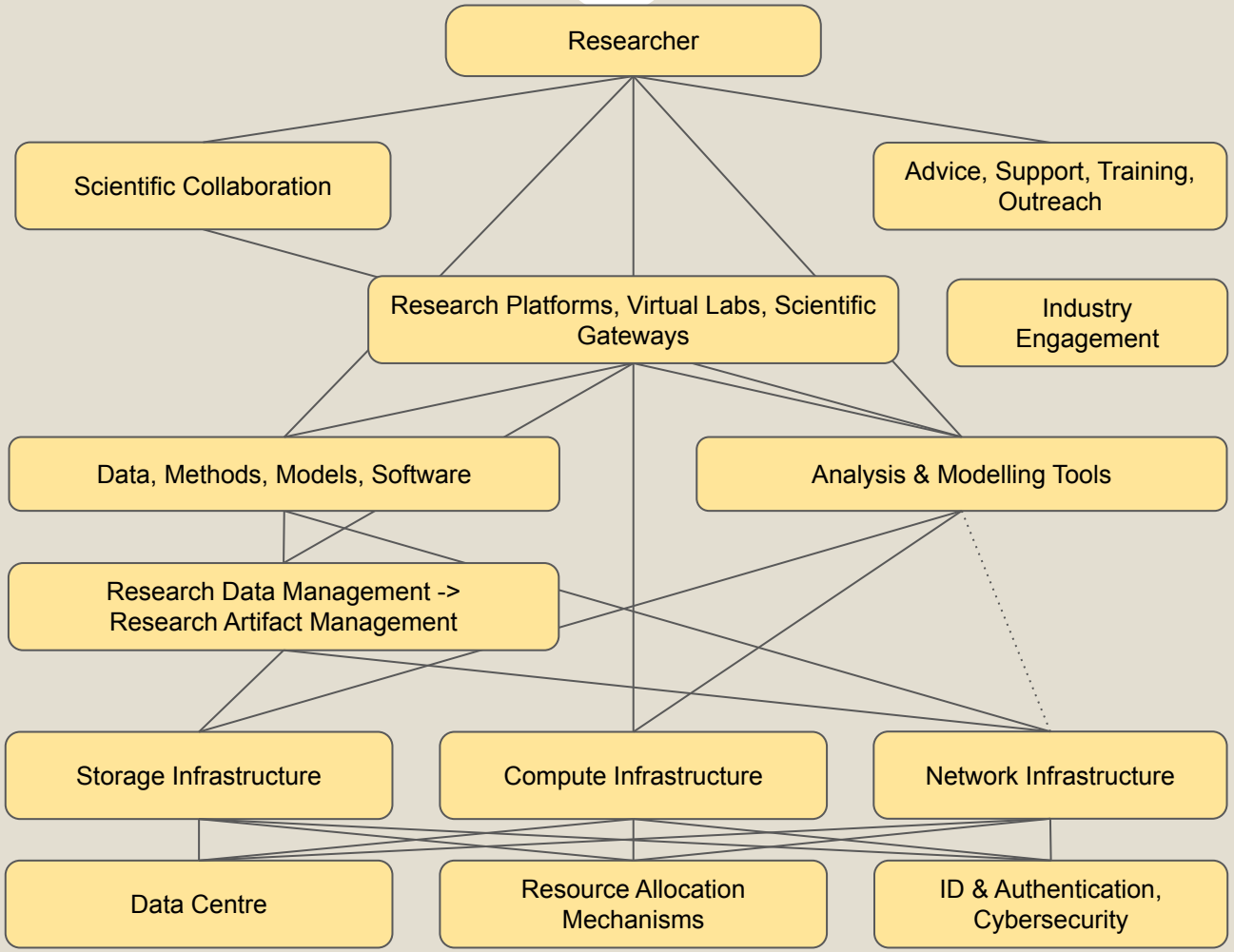
They also confirm:

- All of these activities continue to evolve
- Researcher needs continue to evolve
- Providers continue to evolve
  - Do It Yourself is popular (easy to underestimate what is needed)
  - Commercial offerings are alluring (cloud, online services are marketed as easy to use and nearly free)

# eResearch Ecosystem Framework (eREF)

## Component by Component Key Learnings and Observations relevant to NZ

88



# Comments on Australia



NCI: Needs to address tension of partner missions vs. services to broader community

Australian Research Data Cloud (ARDC)

- Integrating Nectar, RDS, ANDS → science-focussed eResearch capability with real capacity to get work done.
- Integrated mandate still includes distributed cloud computing, supporting "bread and butter" campus requirements → there could be tensions over local partner expectations vs. wider services and sustainability

Mandates of ARDC, NCI and Pawsey still largely overlap.

- E.g. NCI's "National Research Data Collection" similar to ARDC
- Recent funding announcements → collaborative division of labour possible.
- "One shot" nature of funding, below requested amounts → expect renewed funding crisis and mandate conflict in ~3 years.

# Comments on Canada



Leadership Council on Digital Research Infrastructure (LCDRI) proposed a federated investment model in 2017:

- Federal government regularly funds ~75-80% of major ARC systems (which includes co-located persistent storage), along with a national "office" to coordinate the ecosystem and lead investment planning.
- Individual institutions host ARC systems through fully funded contracts/SLAs
- Institutions and provinces fund ~75-80% of local support
- Governance of ecosystem not clear:
  - CANARIE seeks a greater/leading role
  - Compute Canada suffers from severe internal governance challenges
  - RDM is seen as key, but has become a "political football"

Government has supported the LCDRI's proposal to the extent it has announced C$572 million over 5 years for DRI in February 2018.

- Ongoing deliberations re: to whom or how these funds will be distributed

Significant investments have also been made in AI, originally expected to drive additional e-infrastructure investments.

# Comments on Singapore



As a small nation, Singapore's eResearch community is correspondingly small, and the players appear to work together well.

- ➢ Higher level "eScience" components (RAM, gateways, software tools) do not appear to be addressed

# Comments on the Netherlands



As a small nation, the Netherlands' eResearch community is correspondingly small, and the players appear to work together well.

- ➢ Good balance between national and institutional activities
- ➢ Recognition that "eScience" investments (gateways, analysis tools) may be somewhat uncoordinated
  - Definite concerns about sustainability

High level of international collaboration by researchers themselves, as well as close proximity of alternative resources through PRACE/EOSC and the EU, reduces need for a self-sufficient ecosystem.

# Comments on Sweden



Swedish National Infrastructure for Computing (SNIC)

- ➢ Evolving from a light coordinating role, responsible primarily for the national allocation process.
- ➢ Has proposed a strategy for SNIC 2 that suggests consolidation from the current 6 centres, but it is unclear if this will happen in practice.
- ➢ Has also been challenged by outside observers to offer a broader vision for SNIC's role in the Swedish eResearch ecosystem. Unclear if the community feels the need for this.

Swedish eScience Research Centre (SERC) set up to enable "smarter" eResearch activities, but role appears to be primarily one of networking/brokering

Sweden's research funding agency (VR) does not seem concerned about effectiveness of the broader ecosystem -- this may reflect an expectation that Swedish institutions need to step up to these responsibilities.

RDM seems to be a particular gap, especially compared to other countries -- policies require data deposition and access, but solutions are being left to institutions.

- ➢ KTH has taken a leadership role in this area, but does not have a mandate or funding.

Access to PRACE/EOSC and other EU resources (eg in software), reduces level of concern

# Comments on Switzerland



CSCS has a national mandate, but has also grown organically through cooperation with several universities

- Tensions with EPFL may limit broader initiatives.
- There is no strategy for higher level "eScience" components (RAM, gateways, software tools), although CSCS has provided good solutions when asked.
  - Reactive approach is not sustainable

# Comments on the United Kingdom



UK ecosystem

- Marked by large number of strongly competitive incumbents
- With Brexit, research and eResearch are the least of their problems
- Gaps in several areas (compute, RAM, gateways) suggest there is an opportunity to "leapfrog" other countries, but
- Slow creation of a new umbrella funding organization, UKRI, that has started a comprehensive roadmap for research infrastructure.

Even with a clear roadmap and compelling case for investment, likely insufficient funding to do more than "muddle through"...

# Comments on Nordic Region



- Norway, Denmark, Finland: eResearch providers have similar wide scope:
  - CSC Finland, Uninett Norway are single providers
  - DeIC Denmark is a federation
- For Nordic eInfrastructure Collaboration (NeIC),
  - Member organizations operate independently
    - E.g. Investments are not complementary
  - Members elect to participate in joint initiatives
    - The most successful initiatives are driven locally and then opened up to the broader community.
    - Not all are successful:
      - E.g., Dellingr: cross-authentication, as a step to single sign on, has not achieved objectives.

# Comments on New Zealand



NZ ecosystem

NZ is a relatively late investor in creating an eResearch ecosystem, with first investments starting 2006, some 10-30 years after other jurisdictions

- Characterised by being a late and risk-averse investor in infrastructure
- As a demand-side driven economy and with a full-cost research funding system, NZ typically faces a challenge in building new research capabilities and a broader base of skills due to a preference for non-interventionist approaches
- Gaps in several areas (broader types of compute, storage, RDM, gateways, skills training, data/compute intensive science collaborations) suggest there is an opportunity to "leapfrog" other countries
- Partly impeded by hesitance in making investments due to a deferred MBIE roadmap for research infrastructure

With NZ's small scale and lack of incumbency from its short history of investments in this area, conditions suggest significant opportunity to leverage and move ahead of other nations

# Australia
## "Horizontal Segmentation"



- Organizational model builds on segmentation of researcher needs, assumes distinct resources and services are needed in the different segments, that there are no economies of scope.
- NCI: Highly specialized eResearch provider, tightly focussed on service to its partners
  - Science partnering comparable to CSCS
  - Integrating meaningful research activities through platforms, research artifact management, modelling tools and underlying infrastructure
  - Interest, but uncertain ability, to broaden reach of successful eResearch capabilities past current partners
- ARDC: Promising merger of ANDS, RDS & NECTAR
  - Recognizes real synergy of data and analysis, value of co-locating storage and compute
  - Leverages more "commodity" infrastructure, loosely coupled technically.
  - Mandate overlaps with NCI (and others); ARDC must also serve many stakeholders
  - Unclear scope for software investment (e.g. in analysis & modelling tools) that gives life to data

Key Takeaways for NeSI:

- Providers will lose time while they resolve mandate overlaps -- better to address at the point of establishment
- Broader exploitation of valuable competencies limited by scope and budget set by partner
  - Limits on ability to broaden uptake and reuse
- Development and maintenance of tools, related software engineering community, needs to be explicitly resourced and encouraged

# The Netherlands
## "Vertical Segmentation"

99

- SURF: Integrated eInfrastructure service provider
  - Mandate broader than research -- also includes shared IT services and federated purchasing across educational sector.
  - More limited engagement with researchers than in other countries
  - More limited role in support, training & outreach than in other countries
  - Complements EU investments (e.g. in PRACE) so not seen as supply constrained
  - At the same time, SURF has strategic intentions to increase this engagement
- NLeSC: eScience partner collaborating directly with many research/eResearch projects:
  - Deliberately set up separate from SURF to ensure trusted advisor/partner role; SURF services do not have to be used; NLeSC can help to focus future SURF service offerings
  - Mechanism is funding, which in turn is used to embed NLeSC experts
  - No mechanisms for sustaining resulting capabilities.
- DANS: Focussed research artifact management provider
  - Complements EUDAT, therefore seen as effective service provider with well-defined mandate.

Key Takeaways for NeSI:

- What seems like natural functional segmentation can leave tensions among providers unresolved
  - Poorly defined mandates
- Efforts to address longer term sustainability could appear like mandate expansion/competition
- Current success of NL ecosystem may be a function of modest demands, with more extreme needs met by EU resources

# Finland's CSC
## Effective Single Provider

100

- Integrated eResearch service provider
  - Extensive service catalog: understands that "one size does not fit all"
  - Balanced focus on service to researchers, regardless of solution
- Collaborative role in dozens of research/eResearch projects:
  - Focus on arctic and marine science
  - Champion of EUDAT (EU's Research Artifact Management) and Elixir (EU life sciences collaboration)
- Mandate actually broader than eResearch -- also includes IT services across educational sector, and CSC engages in fee-for-service work in both research and other public-sector areas.
  - Revenue diversity is good, but highlights risk of being seen "just" as shared service IT provider
  - Balanced by CSC's focus on relevance and competency to research sector, built over decades.

Key Takeaways for NeSI:

- Balance benefits of shared service organization with intense focus on the customer.
- For research, embrace multiple solutions if that is what the market demands.

# Switzerland's CSCS
## Strong partnership with science

101

- Takes a prominent role in several research projects:
  - Human Brain
  - Materials at Exascale (MaX)
  - CHIPP/WLCG Worldwide LHC Computing Grid
- Direct support of computational science needs of MeteoSwiss
- Significant personnel dedicated to computational science (23+), software engineering (17)
  - ~½ of non-management personnel at CSCS
- Direct role in improving, optimizing a range of computationally intensive modelling tools (including weather, molecular dynamics)
- Limited advocate for research platforms except in direct support of scientific collaborations above
- Earned trust of other Swiss institutions to enter into hosting and support arrangements for requirements outside of ETH Zurich

Key Takeaways for NeSI:

- Partnering with science can increases eResearch provider's relevance and competency, but requires significant investment in computational science and software engineering.
- CSCS wants to broaden adoption beyond current customer base
  - Considering product management techniques -- but this study suggests leveraging "science leadership" would be more successful.

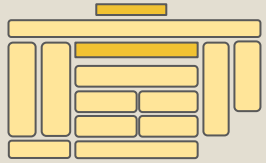
# eREF: Researchers, Data, Methods, Models and Code



Key Learnings:

- Worldwide, more disciplines, and more researchers within each discipline, are adopting and integrating eResearch techniques into their activities
- "data, methods and models" have always been a critical aspect of research:
  - In eResearch, Data, Methods and Models -- as well as Codes -- are more clearly identified, manipulated and shared
  - Data volumes, the size of input and output data sets, extent of data distribution, growing rapidly
  - Model complexity growing exponentially in quest for accuracy: mesh resolution, time steps, numbers of particles, atoms, actors, equations, dimensions
  - Methods are encapsulated in workflows and techniques like ensemble statistics, parameter sweeps
  - Efficient Code allows more complex Models to execute in realistic timeframes.

Observations:



- Outreach to / adoption by / diffusion among -- researchers -- are critical.
- Create centre(s) of expertise...
  - Able to navigate the changing landscape of data, methods, models, codes...
  - That will draw in collaborators who will
    - sustain the expertise,
    - build reputation, and
    - motivate others to adopt.

# eREF: Partnering with Science

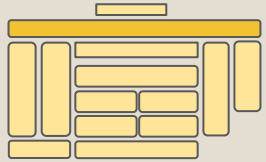


Key Learnings:

- Deliver value to the most demanding "client" researchers, who often set the tone of the community in embracing or rejecting eResearch providers
  - NLeSC "We want to be as good as you"
- Be guided by working scientists in setting priorities and making choices
  - Nordic, Finland: more than just "IT managers"
- Focus on collaborating with "client" research groups, rather than advancing a separate science agenda

Observations:

- Science Advisory Board or similar
- Chief Science Officer / Science Leads
- Visiting posts / Fellowships
- Foster mobility: At NLeSC 40% of positions are time-limited
- Longer term positions with deep expertise
  - Data science/analytics
  - Computational science
  - Possibly embedded in strategic science areas (e.g. cross appointments in genomics, climate, ecology, ...)

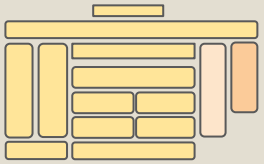


# Research Software & Data Engineering



Key Learnings:

- Software is integral to many of Components in the eResearch Ecosystem Framework
- Software Sustainability Institutes (SSI) have been established in the US, UK and elsewhere to promote best practices in research software development and maintenance.
- The Research Software Engineering Association was established in the UK to recognize the contribution that professional research software developers make to the eResearch Ecosystem. Similar initiatives are underway in several other countries.
- For Challenge Seeking Ecosystems, investment in RS&DE can approach 50% of all FTEs



Observations:

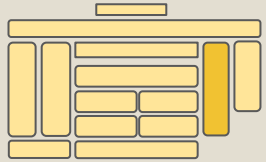
- Despite acknowledge and investment, best practices are not well defined
- In most cases SSI investments use the techniques of "communities of practice" to advance the area. This is fine, but more robust practices will be critical.

# eREF: Advice, Support, Training, Outreach



Key Learnings:

- "Adoption" of eResearch techniques is low (typically ~10% of researchers), yet the benefits are great (correlation with increased scientific impact, widely transferable skills)
- Traditional training and support mechanisms do not scale well



Observations:

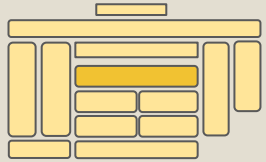
- Focus resources on "scaleable" capabilities
  - Curriculum development
    - Cover the full range of eResearch activities: e.g. "Introduction to Gateways"
  - Train-the-ambassador-and-the-trainer: e.g. Campus Champions in US
  - Skills testing/certification
  - Knowledge base
- Extended support assignments ("embedding")
  - Explicit capture of outcomes (e.g. code speed-up, time-to-result) that are then promoted
  - Teach researchers to fish -- promote the research benefits of learning to fish
- Lead by creating non-academic career paths (a la "Research Software Engineer" in UK).
  - E.g. RSE internships at NeSI, permanent positions

# eREF: Platforms, Labs, Gateways



Key Learnings:

- Gateways...
  - lower the barriers to adoption of eResearch techniques
  - insulate users from many details raised by other components
  - efficiently amplify usage for providers (gateway accounts ~3-10x more than direct infrastructure accounts)
  - build communities and support collaboration through gateway features and self-support, rather than eResearch personnel



Observations:

- Build internal capabilities to support selected key gateways
  - Participate/contribute to the selected open source gateway communities
  - Integrate with other components (e.g. research data management)
  - Use selected gateways as platform for training and service delivery
    - e.g. "Orchestrating Data Analytics on Mahuiki using Jupyter Notebooks", "Data Sharing and Collaboration using Airavata"
- Don't "judge" non-supported platforms
- Focus on enabling science rather than promoting products
  - "Product Management" (managing the features and functionality of gateways and other tools) is essential for NeSI to remain effective internally, but not interesting to the research community

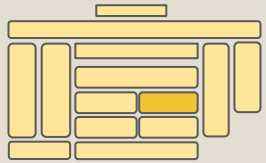
# eREF: Analysis & Modelling Tools



Key Learnings:

- Expertise in these tools elevates a "shared IT service" to an eResearch provider
- The range of science requirements is broad, challenging a modest team.
- Most demanding science requires expertise on underlying compute infrastructure
- Leading analytics environments evolving to look more like HPC environments
- Processing of distributed data an important frontier as datasets become too big for one system

Observations:



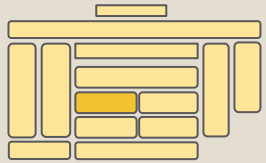
- Focus on fundamentals and culture
  - Help researchers optimize their own codes
    - Point them to community codes when appropriate
  - Support local research communities as they contribute to international community codes
    - Encourage use of collaborative test/validate platforms
  - Provide the right environments for more packaged codes/tools
    - Hardware/OS/containers/etc.
- Link expertise to Compute Infrastructure to enable supercharged performance when needed
  - Support special purpose compute investments as investigations of new technology
- Contribute to and support the broader research software engineering community, including collaborative platforms for test and validation.

# eREF: Research Artifact Management



Key Learnings:

- Few countries have clear strategy, fewer funding the effort
- Nevertheless a key focus of "Open Science" initiatives
- Ultimate objective is to change research culture, rather than adopting specific tools
  - "Data Champions" embedded in faculties (Cambridge, TU Delft)



Observations:

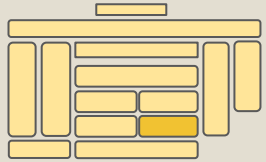
- Solutions may ultimately sit in other components:
  - Policies around preserving Data, Methods, Models should be owned by Researchers
  - eResearch expertise and DevOps can automate better practice artifact management processes in middleware, particularly Gateways.
  - Underlying Storage Infrastructure features should enable those processes
  - eResearch can Support/Train data champions
- Help the research community make good choices, don't push a "perfect" solution
  - Google vs. Lycos vs. Altavista

# eREF: Compute Infrastructure



Key Learnings:

- Total science-driven demand for compute capacity continues to grow exponentially
- Explosion of architectures (HPC, HTC, accelerators), hosting strategies (on-, off-premise, hybrid), precision (double, single, integer, etc) continues -- there are no obvious "best practices"
  - Hardware increasingly being optimized for applications → "appliances"
- Pace of price-performance improvements expected to continue, even without Moore's Law
- Where benefits of scale can be achieved through consolidated investment in common technologies, these can be offset by associated governance and funding challenges
  - Flagship investments are best justified by science leadership/first mover advantage, rather than trying to address alleged market failure
  - EU countries have benefitted from Tier 0 investments (PRACE) to relieve pressure for large local HPC investments.



Observations:

- Leading research projects/missions should guide flagship investments
  - eResearch providers focus on maintaining local awareness of international developments
  - Guide investments to enable use by broader base, and then create ecosystem supports to ensure that usage.
- Support interoperability across ecosystem, through architecture leadership, middleware
- Sustained investment is needed to keep up with demand, assuming performance improvements continue (likely through use of new architectures)
- Be open to Tier 0 partnerships with other countries

# eREF: Storage Infrastructure

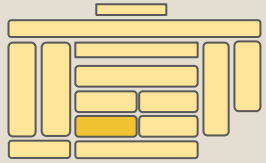


Key Learnings:

- Volume of active research data growing exponentially, in total and for individual data sets
- Volume of archival data also growing exponentially, but lagging active data
- "Kryder's Law" pace of price-performance improvements expected to continue
- Growing volumes of data will drive co-located compute investment
  - Bring the compute to the data
  - Rethinking tightly coupled algorithms to work on distributed data
  - Future networking performance improvements unlikely to change these dynamics

Observations:

- Investment in persistent storage is strategic -- a new competitive advantage in global research
  - $ amount is smaller, but year-over-year growth required, vs. "constant" compute investment.
- Adopting architectures ("software defined storage") that enable interoperability, federation, unified search, authenticated access -- all critical for sustainability
- Federation plus geo-replication creates cost-sharing model for long term preservation schemes.



# eREF: Network Infrastructure



Key Learnings:

- Improving price-performance favours long-haul, high bandwidth routes
  - But not improving fast enough to alter dynamics of data-compute co-location
- Last mile -- both on campus and in the field -- remains challenging everywhere
  - Edge computing potentially offers a trade-off, but no easy solutions today
  - 5G offers nearer term solutions for many needs (e.g. climate science)
- Submarine connections remain particular challenge for islands
- Network infrastructure often tasked with security (at the expense of performance) when "higher level" controls would be more effective.

Observations:

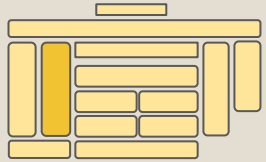
- Network investments must keep pace with
  - Institutional needs
  - Demands of "Internet of Things," dense observation systems
  - International connectivity needed to support international collaboration

# eREF: Service & Resource Governance



Key Learnings:

- Many countries "piggy back" allocations on peer-reviewed funding processes
  - Assigning awards to specific resources handled as a technical process
- Sustained support services also treated as allocated resource in many countries
- Pressure on competitive processes eased by access to "upper tier" resources (e.g. PRACE)
- Persistent storage still not a key allocated resource in most countries
- Most countries make resources available for free, except for industrial use.
  - Some countries (e.g. Denmark, Norway) charge nominal amounts, which are typically covered by related funding awards (which are increased to cover these fees)



Observations:

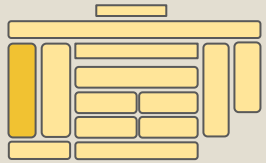
- Align resource allocation policies with top level strategy, governance and funding
  - Ideally create "one window" access process
  - Strategically, some access should be "free" (i.e. small resource quanta for grad students, new faculty) -- so funding, procurement, etc. should reflect this
- Piggy back, or at least harmonize, competitive allocation processes for research funding as well as access to other resources
  - Avoid double jeopardy
  - Lower administrative burdens
- Persistent storage of active data and long-term preservation of non-published data seem well suited to peer-reviewed competition
- Although fees-for-service provide economic incentives for efficient usage, this usually penalizes the start-up user and the largest users (whose needs are driven by the science)

# eREF: ID & Authentication, Cybersecurity



Key Learnings:

- Identity federations extending internationally
- Federated mechanisms for authenticating identities and authorizing resource access are maturing
- Two factor authorization (2FA) is becoming norm for implementing robust ID management, enabling single sign on for users
- Best practices in cybersecurity procedures are well known, but often challenging for institutions to implement, primarily because of the initial labour requirements.



Observations:

- Several roles are needed:
  - Identity provider of last resort, plus assistance setting up appropriate service providers (e.g. application frameworks, reference architectures, etc.)
  - Cybersecurity centre of excellence, to help research organizations implement effective policies and procedures
  - Ongoing cybersecuity audit capability, to provide objective assessments and corrective guidance

# eREF: Data Centre and Cloud Hosting

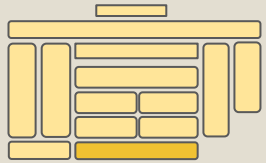


Key Learnings:

- Latest HPC systems are increasingly high-power, high-density, liquid cooled.
  - Liquid cooling will increasingly be possible with "ambient temperature" water systems, reducing or eliminating the need for chilled water systems.
- Many new Compute Infrastructure architectures, and future appliances, will not require special environments.

Observations:

- Existing institutional data centres will continue to be viable sites for new eResearch investments (assuming availability of some air cooling and the higher levels of power needed).



# Common Organizational Structures



Traditional Players

New Initiatives

“Diversity-based Evolution” vs. “Single Provider”

# Traditional Players



## HPC-centred Facilities

- Mindset evolves from keeping the machines running
- Users assumed to know how to run their codes, but increasingly facilities need to support and/or supply appropriate analysis or modeling software
- Visualization often integrated to enable pre- and post-processing of data, while minimizing data movement

## Cluster-centred Facilities

- Grid and cluster-based facilities provide capacity solutions for a broad range of science needs. Early investments into interoperability, workload portability

## Network-centred Facilities

- In most countries, NRENs have targeted a broad range of activities as natural expansion of their mandates.

# Traditional Players Expanding Roles to fill service gaps



- Various HPC-Centred facilities have expanded services, particularly in consulting, platforms, persistent storage.
  - Service expansion has sometimes been prioritized to improve a given centre's chance of surviving consolidation of HPC centres, as much as to serve the community better
- Network-centred facilities want to leverage leadership in access management, as well as joint procurement capabilities, to justify a role as the natural provider of higher value services. Cloud computing is a notable example.
- Cluster-centred facilities are expanding to provide "open science cloud" services (see below)

# New organizations & facilities arising to fill gaps



- *eScience Initiatives:* Facilitating integration of computational and data skills, software development and engineering into strategic research projects.
- *Data Science Initiatives:* Facilitating collaborations with data scientists, to enable new insights around research data.
- *National and international RAM* responding to growing need for research data management functions and activities.
- *Institutional storage initiatives*, addressing institutions' pre-existing responsibilities to preserve research results, sometimes mandated by government (e.g. in Sweden and UK)
- "*Open Science Cloud*" initiatives provide:
  - "Non-peak" compute infrastructure (potentially available through the cloud),
  - Broadly available research data management capabilities
  - Platforms, gateways and/or virtual labs making eResearch capabilities easier to access/use
- *National and international training, community support initiatives*, such as the Carpentries and Research Software Engineering associations (RSE)
- *Data Analytics Facilities:* Compute and Storage Infrastructure purpose built to optimze data analytics work.

# Diversity-based Evolution of eResearch Ecosystems



Diversity-based ecosystems are the most common:

- *Australia**, US, China, Japan, *UK**, Poland, France, Brazil, *Singapore** (Europe in general)

Not all "types" of players are active in every ecosystem. Even for similar "types" of player, mandate, funding and precise roles can vary across ecosystems

Early impressions:

- Researchers need to assemble resources from multiple providers, and interoperability can be a challenge.
- Not all researchers have access to the services they need. This is a natural result of a diversity-based approach, but may not be consistent with the publicly-funded nature of eResearch services.
  - *This impression may be the result of observational bias: There is no obvious provider or providers of certain services, so we assume the services are not being provided – but this may not be true.*
- There is evidence of unbalanced investment

**Considered as part of current benchmarking study*

# Single/Federated Provider Approach



Public funders want to avoid competition and overlap:

"Monopoly" eResearch providers avoid this problem

- Effective "monopoly" providers have appeared in several countries:
  - With government mandate: *Netherlands**, Finland
  - Without such a mandate: *Switzerland**, Spain
- Key risks: seen as "government IT" and not as competent partner for leading research groups

Federated providers harness diversity of existing ecosystem and

- avoid competition and overlap
- leverage complementary investments/skill sets
- achieve useful scale
- provide more complete range of services

- Germany, *Sweden**, Norway, Denmark, *Nordic region**
- Key risks: governance, coordination, accountability and funding equity are challenges
- *Canada** introduced federation to integrate diverse players, but now looking to differentiate roles between one national and many other players