We cannot wait for the outcome….

- design principles
  - Standardisation
  - Openness
  - Transparency
  - Reversibility and **traceability**
  - Well-documented decision making
  - **Governance, Risk and Compliance**
- secondary indicators
  - **Trust** (what is trust?)
  - Proof of **trustworthiness**
  - **Certification**
  - Predictive measures of likely success
  - “Best practices”
  - Internal **process metrics**

Digital Preservation

Governance, Risk and Compliance

Christoph Becker

http://www.ifs.tuwien.ac.at/~becker

Agenda

- Trust
- Compliance
- Risk Management
  - DP and Risks
  - DRAMBORA
- Governance, Risk, Compliance and IT
  - What is IT Governance?
  - The relevance of IT Governance for DP
- COBIT: Control Objectives for IT
  - Goals and Framework
  - Example processes

Trust

- Producers and consumers need trust in a repository
- What is trust?
- Concepts
  - being able to predict something
  - Confidence of producers and consumers
  - Reliability, authenticity
  - A trusted party is presumed to seek to fulfill expectations (legal obligations, policies, ethics, contracts…)
- How do we achieve trust?
- Options
  - Standards compliance...?
  - Certifications...?

Trust in a Repository

- Critical services require trust
- RLG/OCLC “Trusted Digital Repositories – Attributes and Responsibilities” (2002)
  - depositors trust information holders
  - users trust digital assets provided by repositories
  - information holders trust third party service providers
- How is trust established, maintained, and secured?
- How to verify trust?

Trustworthy Repositories Principles

CRL-RLG-OCLC-Nestor-DPE-DCC criteria and checklists

I. The repository commits to continuing maintenance of digital objects for identified community/communities.

II. Demonstrates organizational fitness (including financial, staffing structure, and processes) to fulfil its commitment.

III. Acquires and maintains requisite contractual and legal rights and fulfils responsibilities.

IV. Has an effective and efficient policy framework.

V. Acquires and ingests digital objects based upon stated criteria that correspond to its commitments and capabilities.

Trustworthy Repositories Principles

VI. Maintains/ensures the integrity, authenticity and usability of digital objects it holds over time.

VII. Creates and maintains requisite metadata about actions taken on digital objects during preservation as well as about the relevant production, access support, and usage process contexts before preservation.

VIII. Fulfils requisite dissemination requirements.

IX. Has a strategic program for preservation planning and action.

X. Has technical infrastructure adequate to continuing maintenance and security of its digital objects.

Audit and Certification Initiatives

- Is declaring conformance to principles sufficient?
- On what basis do we establish trust?
- RLG- National Archives and Records Administration Digital Repository Certification Task Force
  - Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC)
  - ISO RAC 16363
- NESTOR
  - Catalogue of Criteria of Trusted Digital Repositories
- DRAMBORA: Self-assessment

TRAC

Criteria checklist

Three groups

A. Organisational Infrastructure

B. Digital Object Management

C. Technologies, Technical Infrastructure & Security

TRAC and Preservation Planning I

A 3.2 Repository has procedures and policies in place, and mechanisms for their review, update, and development as the repository grows and as technology and community practice evolve.

- Planning procedure
- Watch Services, triggers
- Update of preservation plans

A3.6 Repository has a documented history of the changes to its operations, procedures, software, and hardware that, where appropriate, is linked to relevant preservation strategies and describes potential effects on preserving digital content.

- History of preservation plans (created, reviewed and updated)
- Plato: Automated documentation of planning activities

TRAC and Preservation Planning II

A3.7 Repository commits to transparency and accountability in all actions supporting the operation and management of the repository, especially those that affect the preservation of digital content over time.

- Solid workflow in consist manner enables informed and well-documented decisions
- Explicit definition of objectives and measurement units
- Change history in plans

B1.1 Repository identifies properties it will preserve for digital objects.

- Objective Tree
- Evaluation results

TRAC and Preservation Planning III

B3.1 Repository has documented preservation strategies.

- Preservation Plan

B3.3 Repository has mechanisms to change its preservation plans as a result of its monitoring activities.

- Watch Services, triggers
- Verification against changes in the environment
- Update of preservation plans

Nestor Criteria & Preservation Planning

8. The digital repository has a strategic plan for its technical preservation measures.

- Preservation Plan
- Triggers for re-evaluation
- Watch Services

9.2 The digital repository identifies which characteristics of the digital objects are significant for information preservation.

- Objective Tree

- Cf. TRAC B1.1

Nestor and TRAC versus PP

- Certification and Audit of repositories
- NESTOR and TRAC
- Plato Preservation Planning approach
  - Documented preservation strategies
  - Identification of significant properties
  - Continuous monitoring and mechanisms to react to changes in the environment

…and in practice?

- Criteria checklists important step
  - Upcoming: audit certificates (ISO 16363)
- Criteria not always helpful
  - How to measure fulfilment
  - How to prove trust
  - How to improve

- Audit and Certification as ultimate goal
- Self-audit as important step

- Governance, Risk and Compliance

DRAMBORA

- Digital Repository Audit Method Based on Risk Assessment
- Self-Audit and Self-Assessment

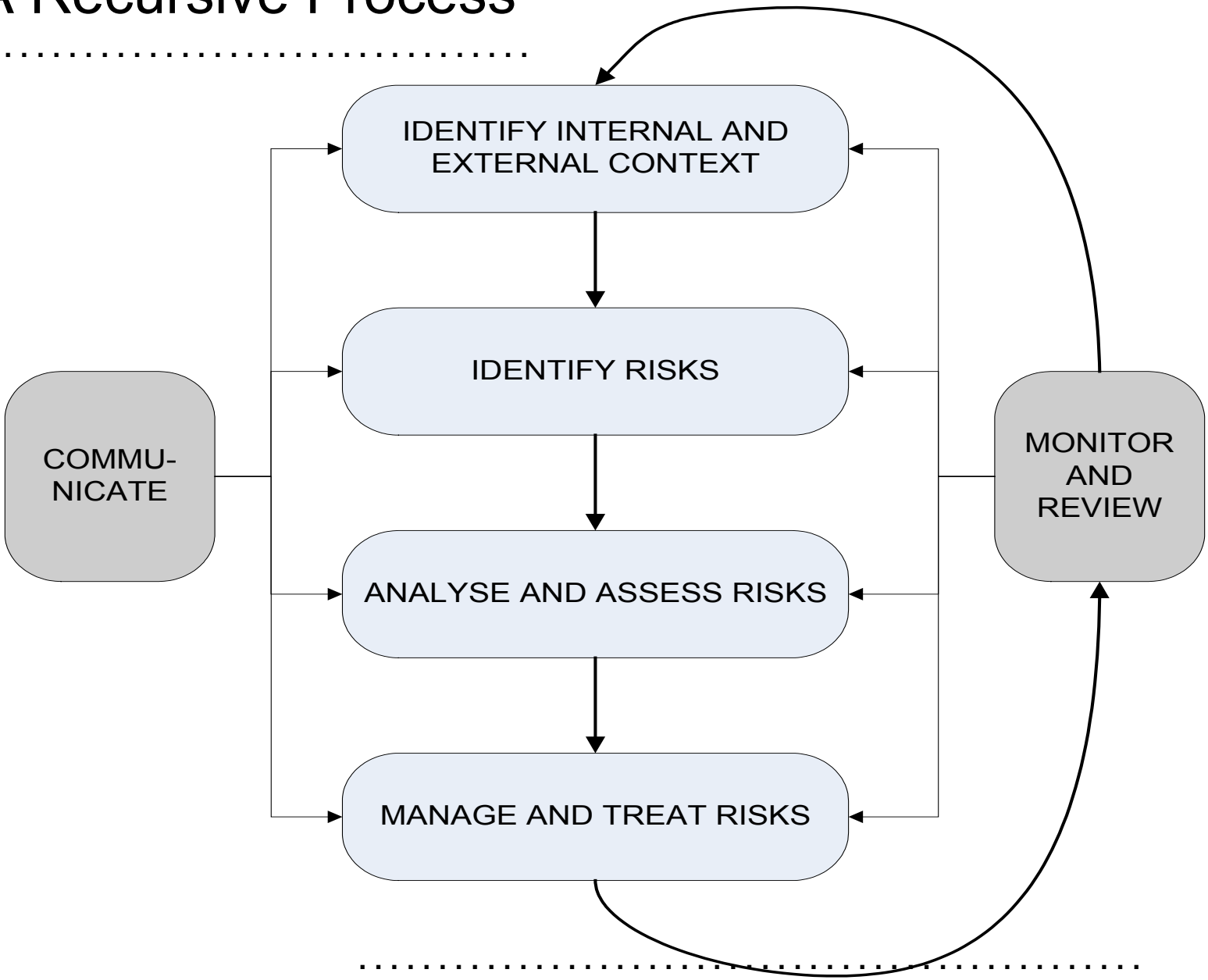
- Evidence based
  - Consistency
  - To ensure conclusions can be validated and replicated
  - Documentary, testimonial, and observational evidence

- Pilot audits
- Risk awareness is low within the community

Risk and Digital Preservation

- Digital Preservation ***is*** Risk Management
- Transform uncertainties into manageable risks
- Standard risk management models in many disciplines
- DRAMBORA is adaption of standard risk assessment procedure, customized to DP

A Recursive Process



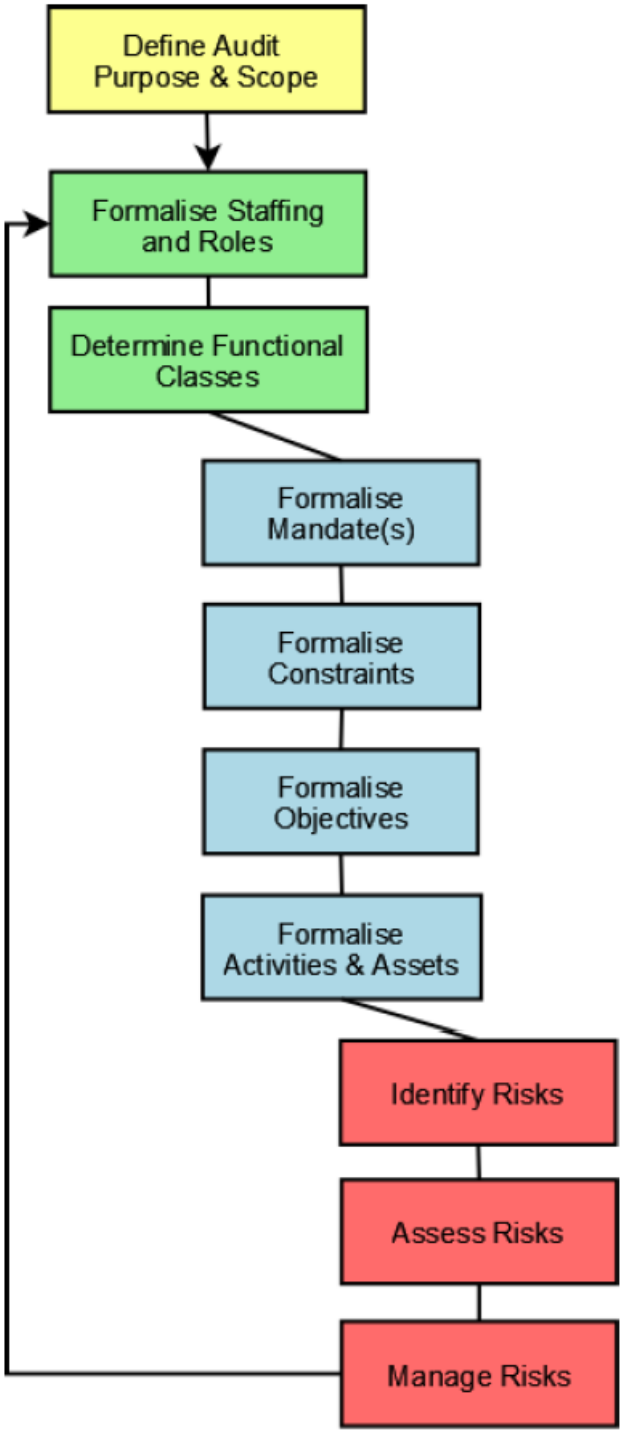
DRAMBORA

- Mechanisms to facilitate internal self-assessment & reporting
- Means to validate appropriateness of repository's efforts
- Generate appropriate documentation

- 4 stages, 10 tasks
- Available at www.repositoryaudit.eu

Process overview

1. Define Audit Purpose and Scope
2. Formalise Staffing and Roles
3. Determine Functional Classes
4. Formalise Mandate(s)
5. Formalise Constraints
6. Formalise Objectives
7. Formalise Activities and Assets
8. Identify Risks
9. Assess Risks
10. Manage Risks



Evidence

- Documentary
  - Mission statement
  - Deposit agreements
  - Business plan, financial reports
  - Job descriptions/profiles
  - System manuals, Technical documents,…
- Testimonial evidence
  - Highlight whether omissions exist in documentation
  - Validate documentation vs. reality
- Observation of practice
  - Less objective, but important
  - Walkthroughs, test objects,…

Defining the Context

- the mandate of your repository
- the goals and objectives
- the policies the repository has in place
- legal, contractual and other regulatory requirements
- standards and codes of practice
- any other things that influence the repository

➢ Well established means for subsequent risk definition and assessment

Organisational Goals

- Associated with functional classes
  - Acquisition & Ingest
  - Preservation & Storage
  - Metadata Management
  - Access & Dissemination

} operation classes

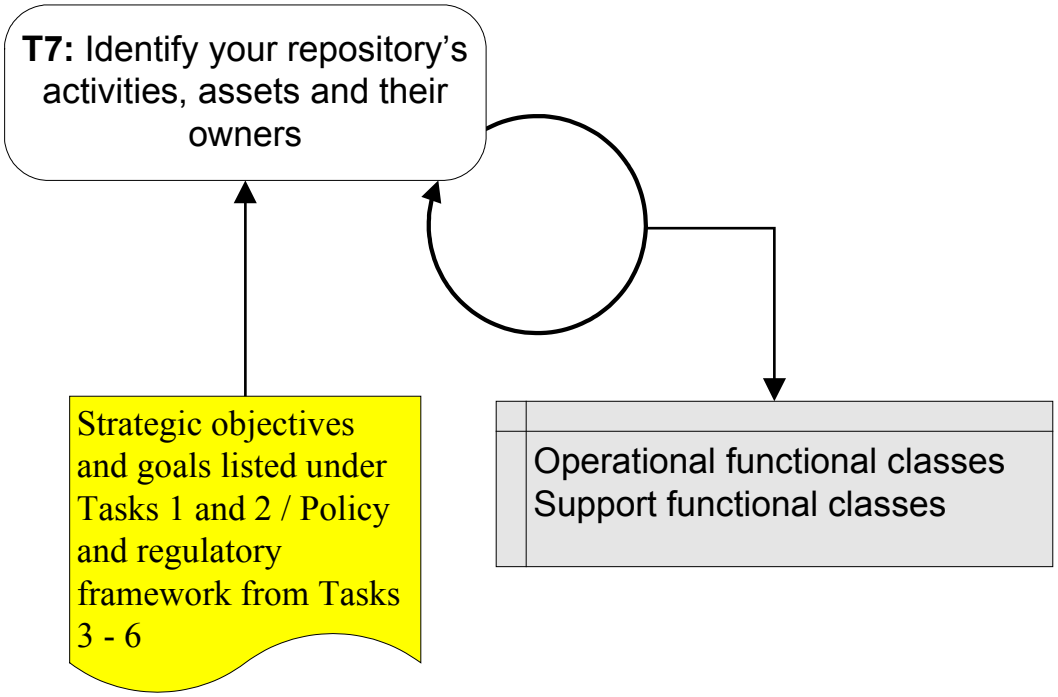
  - Organisation & Management
  - Staffing
  - Financial Management
  - Technical Infrastructure & Security

} supporting classes

Activities, Assets and their Owners

Using the digital repository self-audit tool – III

Stage 3: Identify activities, assets and their owners



Activities, Assets and Owners

- Conceptual model of what the repository does
  - split broad level mission and goals into more specific activities or work processes
  - assign to individual responsible actors
  - link to one or more key assets

Organisational Assets

- Includes:
  - information (databases, data files, contracts, agreements, documentation, policies and procedures)
  - software assets
  - physical assets
  - services and utilities
  - processes
  - people
  - intangibles, such as reputation

Identifying Risks

- Assets & Activities associated with vulnerabilities – characterised as risks
- Build structured list of risks, according to associated activities and assets
- No single methodology – brainstorming structured according to activities/assets is effective
- Iterative process

Kinds of risk

- Assets or activities fail to achieve or adequately contribute to relevant goals or objectives
- Internal threats pose obstacles to success of one or more activities
- External threats pose obstacles to success of one or more activities
- Threats to organisational assets

Stage 5: Assess Risks

- Fundamental issues are:
  - probability of risks
  - potential impact of risks
  - Relationships between / groupings of risks
- A risk assessment must be undertaken for each identified risk

Risk Assessment

- For each risk auditors must record:
  - example manifestations of risk
  - probability of its execution
  - potential impact of its execution
  - relationships with other risks
  - risk escalation owner
- Determine likelihood and impact by considering
  - Historical experiences
  - Mitigation/avoidance measures already in place
  - Experiences beyond repository itself
    - Relevant research
    - Expert opinion (e.g. legal, technical, environmental)
    - Experiences of comparable organisations

Risk Impact

- Impact can be considered in terms of:
  - impact on repository staff or public well-being
  - impact of damage to or loss of assets
  - impact of statutory or regulatory breach
  - damage to reputation
  - damage to financial viability
  - deterioration of product or service quality
  - environmental damage
  - *loss of digital object authenticity and understandability is ultimate expression of impact*

Risk Impact Score	Interpretation
0	*Zero* impact, results in **zero loss** of ability to ensure digital object authenticity and understandability[1]
1	*Negligible* impact, results in **isolated but fully recoverable loss** of digital object authenticity and understandability
2	*Superficial* impact, results in **widespread but fully recoverable loss** of digital object authenticity and understandability
3	*Medium* impact, results in **total but fully recoverable** loss of digital object authenticity and understandability
4	*High* impact, results in **isolated loss, including unrecoverable loss** of digital object authenticity and understandability
5	*Considerable* impact, results in **widespread loss, including unrecoverable loss or loss that is recoverable only by third party** of digital object authenticity and understandability
6	*Cataclysmic* impact, results in **total and unrecoverable loss** of digital object authenticity and understandability

[1] Note that we use understandability in its broadest sense to encapsulate technical, contextual, syntactical and semantic understandability.

Risk Likelihood

Risk Probability Score	Interpretation
1	Minimal probability, occurs once every **100 years or more**
2	Very low probability, occurs once every **10 years**
3	Low probability, occurs once every **5 years**
4	Medium probability, occurs once **every year**
5	High probability, occurs once **every month**
6	Very high probability, occurs **more than once every month**

Risk relationships

Risk Relationship	*Definition of Risk Relationship*
Explosive	where the simultaneous execution of *n* risks has an impact in excess of the sum of each risk occurring in isolation
Contagious	where a single risk's execution will increase the likelihood of another's
Complementary	where avoidance or treatment mechanisms associated with one risk also benefit the management of another
Domino	where avoidance or treatment associated with a single risk renders the avoidance or treatment of another less effective
Atomic	where risks exist in isolation, with no relationships with other risks

Sample risk 1

Risk Identifier	R05
Risk Name	Repository loses mandate
Risk Description	Basis for repository's existence is withdrawn or substantially altered, rendering it incompatible with business activities
Is this Risk relevant?	Is the mandate subject to ongoing review? Is the primary repository service contract subject to renewal or renegotiation?
Example Manifestation	Scope of repository responsibility is changed by legislative amendment
Nature of Risk	Personnel, management and administration procedures
Probability	2
Potential Impact	4

Sample risk 1: Mitigation

- Avoidance
  - Seek all available certifications to publicly demonstrate operational effectiveness
  - Promote organisational transparency
- In the event of execution
  - Establish arrangements for succession
  - Establish contingency plans
  - Establish exit strategy

Sample risk 2

Risk Identifier	R66
Risk Name	Preservation strategies result in information loss
Risk Description	Exposure of an archived object to preservation plans result in loss or damage to one or more of its significant characteristics
Is this Risk relevant?	Does repository offer a definition of acceptable loss that may result from preservation activities?
Example Manifestation	Migration strategy results in loss of 'look and feel' of archived documents, regarded as essential properties by user community
Nature of Risk	Operations and service delivery
Probability	4
Potential Impact	3

Sample risk 2: Mitigation

- Avoidance
  - Evaluate preservation strategies in controlled environment prior to execution
  - Ensure procedures are reversible in the event of unexpected or inappropriate results
- In the event of execution
  - Define policies to describe the acceptable levels of loss tolerated by the repository

Stage 6: Manage Risks

- Combination of avoidance, tolerance and transfer
  - avoid circumstances in which risk arises
  - limit likelihood of risk
  - reduce potential impact of risk
  - share the risk
  - Transfer to others

The Result

- Risk score for each risk quantifies risks' severity
- Composite risk score for each category
- Illustrates vulnerabilities
- Facilitates resource investment

DRAMBORA Outcomes

- Documented organisational self-awareness
- Catalogued risks
- Risk scores
  - The self-audit produces a composite risk score for each functional class.
  - This numeric result can be compared with risk scores of other functional classes and allows the identification of the areas of repository work that are most vulnerable to threats
- Understanding of infrastructural successes and shortcomings
- Preparation for full scale external audit

Summary

- Criteria for trusted repositories
  - TRAC, Nestor, Trustworthy Repository Principles
  - Relation to Preservation Planning
- DRAMBORA: Risk-based self assessment
  - Documented self-awareness
  - Risk register as basis for ongoing management
  - Preparation for external audit
- Building Trust

A Digital Archive Disappears

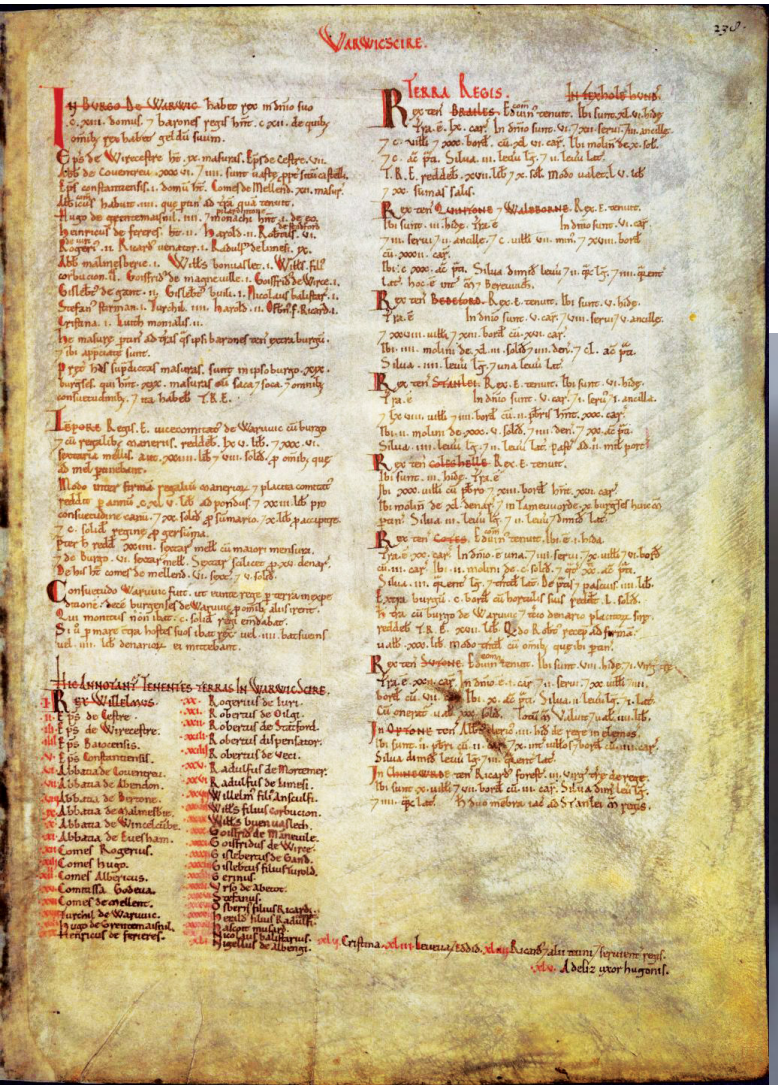
- In 1991, … [an] ambitious digital preservation project sought to capture and archive the content of numerous nascent electronic journals. The project attracted considerable attention among authors and editors of e-journals… “to create a significant collection of electronic journals on the Internet, which scholars, libraries, and individuals around the world can access via the Web.” (info.lib.uh.edu/pr/v7/n4/mace7n4.html).
  Alas, **essential funding never appeared**, and CICNet itself ceased operations in 1997. The CICNet Journal **Archive vanished** with it. Ironic, indeed, to lose not a mere collection but an archive whose purpose was to prevent loss of electronic content. How many pioneering e-journals, many of them hosted on now-defunct Gopher servers, were lost for eternity? [2]

Lost Bush emails

- Under the Presidential Records Act, White House is legally obliged to keep copies of all communication, including emails
- In 2007, 22 million emails were declared missing
- With tremendous effort emails could be recovered
- Melanie Sloan (CREW executive director): **We may never discover the full story of what happened here. It seems like they just didn't want the e-mails preserved**."
- *[http://www.guardian.co.uk/technology/2009/dec/15/bush-emails-recovered]*

The Domesday book

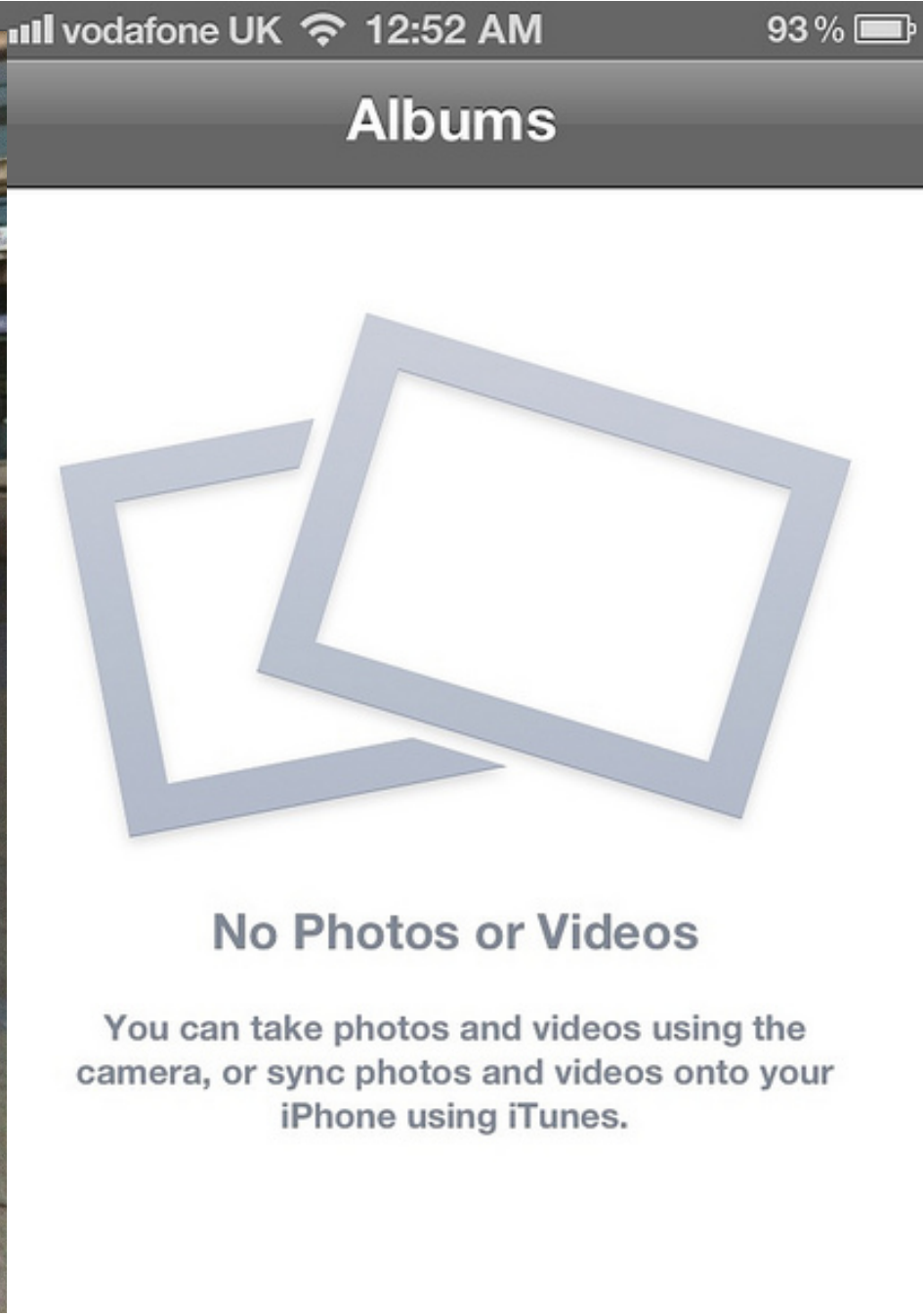
- http://www.ariadne.ac.uk/issue36/tna


WARWICSCIRE.

TERRA REGIS.

Tevatron (particle physics)

- 20 PB of data accumulated over 26 years
- *Both D0 and the CDF* ***expect to lose their dedicated computing infrastructure*** *over the next five years. A gradual loss of knowledge about how to deal with the complex data, which includes raw detector readouts, reconstructed particle trajectories and higher-level analyses, could also present a serious hurdle to exploiting the data in the future.*
- http://www.nature.com/news/2011/110527/full/474016a.html

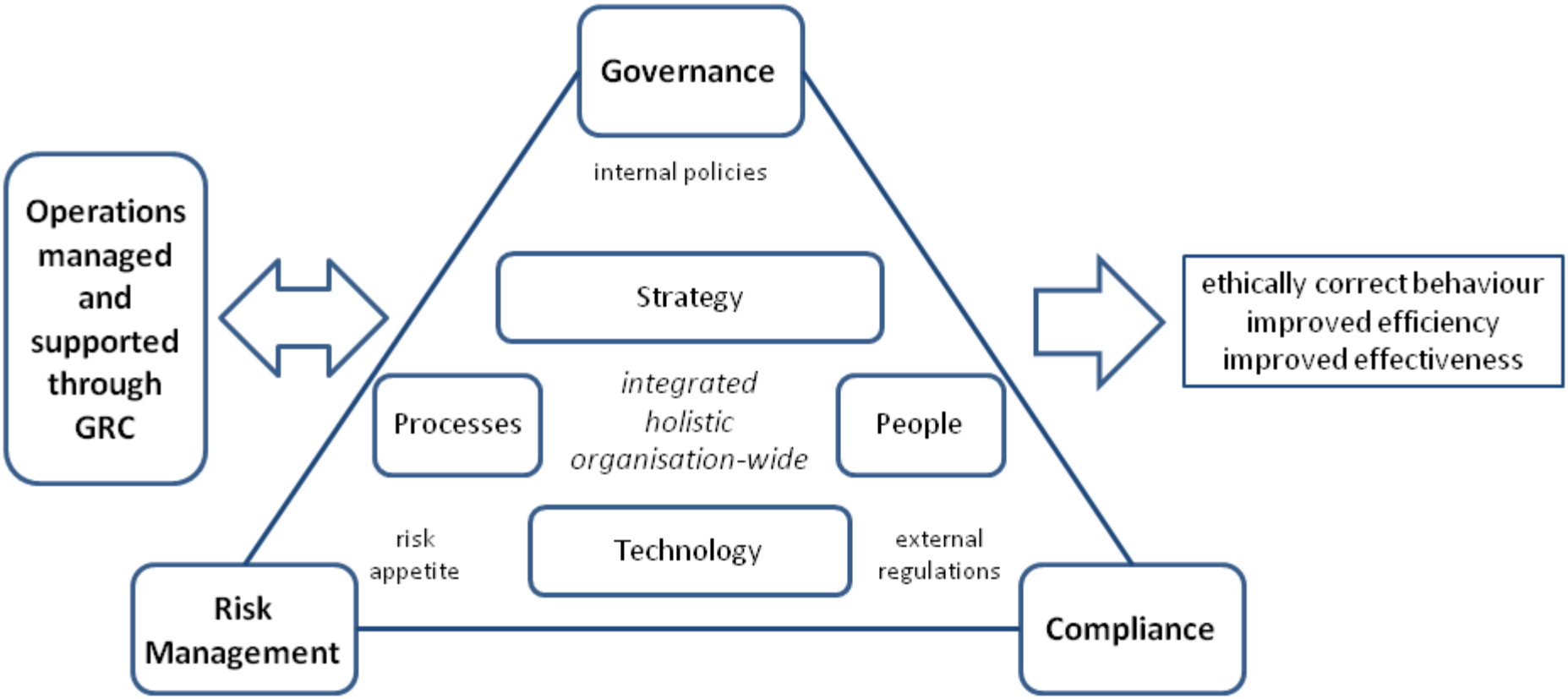


http://www.flickr.com/groups/2121762@N23/pool/

Given a library caring for a massive scanned newspaper collection...

- Activity: Preservation planning
- Assets?
- Owners?
- Risks?
- How to assess, how to mitigate?

Governance, Risk and Compliance

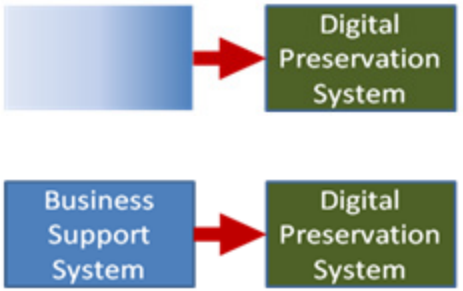


Risk, Compliance... Governance

- Digital Preservation vs. Information Management
  - Information Management needs IT
  - Information Technology needs controls
- What is IT Governance?
  - Expectations, goals, responsibility, performance, control
  - Systems, Organizations and Goals
  - Strategic alignment of business and technology
  - Enterprise Architecture
- Transparency, compliance, trust: Governance

Digital Preservation Scenarios

- DP is a concern in different scenarios
- Digital Preservation System" (DPS)
  The business is digital preservation
- "Systems of Systems" (SoS)
  The business is supported by a system



- Digital Preservation Ready System" (DPR):
  The business transposes the digital preservation concerns in non-functional requirements

- Strategic Alignment: technology <> business

IT Governance

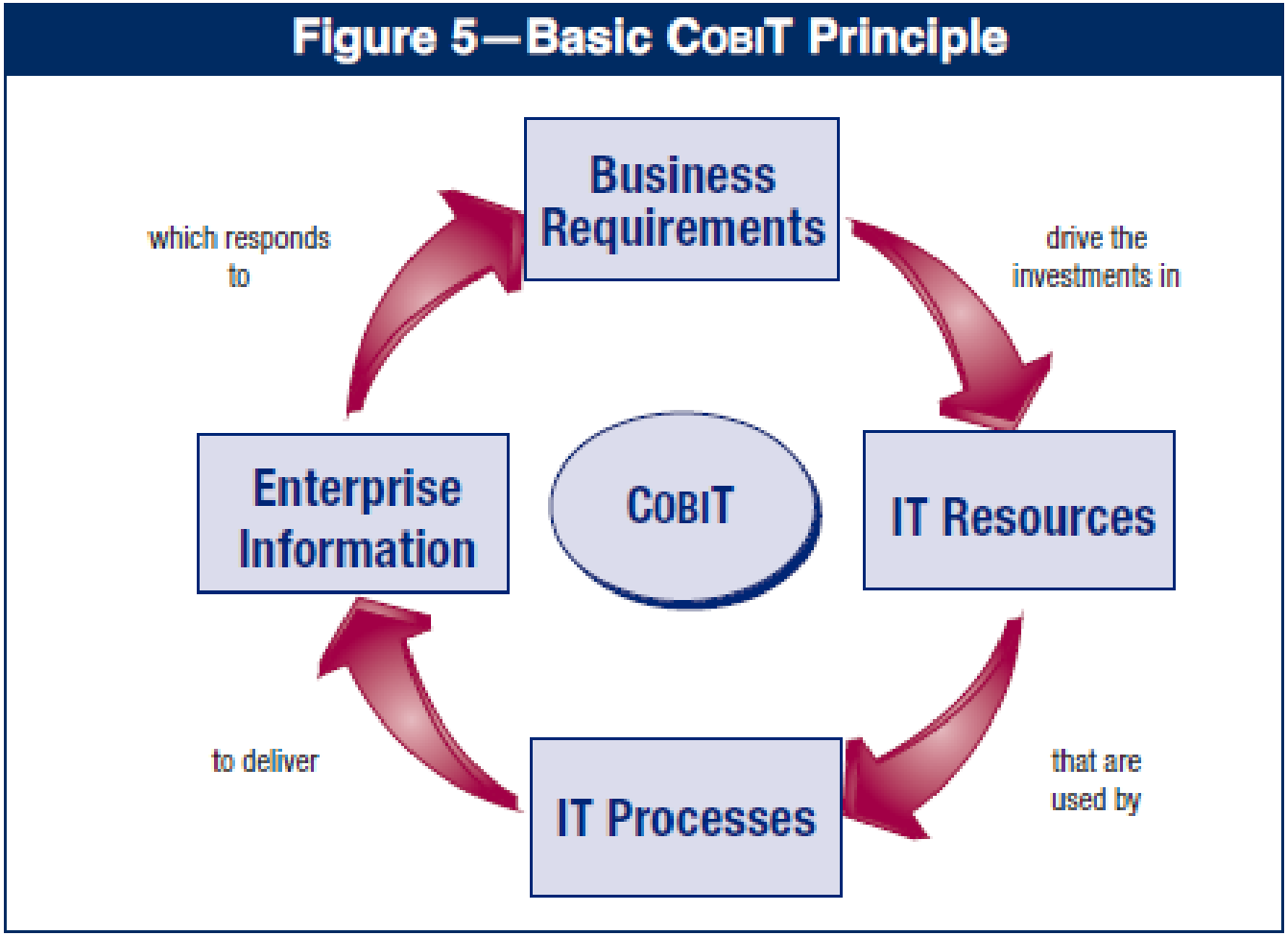
- DP is an IT-supported business and needs IT Governance
  - Key discipline for decision making and communication within IT-supported organisations
  - Proactive identification of problems
  - Early action to minimise impact
  - Gaining importance because of regulatory requirements and compliance
- IT Governance
  - "*the leadership, organisational structures and processes that ensure that the enterprise's IT sustains and extends the organisation's strategies and objectives*." [COBIT]

IT Governance: COBIT

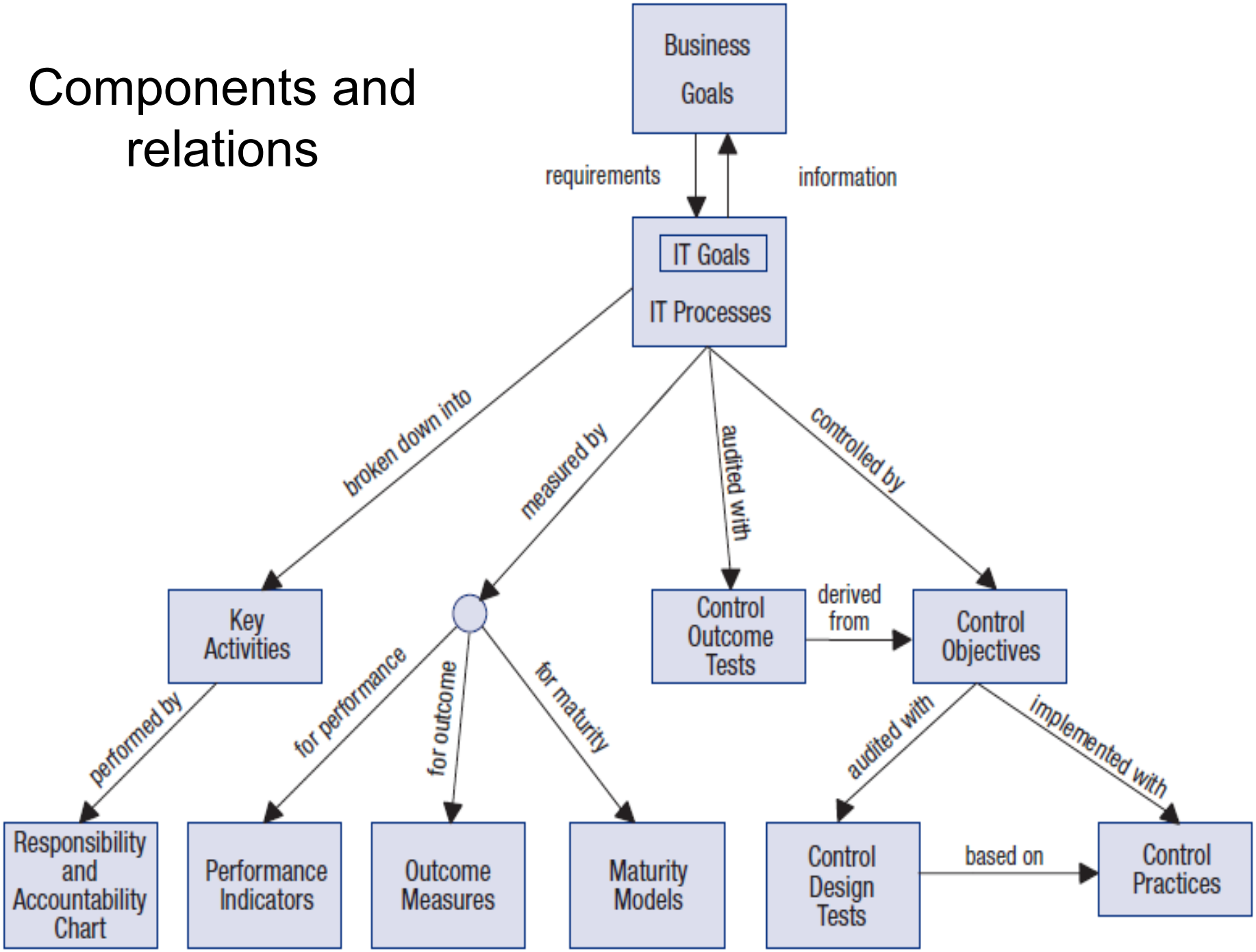
- Control Objectives for IT
  - set of best practices, measures and processes to assist the management of IT systems
- Goal: Enable businesses to deliver against business and governance requirements
  - Link and align technology to business requirements
  - Make performance against these measurements transparent
  - Organise activities into accepted process model
  - Identify major resources to be leveraged
  - Define management control objectives to be considered

Principles

- business-focused
- process-oriented
- controls-based
- measurement-driven

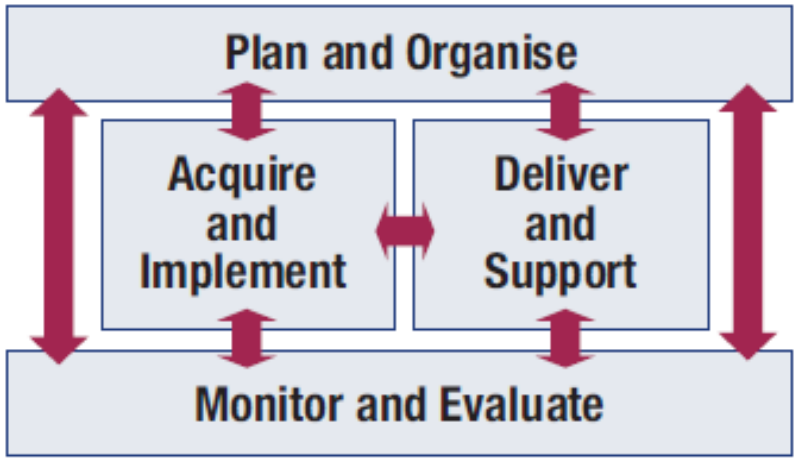


Components and relations



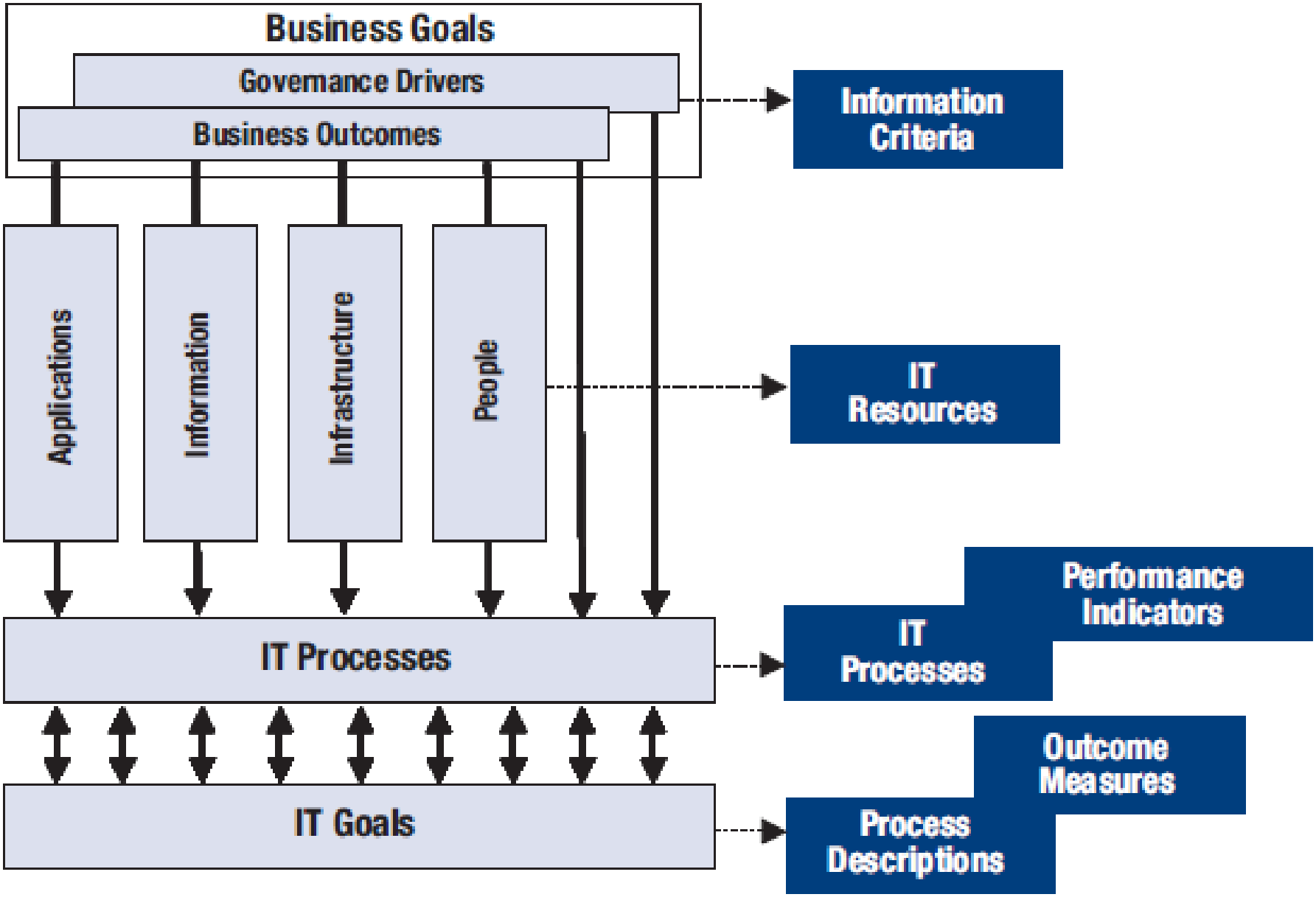
COBIT domains and information criteria

- Domains:
  - Plan and Organise
  - Acquire and Implement
  - Deliver and Support
  - Monitor and Evaluate



- Information Criteria
  - **Effectiveness:** relevant and pertinent to the business process as well as being delivered in a timely, correct, consistent and usable manner.
  - **Efficiency**: provision of information through the optimal (most productive and economical) use of resources.
  - **Confidentiality**: protection from unauthorised disclosure.
  - **Integrity:** accuracy and completeness, validity to business
  - **Availability:** available when required, safeguarded
  - **Compliance** to laws, regulations, contracts (external and internal)
  - **Reliability**: enable management to operate and exercise responsibilities

Control, Alignment, Monitoring



Some COBIT4 processes

- Plan and Organise
  - PO1 Define a Strategic IT Plan
  - PO9 Assess and Manage IT Risks
- Acquire and Implement
  - AI1 Identify Automated Solutions
  - AI3 Acquire and Maintain Technology Infrastructure
- Deliver and Support
  - DS1 Define and Manage Service Levels
  - DS4 Ensure Continuous Service
- Monitor and Evaluate
  - ME3 Ensure Compliance With External Requirements

PO9 Assess and Manage IT Risks

- A risk management framework is created and maintained. The framework documents a common and agreed-upon level of IT risks, mitigation strategies and residual risks. Any potential impact on the goals of the organisation caused by an unplanned event is identified, analysed and assessed. Risk mitigation strategies are adopted to minimise residual risk to an accepted level. The result of the assessment is understandable to the stakeholders and expressed in financial terms, to enable stakeholders to align risk to an acceptable level of tolerance.
- Control Objectives
  - PO9.1 IT Risk Management Framework
  - PO9.2 Establishment of Risk Context
  - PO9.3 Event Identification
  - PO9.4 Risk Assessment
  - PO9.5 Risk Response
  - PO9.6 Maintenance and Monitoring of a Risk Action Plan

PO9 Assess and Manage IT Risks

A RACI chart maps all activities onto roles:
Responsible, **A**ccountable, **C**onsulted, **I**nformed

RACI Chart

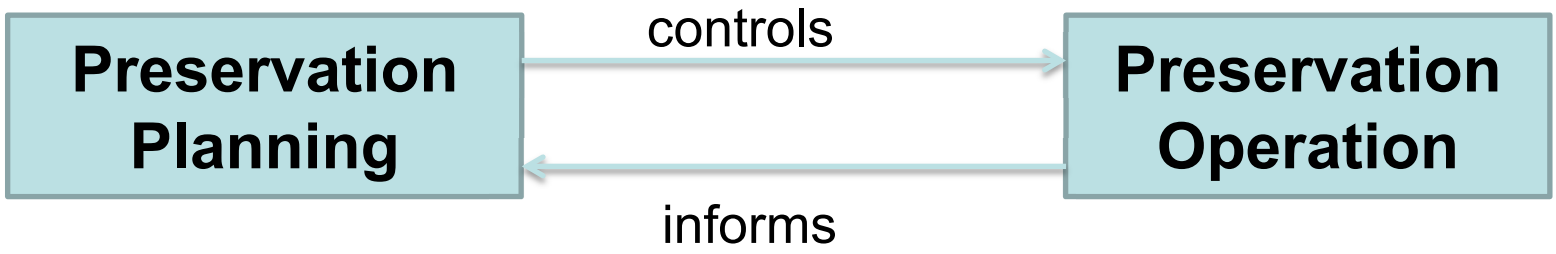
Activities \ Functions	CEO	CFO	Business Executive	CIO	Business Process Owner	Head Operations	Chief Architect	Head Development	Head IT Administration	PMO	Compliance, Audit, Risk and Security
Determine risk management alignment (e.g., assess risk).	A	R/A	C	C	R/A	I					I
Understand relevant strategic business objectives.		C	C	R/A	C	C					I
Understand relevant business process objectives.				C	C	R/A					I
Identify internal IT objectives, and establish risk context.					R/A		C	C	C		I
Identify events associated with objectives (some events are business-oriented [business is A]; some are IT-oriented [IT is A, business is C]).	I			A/C	A	R	R	R	R		C
Assess risk associated with events.				A/C	A	R	R	R	R		C
Evaluate and select risk responses.	I	I	A	A/C	A	R	R	R	R		C
Prioritise and plan control activities.	C	C	A	A	R	R	C	C	C		C
Approve and ensure funding for risk action plans.		A	A		R	I	I	I	I		I
Maintain and monitor a risk action plan.	A	C	I	R	R	C	C	C	C	C	R

A **RACI** chart identifies who is **R**esponsible, **A**ccountable, **C**onsulted and/or **I**nformed.

Core Preservation Capabilities

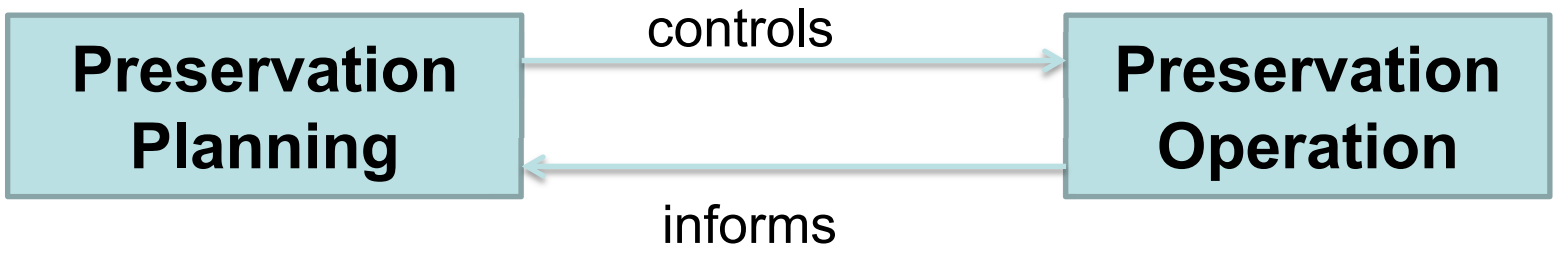
Preservation Planning	Preservation Operation
Monitor, steer and control the preservation operation of content	Control the deployment and execution of preservation plans.

Core Preservation Capabilities



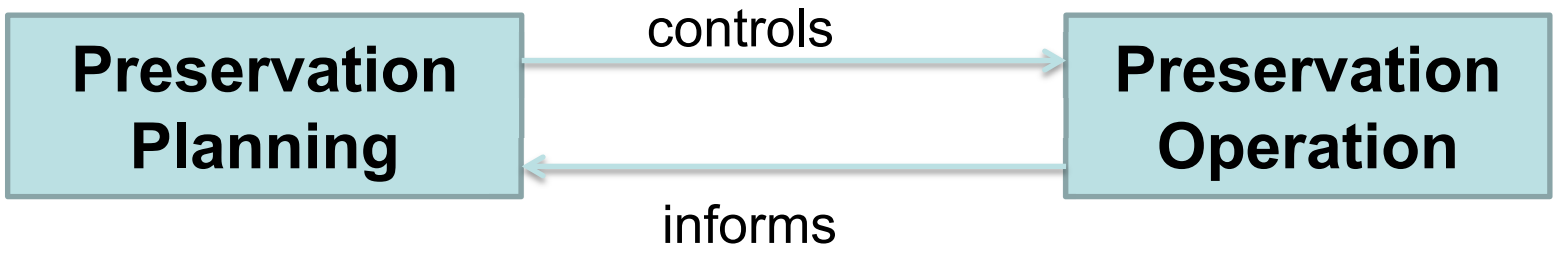
Preservation Planning	Preservation Operation
	Run operations and report on them
	•Analyze content •Execute preservation actions •Ensure adequate provenance trail •Handle preservation metadata •Conduct Quality Assurance •Provide reports and statistics

Core Preservation Capabilities



Preservation Planning	Preservation Operation
Monitor and control operations	Run operations and report on them
•Influencers and Decision making •Options diagnosis •Specification and delivery •Monitoring	•Analyze content •Execute preservation actions •Ensure adequate provenance trail •Handle preservation metadata •Conduct Quality Assurance •Provide reports and statistics

Core Preservation Capabilities



Preservation Planning	Preservation Operation
Monitor and control operations	Run operations and report on them
•Influencers and Decision making •Options diagnosis •Specification and delivery •Monitoring	•Analyze content •Execute preservation actions •Ensure adequate provenance trail •Handle preservation metadata •Conduct Quality Assurance •Provide reports and statistics
"Migrate this set of images (in TIFF-5) to JP2 using ImageMagick 6.3 with parameters a,b,c"	*•Analyse original* *•Migrate, analyse output* *•Conduct quality assurance* *•Provenance, metadata, Reporting*

Maturity models

- Each process can have different maturity levels ...in different dimensions:
  - Awareness and Communication
  - Policies, plans and procedures
  - Tools and Automation
  - Skills and Expertise
  - Responsibility and Accountability
  - Goal Setting and Measurement
- Levels
  0. Non-existent
  1. Initial/Ad Hoc
  2. Repeatable but Intuitive
  3. Define Process
  4. Managed and Measurable
  5. Optimised

A Maturity Model for Operations

Coming from Software Engineering, the CMM has been shown to be a powerful instrument for assessment and improvement

	Awareness and Communication	Policies, Plans and Procedures	Tools and Automation	Skills and Expertise	Responsibility and Accountability	Goal Setting and Measurement
1	Initial / ad-hoc					
2	Repeatable, but Intuitive					
3	Defined					
4	Managed and Measurable					
5	Optimized					

...Governance, Risk, Compliance and DP

- Governance
  - DP is an IT-supported business and needs proper IT Governance
  - IT Governance frameworks do not explicitly address DP concerns
  - Planning exercises Control based on Objectives
- Risk
  - DRAMBORA customizes Risk Management for DP
  - Currently: lack of integration with other RM frameworks
- Compliance
  - OAIS and TRAC / RAC
  - But: no maturity model established yet

Questions?

becker@ifs.tuwien.ac.at
www.ifs.tuwien.ac.at/~becker