# A Capability Model for Digital Preservation

## Analyzing Concerns, Drivers, Constraints, Capabilities and Maturities


Christoph Becker
Vienna University of Technology
Vienna, Austria
www.ifs.tuwien.ac.at/~becker

Gonçalo Antunes,
José Barateiro,
Ricardo Vieira
INESC-ID Information Systems Group, Lisbon, Portugal


SCAPE
SCAlable Preservation Environments

SHAMAN
Sustaining Heritage Access through Multivalent ArchiviNg

FACULTY OF !NFORMATICS

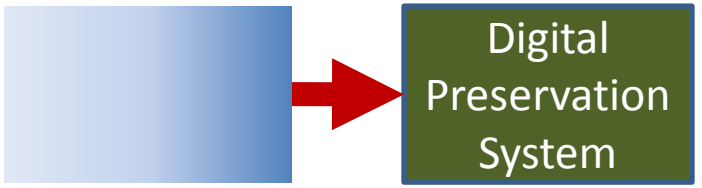
# Digital Longevity

- Numerous reference models, frameworks and concepts
  - OAIS and trust: TRAC, RAC (ISO 16363), NESTOR…
  - Records Management: MoReq, ISO 15489…
  - Risk: DRAMBORA…
  - Planning: PLATO, PLATTER
  - Economics: BRTF, LIFE….
- Yet, we still lack a holistic view
  - Maturity of the field is unclear and evolving
  - Integration into Information Systems and Information Technology fields is unclear
  - How does Digital Preservation relate to, e.g., IT Governance?
  - How can we assess and improve organizational capabilities?
  - How can we deal with non-repository scenarios?

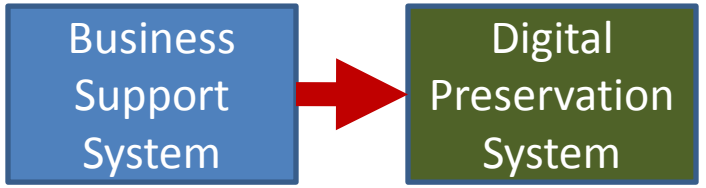
# About systems, requirements and preservation

Scenarios of systems
and their perceived relevance of digital preservation requirements

The Digital Preservation System (DPS):
DP as *functional requirements*




The Systems of Systems (SoS): Business system delegates DP responsibility to a DPS




The "Digital Preservation Ready" System (DPR):
Longevity as a *non-functional requirement*!

# Background

- Enterprise Architecture (EA)
  - Holistic system architecture approach to information systems and technology in organizations
- IT Governance
  - decision making and communication within IT-supported organizations
  - leadership, organisational structures and processes
  - ensure that the IT sustains the organisation's objectives
  - COBIT: Control Objectives for IT
- The goals of Reference Architectures
  - Process
  - Stakeholder concerns
  - Independent of business domain and organization

# Make DP ubiquitous in systems

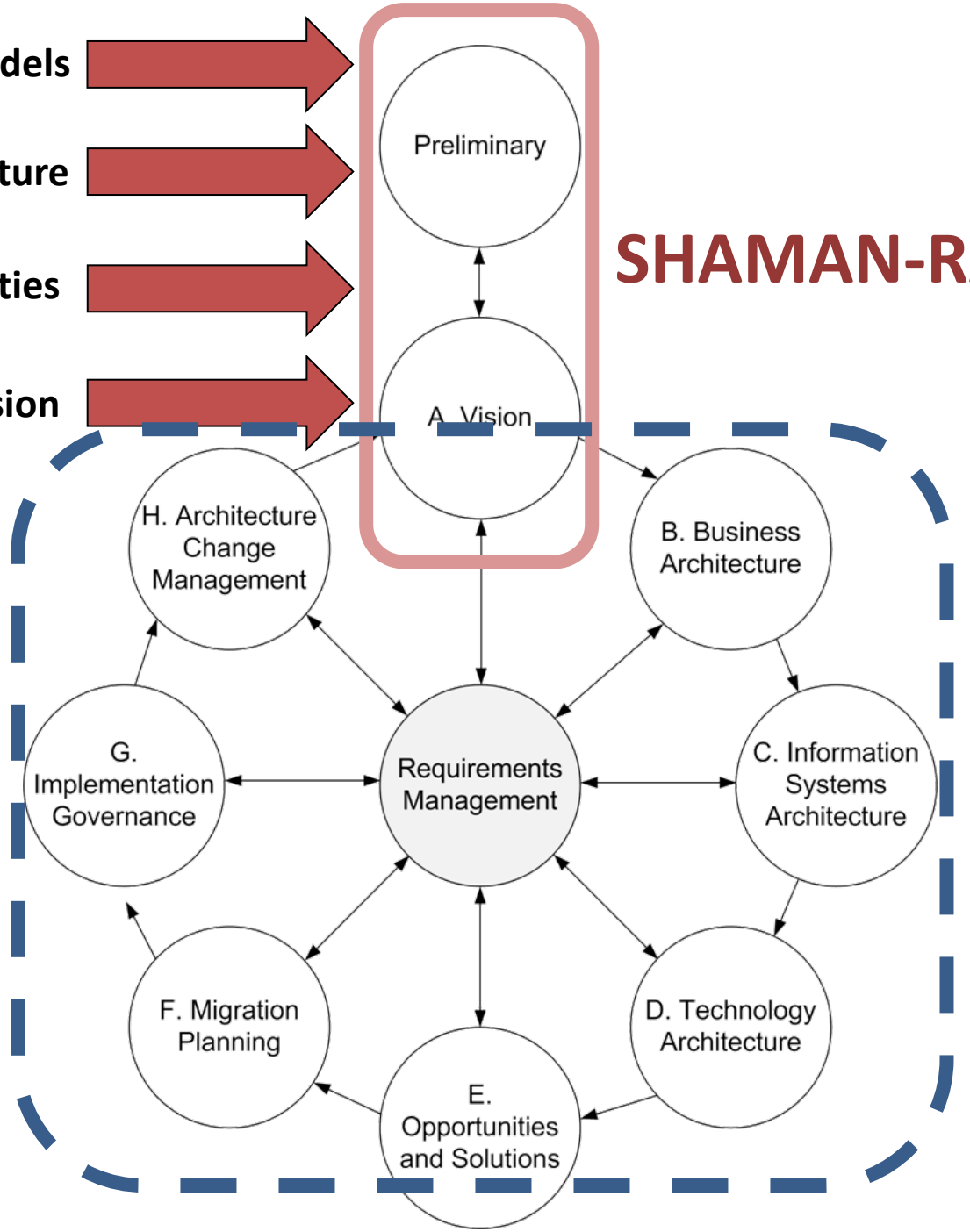
- Consider clearly defined Goals…
  - Business process
  - Stakeholder concerns
- Align business and technology according to the best references in Enterprise Architecture (EA)…
  - TOGAF Architecture Development Method
- Follow IT Governance best practices…
  - COBIT (goal-oriented, process-oriented, control-based)
  - Maturity Model based on CMM
- **Define a Reference Architecture for reuse in the EA processes where DP is a relevant concern**

Consolidation of DP Reference Models

Contextualization of a DP Architecture

Modeling of DP Capabilities

Creation of a DP Architecture Vision



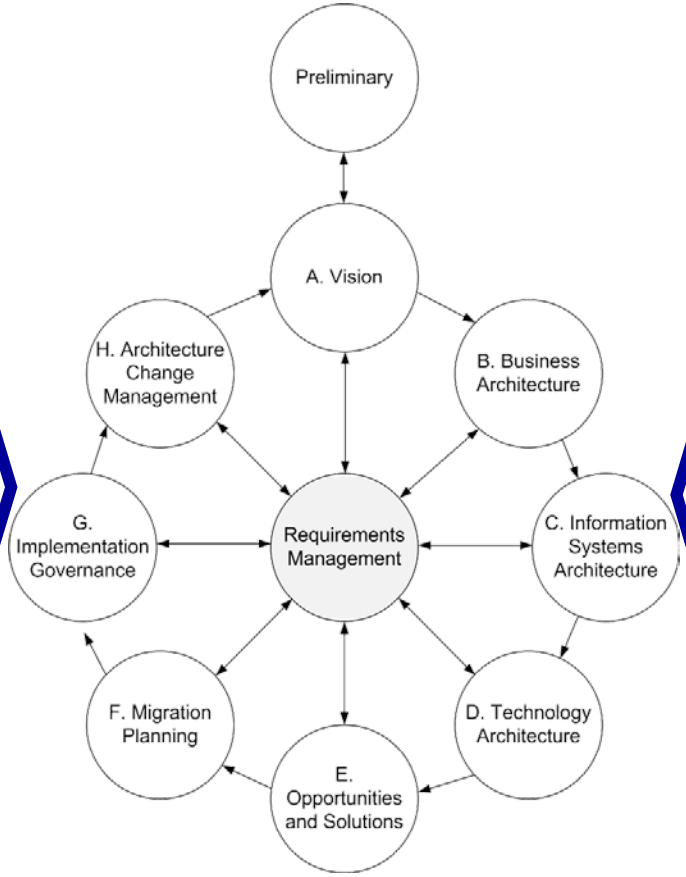

SHAMAN-RA

CONCRETE ARCHITECTURES

# A Capability-based Reference Architecture

**Domain Knowledge**

- SHAMAN-RA v1.0
- OAIS
- TRAC/RAC
- TDR 2002
- NESTOR
- Planets Planning method
- Planets Functional Model
- PREMIS
- BRTF Sustainability Report
- DRAMBORA
- PARSE.Insight
- …

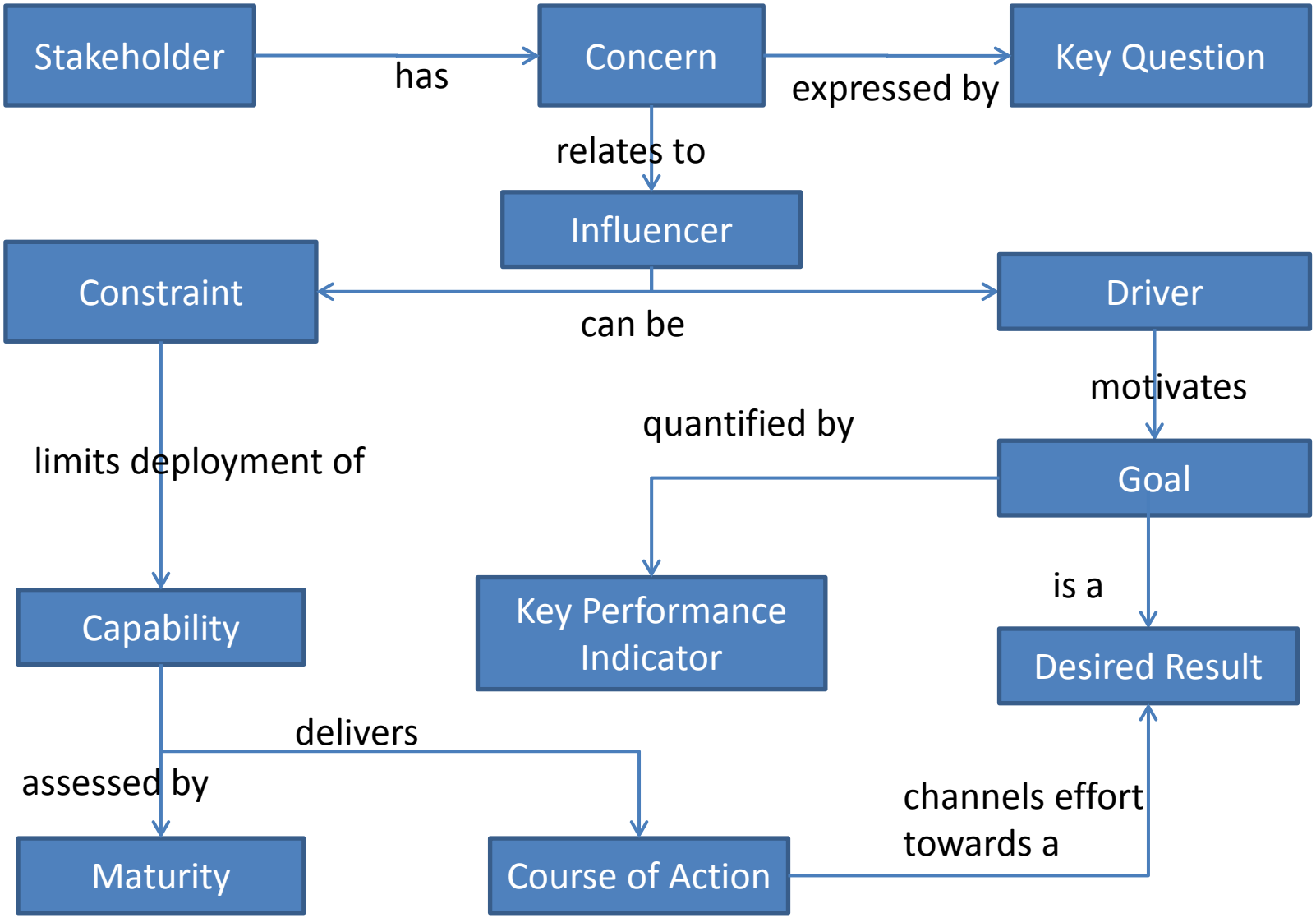



**Standards and Best-Practices**

- OMG UML
- OMG BMM
- OMG SBVR
- OMG OSM
- ISO 27000: Security
- ISO 31000: Risk Manag.
- IEEE Std. 1471-2000
- Zachman Framework
- COBIT
- DoDAF
- …

| Stakeholders | Concerns | Influencers | Goals | Capabilities |
|---|---|---|---|---|

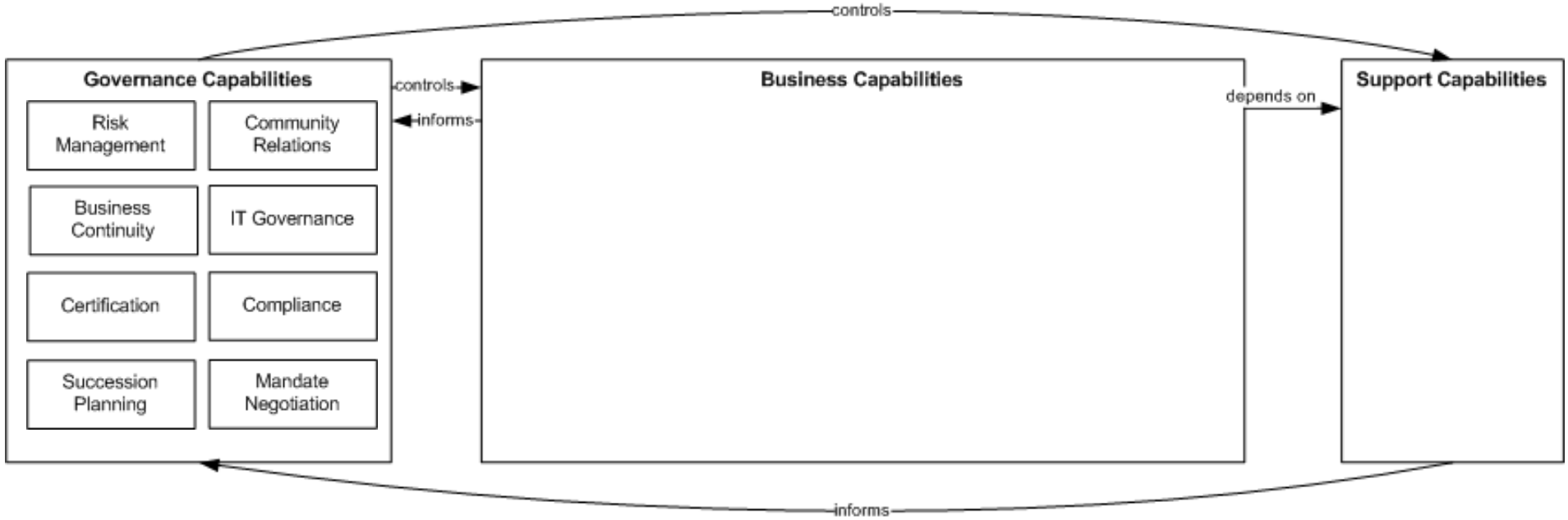
# Metamodel of key elements



# Digital Preservation Capabilities

A **capability** is an "*ability that an organization, person, or system possesses. Capabilities are typically expressed in general and high-level terms and typically require a combination of organization, people, processes, and technology to achieve*"

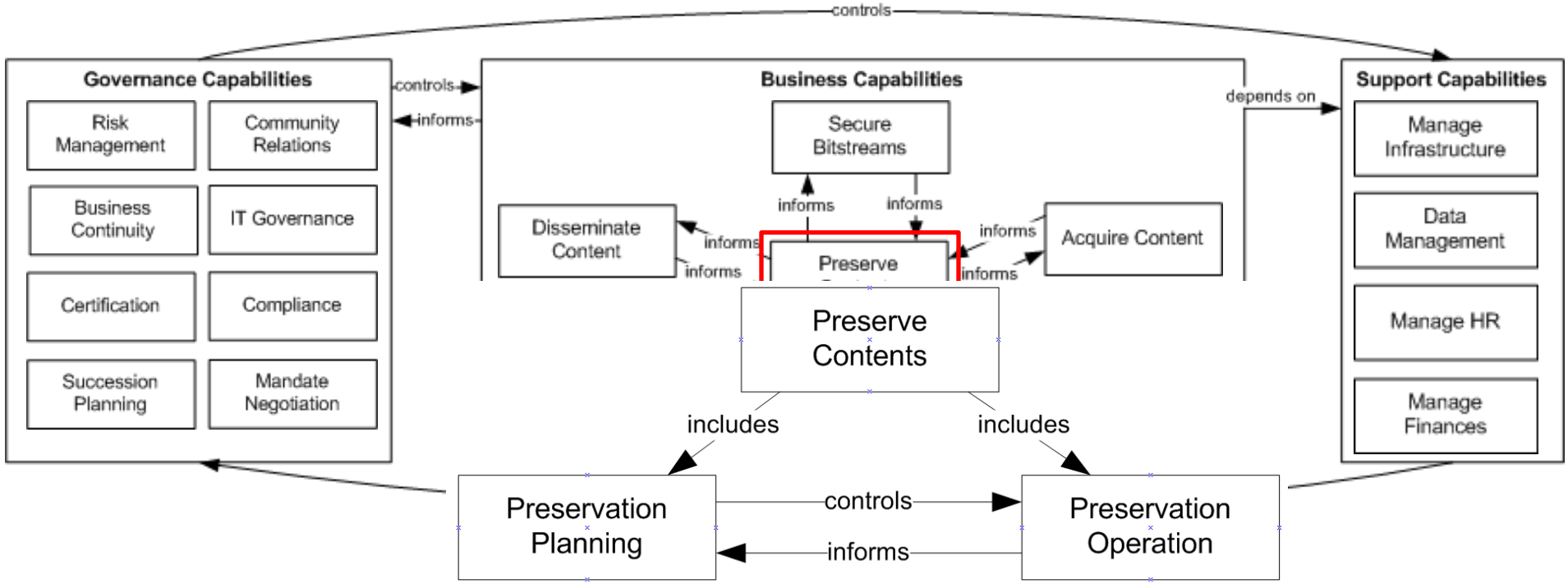
*A capability can control, inform, include, or depend on another capability*



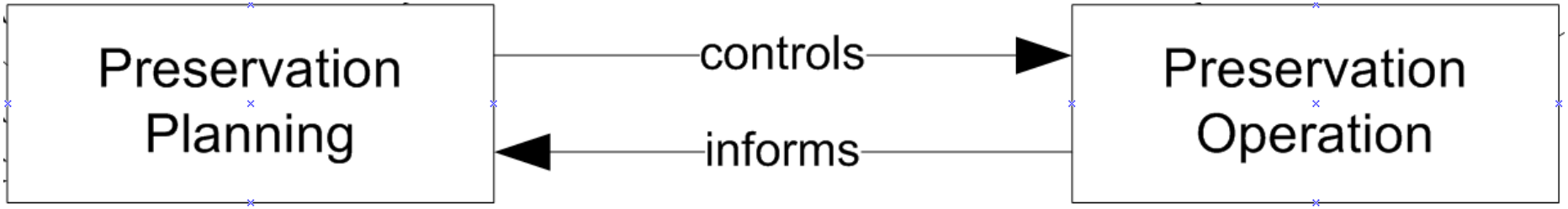
# Digital Preservation Capabilities

A **capability** is an "*ability that an organization, person, or system possesses. Capabilities are typically expressed in general and high-level terms and typically require a combination of organization, people, processes, and technology to achieve*"

**Preserve Contents** is the ability to **maintain content authentic and understandable to the defined user community over time** and assure its provenance.

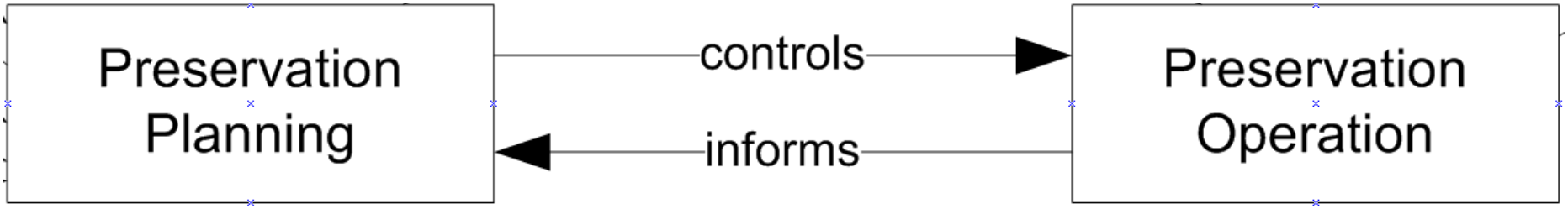


# Core Preservation Capabilities



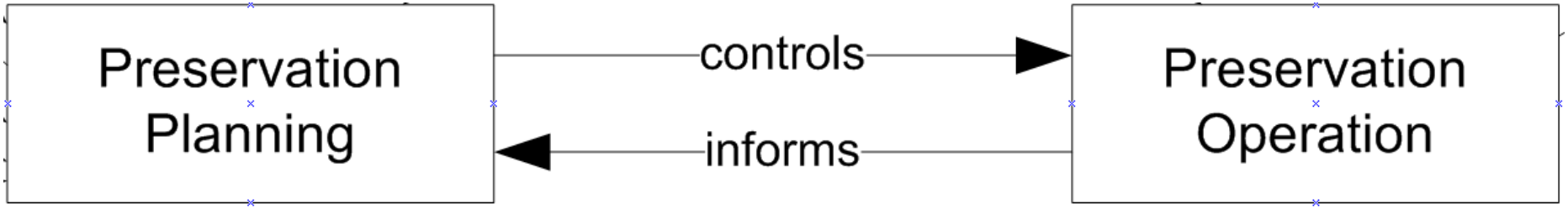

| Preservation Planning | Preservation Operation |
| --- | --- |
| Monitor, steer and control the preservation operation of content | Control the deployment and execution of preservation plans. |
| | |
| | |

# Core Preservation Capabilities



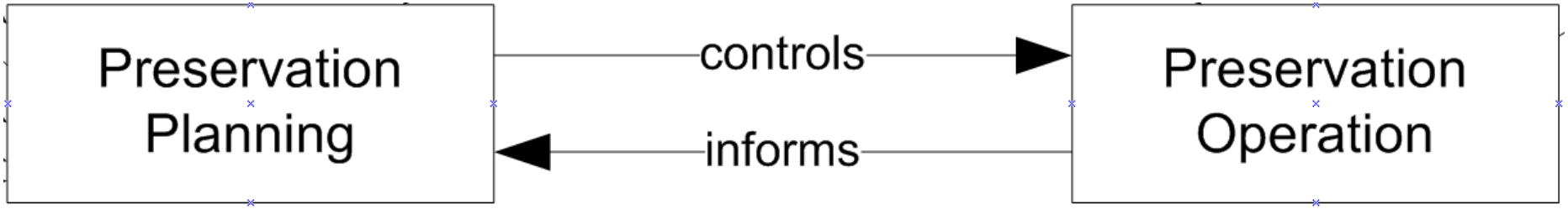

| Preservation Planning | Preservation Operation |
|---|---|
| Monitor, steer and control the preservation operation of content | Control the deployment and execution of preservation plans. |
| •Influencers and Decision making<br>•Options diagnosis<br>•Specification and delivery<br>•Monitoring | |
| | |

# Core Preservation Capabilities




| Preservation Planning | Preservation Operation |
|---|---|
| Monitor, steer and control the preservation operation of content | Control the deployment and execution of preservation plans. |
| •Influencers and Decision making<br>•Options diagnosis<br>•Specification and delivery<br>•Monitoring | •Analyze content<br>•Execute preservation actions<br>•Ensure adequate provenance trail<br>•Handle preservation metadata<br>•Conduct Quality Assurance<br>•Provide reports and statistics |
| | |

# Core Preservation Capabilities




| Preservation Planning | Preservation Operation |
|---|---|
| Monitor, steer and control the preservation operation of content | Control the deployment and execution of preservation plans. |
| •Influencers and Decision making<br>•Options diagnosis<br>•Specification and delivery<br>•Monitoring | •Analyze content<br>•Execute preservation actions<br>•Ensure adequate provenance trail<br>•Handle preservation metadata<br>•Conduct Quality Assurance<br>•Provide reports and statistics |
| *"Migrate this set of images (in TIFF-5) to JP2 using ImageMagick 6.3 with parameters a,b,c"* | *•Analyze original*<br>*•Migrate, analyze output*<br>*•Conduct quality assurance*<br>*•Provenance, metadata, Reporting* |

# Preservation Planning Capabilities

- Planning Operational Preservation:
  - make drivers and goals operational and assess options against these criteria to deliver efficient decisions and operational plans
  1. Influencers and Decision Making
  2. Options Diagnosis
  3. Specification and Delivery
- Monitoring
  1. Internal Monitoring
  2. External Monitoring

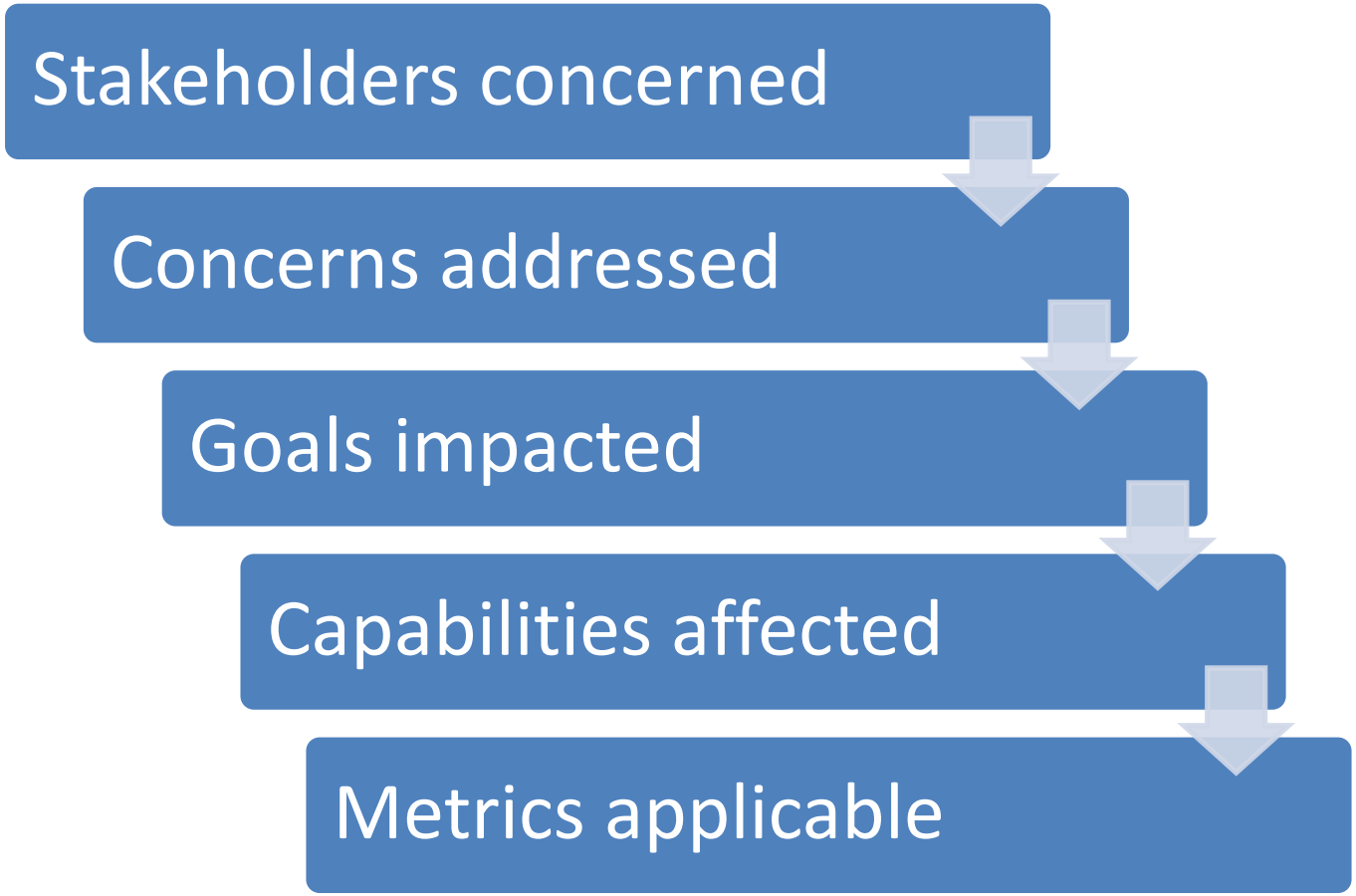
# Preservation Operation Capabilities

- Analysis
- Action
- Quality Assurance
- Preservation Metadata
- Plan Deployment
- Reporting and Statistics

# Example: Measuring Reporting and Statistics

- Timeliness
- Currentness
- Completeness
- Relevance
- Correctness
- Understandability

# Analysing drivers and constraints



# A Capability Maturity Model for Preservation Operation

The CMM has been shown to be a powerful instrument for assessment and improvement in Software Engineering

| | 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 | | | | | |

| | Awareness and Communication | Policies, Plans and Procedures | Tools and Automation | Skills and Expertise | Responsibility and Accountability | Goal Setting and Measurement |
|---|---|---|---|---|---|---|
| **1** | Need for operations recognized | Some uncontrol-led , undocumen-ted operations | Some tools employed ad-hoc | No aware-ness of skills / expertise | None assigned | Operations react to incidents, not tracked |
| **2** | Awareness of role of opera-tions, no formal reporting, individual communication | Operational procedures emerge, informal, intuitive. Indivi-dual procedures, ad-hoc QA | Automated tools used incoherently based on need and availability | Hands-on experience, informal peer training | Responsibility emerges, undocumented, no accountability | Individual awareness of short-term goals, no consistent definition |
| **3** | Understanding of role of operations, reporting guidelines not enforced | Defined process, driven by components and services. | Plans deployed according to spec, but mostly manual | Formal training plan, but training still individual initiative | Responsibility assigned, accountability not always provided | Operational goals specified, no formal metrics or alignment to goals |
| **4** | … | … | automated system controls automated operations | skills and expertise defined for all roles | Responsibility and accountability defined for all roles | Metrics aligned with goals, compliance enforced |
| **5** | … | … | … | … | RACI plan fully traceable to ops | … |

| | Awareness and Communication | Policies, Plans and Procedures | Tools and Automation | Skills and Expertise | Responsibility and Accountability | Goal Setting and Measurement |
|---|---|---|---|---|---|---|
| **1** | Need for operations recognized | Some uncontrol-led , undocumen-ted operations | Some tools employed ad-hoc | No aware-ness of skills / expertise | None assigned | Operations react to incidents, not tracked |
| **2** | Awareness of role of opera-tions, no formal reporting, individual communication | Operational procedures emerge, informal, intuitive. Indivi-dual procedures, ad-hoc QA | Automated tools used incoherently based on need and availability | Hands-on experience, informal peer training | Responsibility emerges, undocumented, no accountability | Individual awareness of short-term goals, no consistent definition |
| **3** | Understanding of role of operations, reporting guidelines not enforced | Defined process, driven by components and services. | Plans deployed according to spec, but mostly manual | Formal training plan, but training still individual initiative | Responsibility assigned, accountability not always provided | Operational goals specified, no formal metrics or alignment to goals |
| **4** | … | … | automated system controls automated operations | skills and expertise defined for all roles | Responsibility and accountability defined for all roles | Metrics aligned with goals, compliance enforced |
| **5** | … | … | … | … | RACI plan fully traceable to ops | … |

| | Awareness and Communication | Policies, Plans and Procedures | Tools and Automation | Skills and Expertise | Responsibility and Accountability | Goal Setting and Measurement |
|---|---|---|---|---|---|---|
| **1** | Need for operations recognized | Some uncontrol-led , undocumen-ted operations | Some tools employed ad-hoc | No aware-ness of skills / expertise | None assigned | Operations react to incidents, not tracked |
| **2** | Awareness of role of opera-tions, no formal reporting, individual communication | Operational procedures emerge, informal, intuitive. Indivi-dual procedures, ad-hoc QA | Automated tools used incoherently based on need and availability | Hands-on experience, informal peer training | Responsibility emerges, undocumented, no accountability | Individual awareness of short-term goals, no consistent definition |
| **3** | Understanding of role of operations, reporting guidelines not enforced | Defined process, driven by components and services. | Plans deployed according to spec, but mostly manual | Formal training plan, but training still individual initiative | Responsibility assigned, accountability not always provided | Operational goals specified, no formal metrics or alignment to goals |
| **4** | … | … | automated system controls automated operations | skills and expertise defined for all roles | Responsibility and accountability defined for all roles | Metrics aligned with goals, compliance enforced |
| **5** | … | … | … | … | RACI plan fully traceable to ops | … |

| | Awareness and Communication | Policies, Plans and Procedures | Tools and Automation | Skills and Expertise | Responsibility and Accountability | Goal Setting and Measurement |
|---|---|---|---|---|---|---|
| **1** | Need for operations recognized | Some uncontrol-led , undocumen-ted operations | Some tools employed ad-hoc | No aware-ness of skills / expertise | None assigned | Operations react to incidents, not tracked |
| **2** | Awareness of role of opera-tions, no formal reporting, individual communication | Operational procedures emerge, informal, intuitive. Indivi-dual procedures, ad-hoc QA | Automated tools used incoherently based on need and availability | Hands-on experience, informal peer training | Responsibility emerges, undocumented, no accountability | Individual awareness of short-term goals, no consistent definition |
| **3** | Understanding of role of operations, reporting guidelines not enforced | Defined process, driven by components and services. | Plans deployed according to spec, but mostly manual | Formal training plan, but training still individual initiative | Responsibility assigned, accountability not always provided | Operational goals specified, no formal metrics or alignment to goals |
| **4** | … | … | automated system controls automated operations | skills and expertise defined for all roles | Responsibility and accountability defined for all roles | Metrics aligned with goals, compliance enforced |
| **5** | … | … | … | … | RACI plan fully traceable to ops | … |

| | Awareness and Communication | Policies, Plans and Procedures | Tools and Automation | Skills and Expertise | Responsibility and Accountability | Goal Setting and Measurement |
|---|---|---|---|---|---|---|
| **1** | Need for operations recognized | Some uncontrol-led , undocumen-ted operations | Some tools employed ad-hoc | No aware-ness of skills / expertise | None assigned | Operations react to incidents, not tracked |
| **2** | Awareness of role of opera-tions, no formal reporting, individual communication | Operational procedures emerge, informal, intuitive. Indivi-dual procedures, ad-hoc QA | Automated tools used incoherently based on need and availability | Hands-on experience, informal peer training | Responsibility emerges, undocumented, no accountability | Individual awareness of short-term goals, no consistent definition |
| **3** | Understanding of role of operations, reporting guidelines not enforced | Defined process, driven by components and services. | Plans deployed according to spec, but mostly manual | Formal training plan, but training still individual initiative | Responsibility assigned, accountability not always provided | Operational goals specified, no formal metrics or alignment to goals |
| **4** | … | … | automated system controls automated operations | skills and expertise defined for all roles | Responsibility and accountability defined for all roles | Metrics aligned with goals, compliance enforced |
| **5** | … | … | … | … | RACI plan fully traceable to ops | … |

# Conclusion

- Capability-based Reference Architecture
  - Supports deployment of preservation *capability* to any scenario
  - Enables customized, technology-independent architecture
  - Facilitates organizational change across technology generations
- Contextual Enterprise Architecture approach
  - Improves Separation of Concerns
  - Enables traceability
- Capability maturity model
  - Provides Decision support mechanism
  - Supports prioritizing improvements
  - Guides gap analysis
  - Facilitates management buy-in

# Thank you!

- Questions?

- Related reading:
  **Control Objectives for DP: Digital Preservation as an Integrated Part of IT Governance** published in the proceedings of the 74th Annual Meeting of the American Society for Information Science and Technology (ASIST)
- www.ifs.tuwien.ac.at/~becker