V1.2.2

Big Data, Small Data, the Digital Thread, and Product Lines: Leveraging Existing Assets Using S*Metadata

Bill Schindel

ICTT System Sciences

schindel@ictt.com

11th Annual INCOSE Great Lakes Regional Conference

SUPERIOR SYSTEM SOLUTIONS FOR TODAY'S COMPLEX ENVIRONMENTS

11 - 14 October 2017 Twin Cities, Minnesota

Abstract: The triumphs of the physical sciences, and their exploitation by engineering disciplines founded on them, have powered much of the acceleration of progress in human life during the last three centuries. Central to this progress has been the extraction of simplifying patterns from the bewildering complexity of Nature, yielding powerful laws that predict, explain, and permit synthesis of engineered systems. Today we are seeing a recapitulation of this history, as information technology enables both Big Data and its projection onto System Patterns that compress and extract meaning.

In this talk, the take-aways we will summarize are how S*Patterns from the world of Model-Based Systems Engineering (MBSE) are used (1) to accelerate the federation of the Digital Thread using existing enterprise databases, engineering tools, telemetry and information systems, and the data they already contain, (2) to harvest MBSE product line engineering (PLE) patterns from existing legacy product information, and (3) to enhance effective life cycle collaboration between different organizations and specialists. Targeted audiences include practicing systems engineers, business and IT strategists, process owners, and organizational leaders.

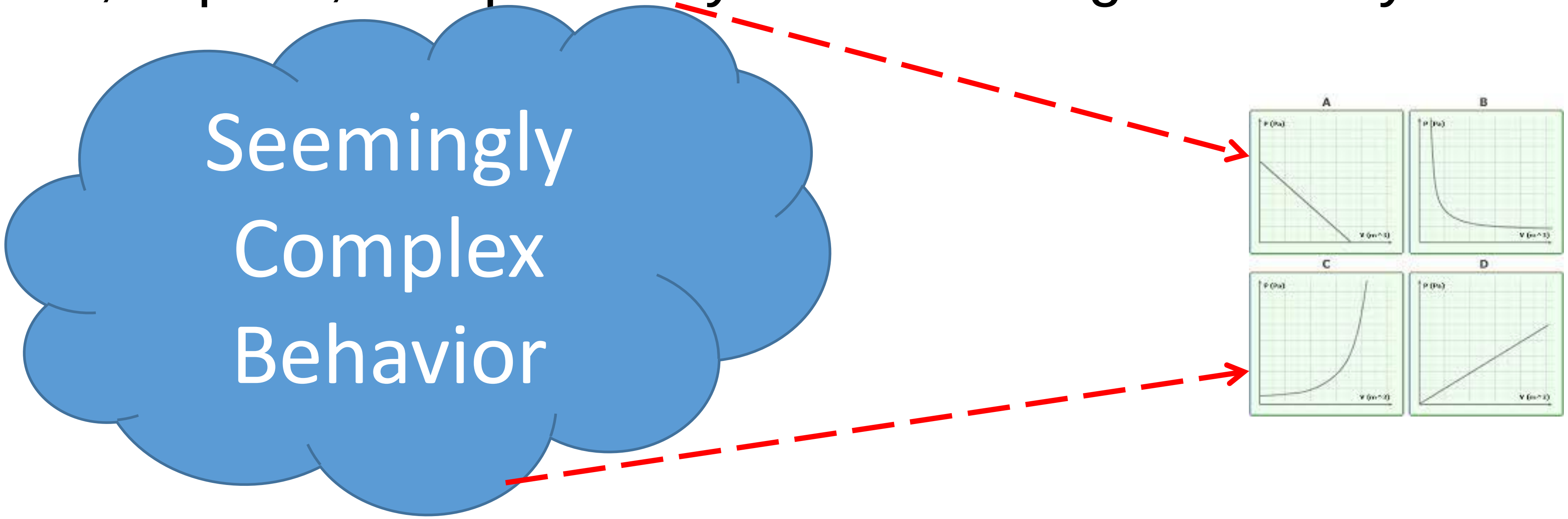
Contents

Compression: Sense-making in an increasingly complex world

- Human-engineered and human-impacted systems are increasing in complexity:
  - Global information networks
  - Systems for distribution of goods and services
  - Agriculture to feed a growing population from fixed lands
  - Economic systems—regional, national, global
  - Urban infrastructure systems
  - Health care delivery and epidemiology
  - Proliferation of cyber-physical systems and the Internet of Things
  - Local, regional, global transportation systems
  - Systems of innovation, engineering, advanced manufacturing, and sustainment
  - Government, other sociotechnical systems
  - Defense systems
- How to make sense, manage, and succeed, across system life cycles, in the face of exploding complexity?

Patterns compress and order complexity

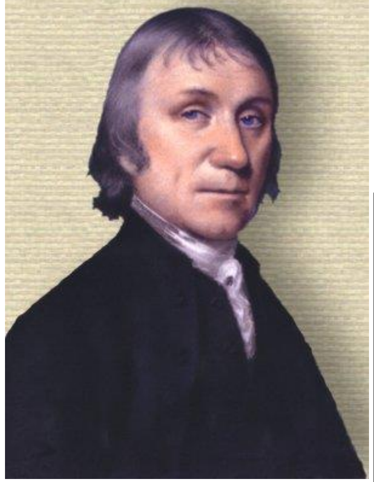
- The triumphs of the physical sciences, and their exploitation by engineering disciplines founded on them, have powered much of the acceleration of progress in human life during the last three centuries.
- Central to this progress has been the extraction of <u>simplifying patterns</u> from the bewildering complexity of Nature, yielding powerful laws that predict, explain, and permit synthesis of engineered systems:



Predictive or explanatory pattern

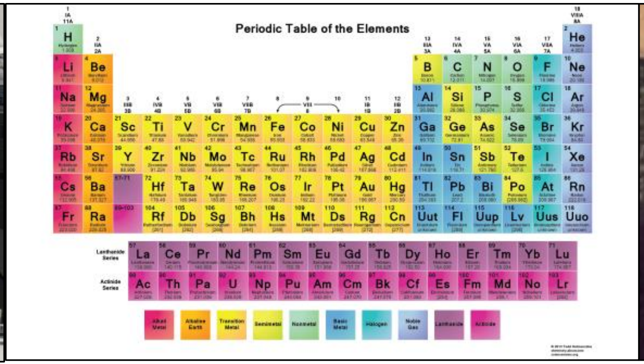
- Today we are seeing a recapitulation of this history, as information technology enables both Big Data and its projection onto System Patterns that compress and extract meaning for the pragmatic uses of society.

Historical Example: Chemistry

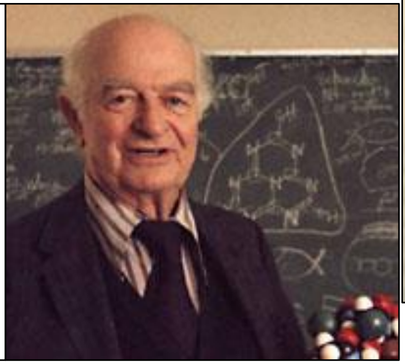


Priestley : Oxygen

Modern Chemist



Periodic Table of the Elements

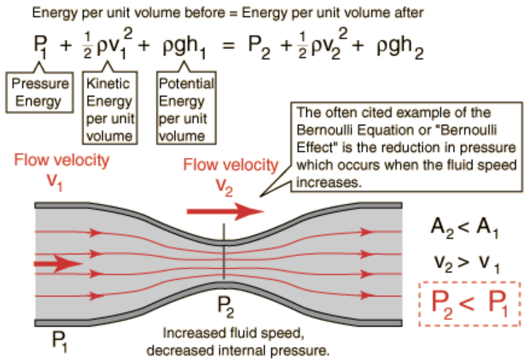


Pauling: Chemical Bond

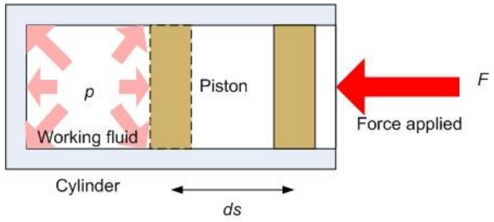
Mendeleev: Periodic Table

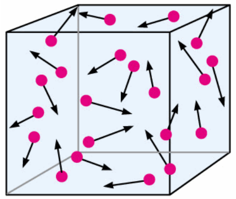
- The transition from rhetoric and mystery of alchemy to the impactful success of chemistry was driven by finding <u>models</u> that compressed the bewildering complexity of nature and behavior of chemically interacting matter into an understandable modeled explanatory and predictive framework:
  - A view that emerged from the scientific discovery and verification of laws of Chemistry.
  - Chemical Elements and their Chemical Properties, organized by the discovered patterns of the Periodic Table.
  - Chemical Bonds, Chemical Reactions, Reaction Rates, Chemical Energy, Conservation of Mass and Energy.
  - Chemical Compounds and their Properties.

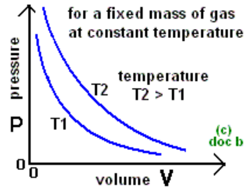
Historical Example: The Gas Laws



$$\frac{N_b}{N_a} = \left(\frac{g_b}{g_a}\right)\left(e^{-(E_b - E_a)/kT}\right)$$







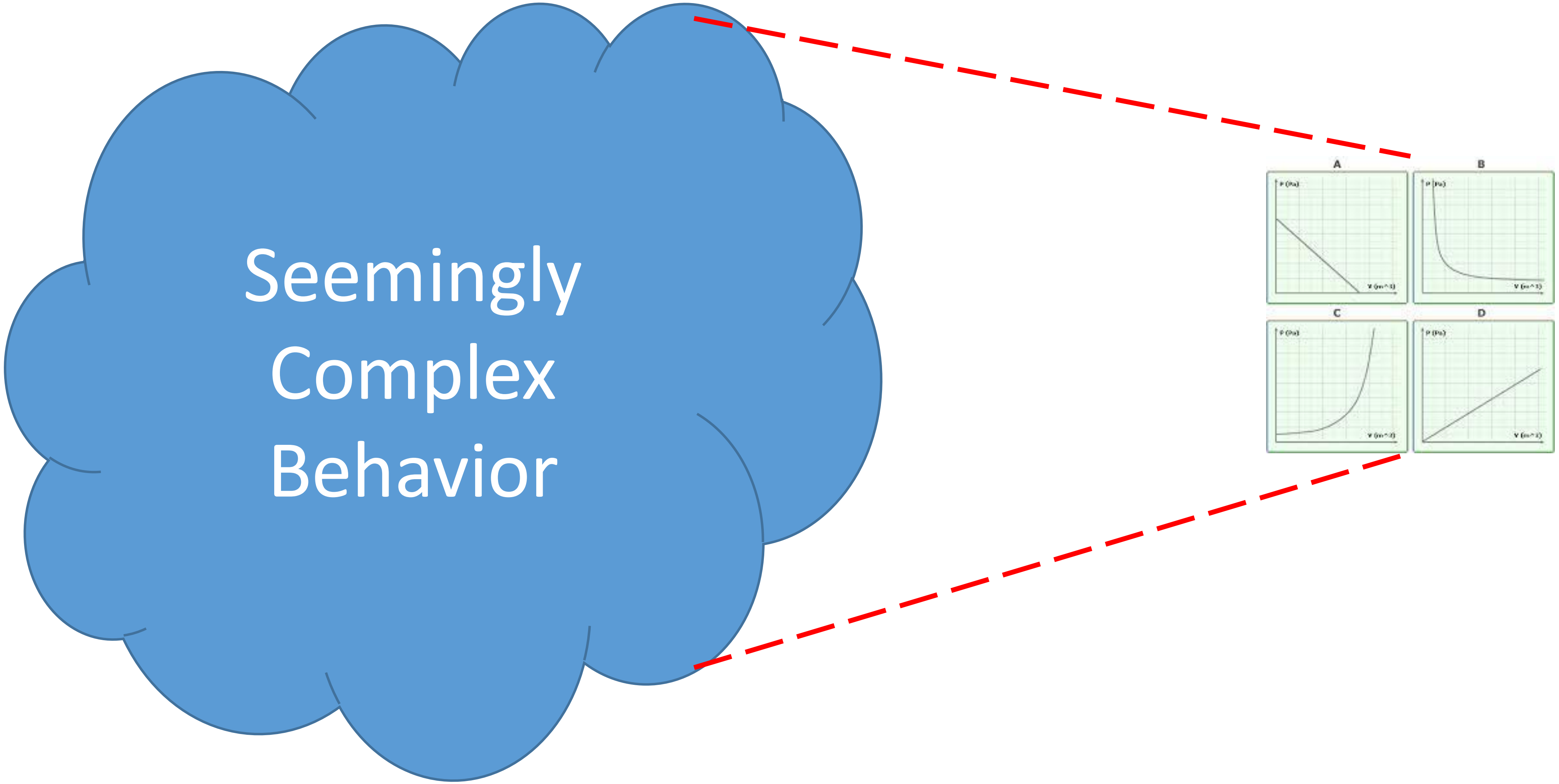
Boyle

Boltzmann

Daniel Bernoulli

- The discovered and verified laws of gases and of compressible and incompressible fluid flow by Boyle, Avogadro, Charles, Gay-Lussac, Bernoulli, and others are rightly viewed as fundamental to science and engineering disciplines:
  - Pressure, Temperature, Flow
  - Laws of Thermodynamics
  - As in the case of Chemistry, the underlying models discovered were based upon characterizing interactions of components entities, and the emergent character of the larger observed systems.

Patterns compress and organize apparent complexity

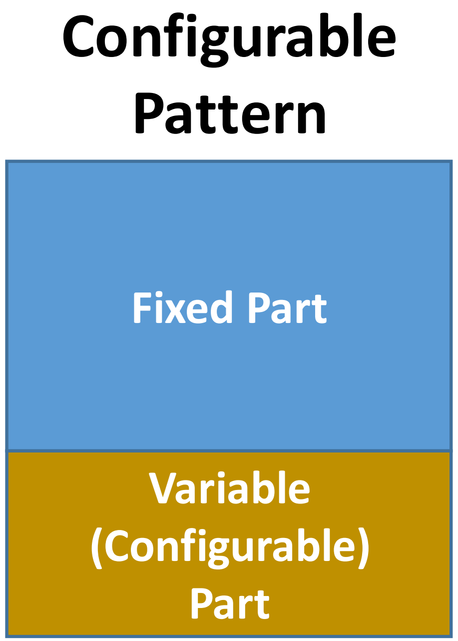


Predictive or explanatory pattern

But, this does not mean that everything our community has been calling "System Models" are credible and capable of that compression.

Patterns compress and organize complexity

- Patterns are recurrences, across time, space, or other indices.
- Most patterns have a fixed (recurring) part and a varying (different in various instances) part, e.g.:
  - the Gas Laws apply to multiple gaseous measurements, with repeating quantitative relationships (the formulae), but varying instance values for the state variables (pressure, volume, temperature, flow rate, density).
  - Database models apply to each record and its relationship to other records, over and over, but the values of data within the records varies.
  - We will see that Patterns provide compression, in the form of the variable part, with the fixed part receding as background.
- Example: The Ethernet 802.xx specification describes a general pattern for local area network use:
  - It can be configured for different bandwidth and other network parameters.
  - This allows us to say (in compressed domain specific language) that we have a “100 MB Ethernet network”.
  - This is much more compressed than repeating the whole Ethernet specification!



Patterns compress and organize complexity

Lawnmower Product Line: Configurations Table

		Units	Walk-Behind	Walk-Behind	Walk-Behind	Riding	Riding	Riding Mower	Autonomous
			Push Mower	Mower	Self-Propelled	Rider	Tractor	Tractor	Autonomous
			Push Mower	Self-Propelled	Wide Cut	Rider	Lawn	Garden	Auto Mower
	Model Number		M3	M5	M11	M17	M19	M23	M100
	Market Segment		Sm Resident	Med Resident	Med Resident	Lg Resident	Lg Resident	Home Garden	High End Suburban
Power	Engine Manufacturer		B&S	B&S	Tecumseh	Tecumseh	Kohler	Kohler	Elektroset
	Horsepower	HP	5	6.5	13	16	18.5	22	0.5
Production	Cutting Width	Inches	17	19	36	36	42	48	16
	Maximum Mowing Speed	MPH	3	3	4	8	10	12	2.5
	Maximum Mowing Productivity	Acres/Hr			1.6				
	Turning Radius	Inches	0	0	0	0	126	165	0
	Fuel Tank Capacity	Hours	1.5	1.7	2.5	2.8	3.2	3.5	2
	Towing Feature						x	x	
	Electric Starter Feature				x	x	x	x	
	Basic Mowing Feature Group		x	x	x	x	x	x	x
Mower	No. of Anti-Scalping Rollers		0	0	1	2	4	6	0
	Cutting Height Minimum	Inches	1	1.5	1.5	1.5	1	1.5	1.2
	Cutting Height Maximum	Inches	4	5	5	6	8	10	3.8
	Operator Riding Feature					x	x	x	
	Grass Bagging Feature		Optional	Optional	Optional	Optional	Optional	Optional	
	Mulching Feature		Standard	Factory Installed	Dealer Installed				
	Aerator Feature					Optional	Optional	Optional	
	Autonomous Mowing Feature								x
	Dethatching Feature					Optional	Optional	Optional	
Physical	Wheel Base	Inches	18	20	22	40	48	52	16
	Overall Length	Inches	18	20	23	58	56	68	28.3
	Overall Height	Inches	40	42	42	30	32	36	10.3
	Width	Inches	18	20	22	40	48	52	23.6
	Weight	Pounds	120	160	300	680	705	1020	15.6
	Self-Propelled Mowing Feature			x	x	x	x	x	x
	Automatic TransmFeature							x	
Financials	Retail Price	Dollars	360	460	1800	3300	6100	9990	1799
	Manufacturer Cost	Dollars	120	140	550	950	1800	3500	310
Maintenance	Warranty	Months	12	12	18	24	24	24	12
	Product Service Life	Hours	500	500	600	1100	1350	1500	300
	Time Between Service	Hours	100	100	150	200	200	250	100
Safety	Spark Arrest Feature		x	x	x	x	x	x	

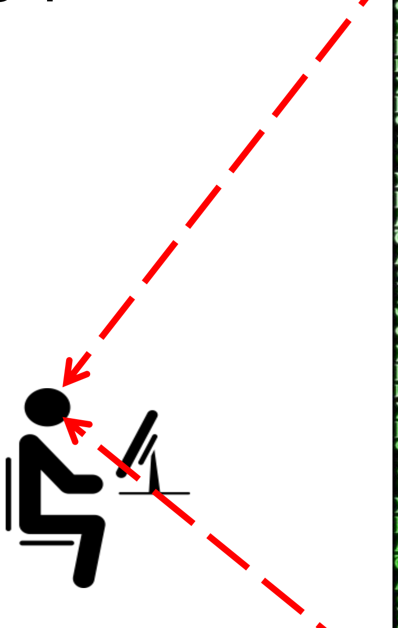
Configurable Pattern

Fixed Part

Variable (Configurable) Part

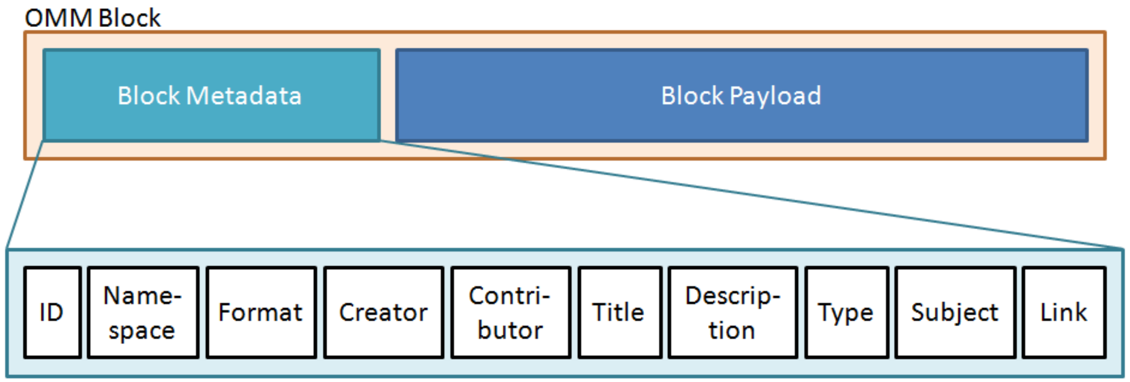
Big Data, Small Data, and Metadata

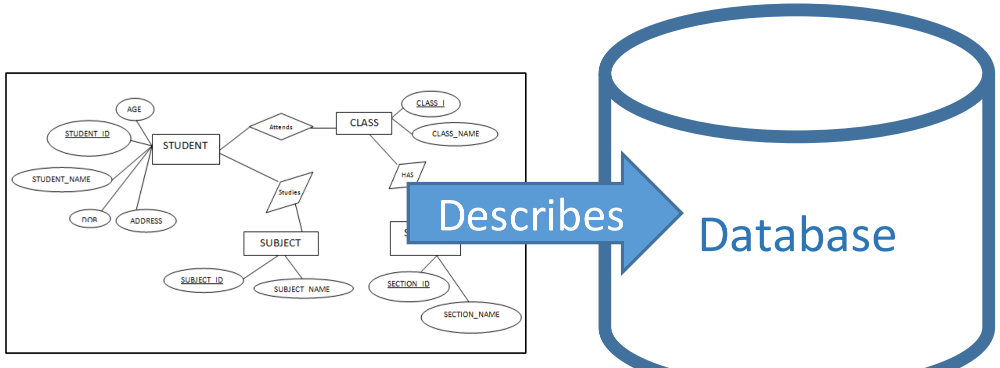
- Typically, we say that “Big Data” is . . .
  - Very large information system data sets, beyond what previous generation information technologies (e.g., relational databases) can effectively handle;
  - Relatively less uniformly structured data, such as messages, social media data, web pages, dissimilar records;
  - Challenging to visualize using earlier generation data views and viewers;
  - Data sets that might be processed by “predictive analytics” algorithms to extract (“learn”) underlying patterns.



Big Data, Small Data, and Metadata

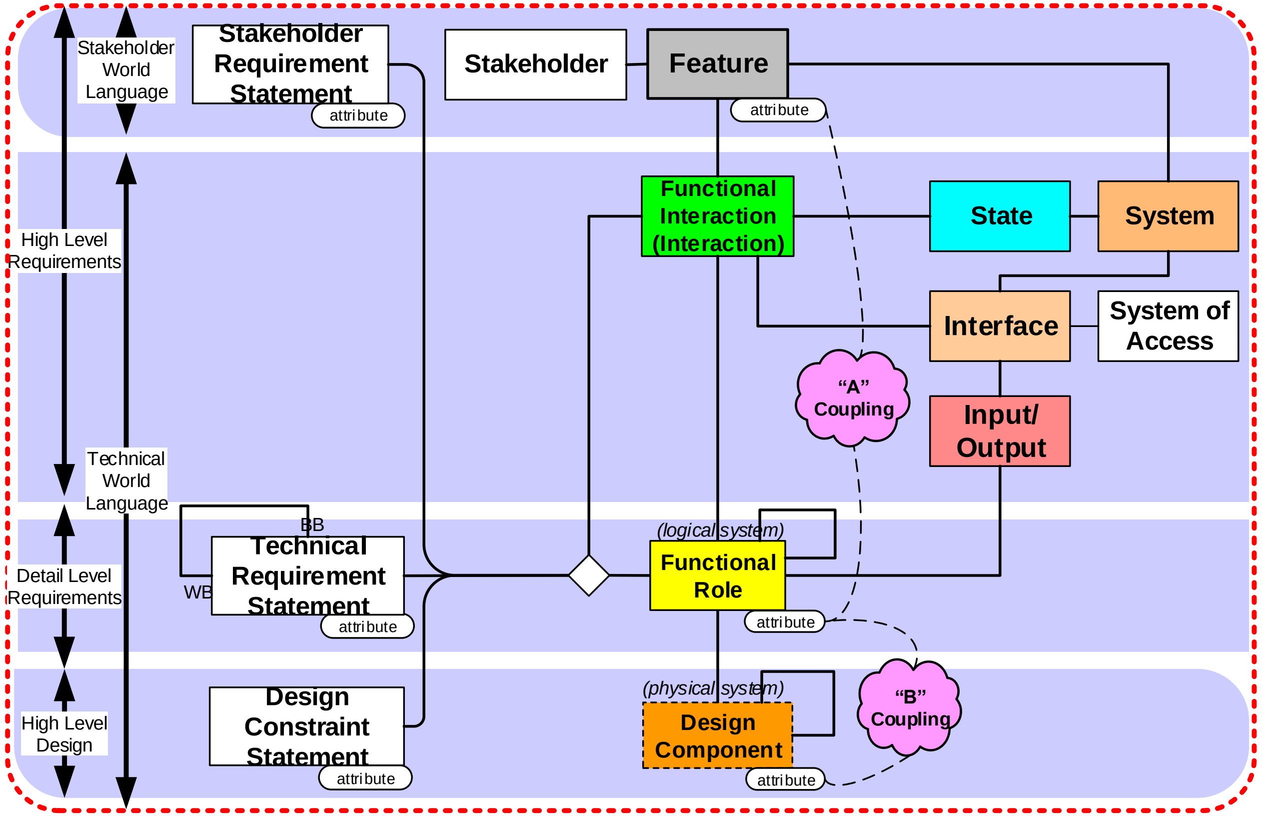
- Typically, we say that “Metadata” is “data about other data”, as in:
  - Database models that describe other data
  - In messages (e.g., phone calls or email message), separated from the “payload” of the message, describing other facts about it, such as origin, destination, time, priority
  - Information about related data ownership, type, time, or other aspects of related data
  - Reference to the relevant S*Pattern elements describing some aspects of an S*Model or other data described by an S*Model.
- The Metadata may be dispersed into and stored or transported with the data it describes, or may be elsewhere but referenced by the data it describes



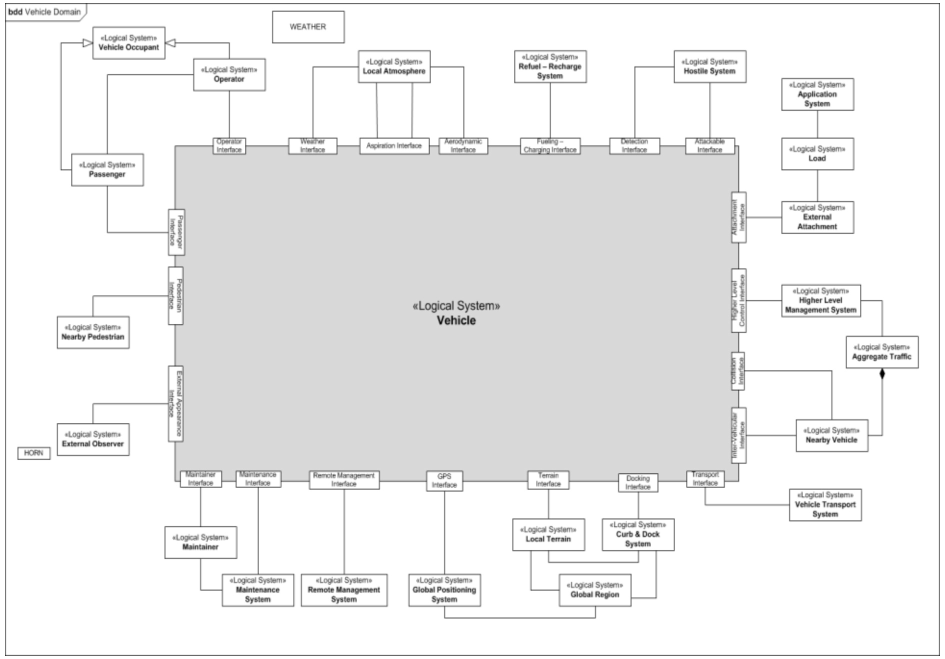


E-R Data Model

- The S*Metamodel:
  - is a model of other models, but language and tool independent (maps to any)
  - describes the smallest amount of information necessary for purposes of engineering or science or life cycle management.
  - is Metadata with respect to any system model it describes.
- An S*Model is any model that corresponds to the S*Metamodel

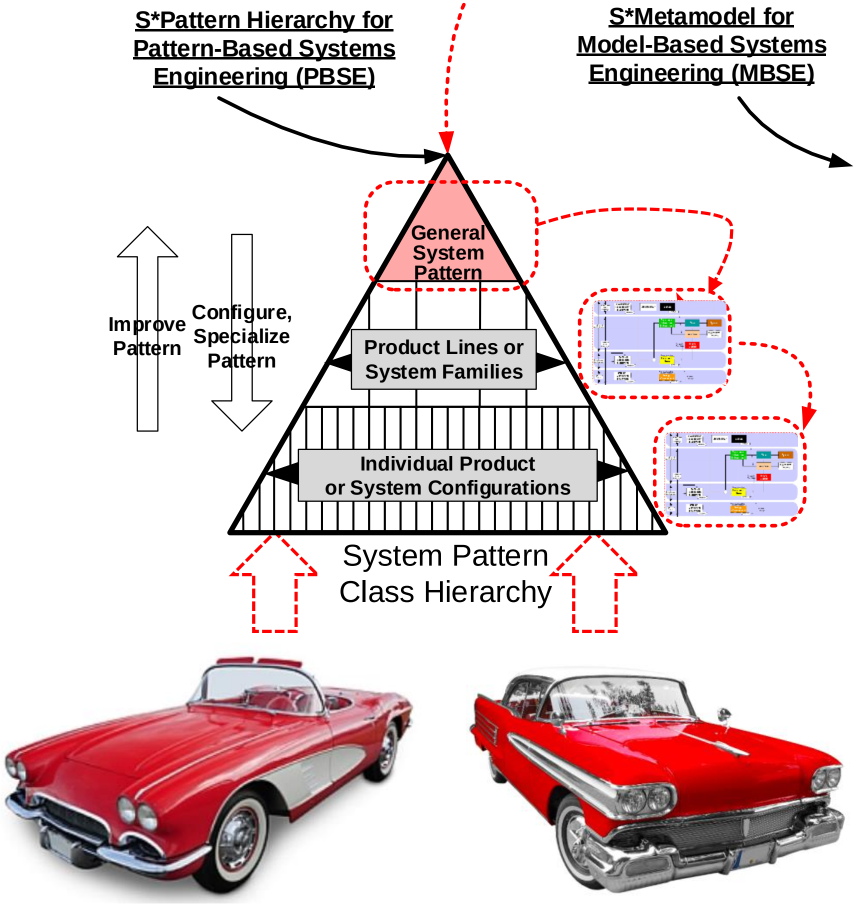


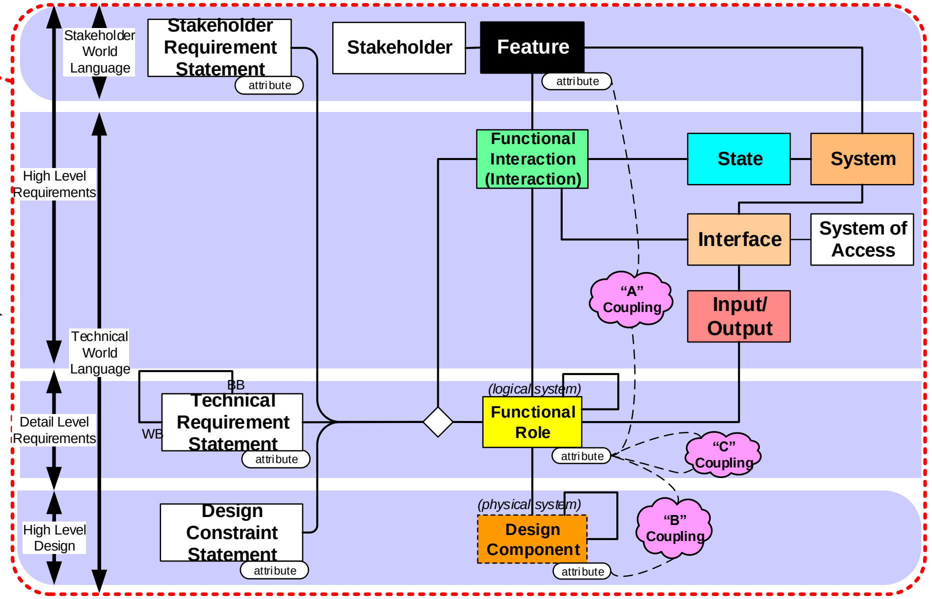
Summary Extract of S*Metamodel



An S*Model, Domain Diagram

Big Data, Small Data, and Metadata





- An S*Pattern is a general, configurable, re-usable S*Model that describes a family of systems, and which can be configured as specific S*Models for different members of the family.
- Each of them provides a compression of the type, consisting of the variable (configurable) values.

Big Data, Small Data, and Metadata

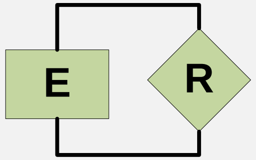
- An S*Pattern is metadata with respect to any specific S*Model configured from the S*Pattern.
- The S*Metamodel is itself an S*Pattern.
- So, the S*Metamodel, S*Models, and S*Patterns are all Metadata with respect to the more specific models they each describe.
- Each describes a Domain Specific Language (DSL) for its own abstract or specific domain, so that “sentences” (models) can be expressed in that DSL . . .

Emergence of Patterns from Patterns: S*Pattern Class Hierarchy

More General

Definition of **Relational Modeling Paradigm**

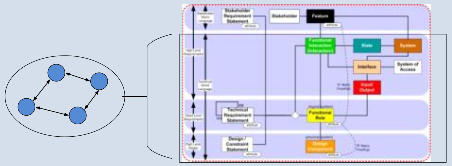
Entity-Relationship Paradigm



E=Entity
R= Relationship

Minimal System S*Metamodel: Definition of (Elementary) System, Material Cause

S*Metamodel Core

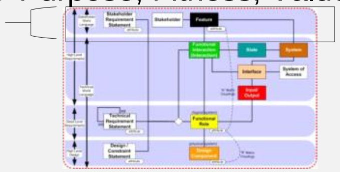


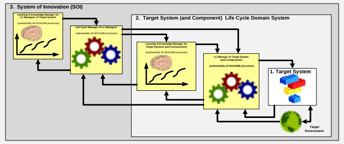
Core S*Metamodel

Emergence & Definition of **System of Innovation**, Fitness, Value, Purpose, Stakeholders, Agility, Final Cause, Formal Cause, Efficient Cause, Intelligence, Management, Science, Living Systems

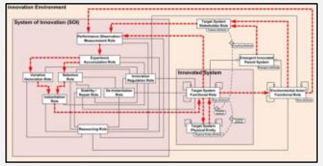
EI Pattern, SOI Pattern, Fitness, Value

S*Purpose, Fitness, Value

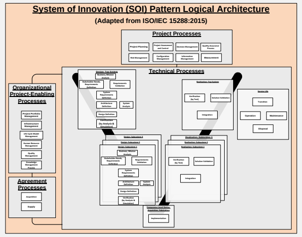




Agile Sys Life Cycle Pattern



System of Innovation Pattern



ISO 15288 System Life Cycle Mgmt Pattern

Emergence & Definition of **Domain Specific Systems**

Domain Specific Pattern

- **Terrestrial Vehicle Domain Pattern**
- **Manufacturing System Pattern**
- **Product Service System Pattern**
- **Space Tourism Pattern**
- **…**
- **Aircraft Flight Control Pattern**
- **Medical Device Pattern**
- **Distribution System Pattern**
- **Socio-Technical Pattern**
- **…**

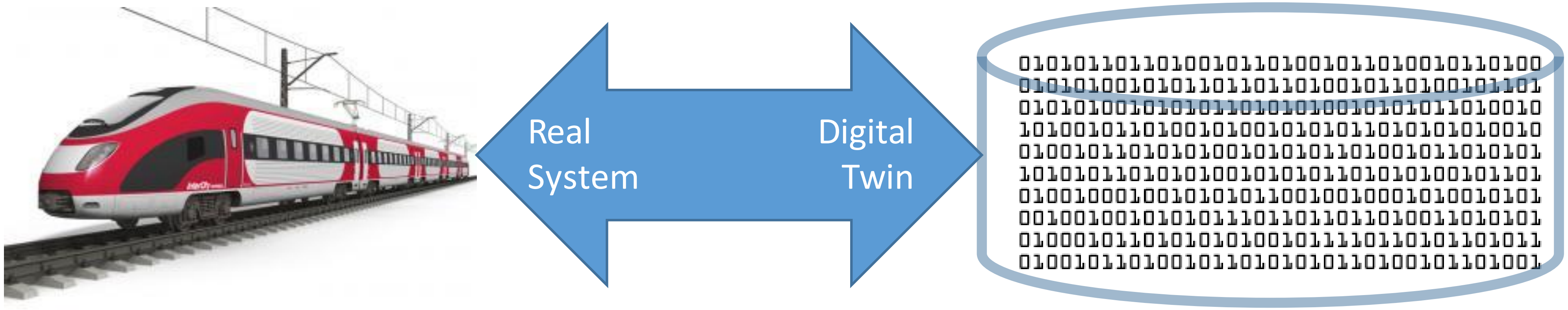
More Specific

The Digital Twin vs. diverse life cycle systems data

What is a Digital Twin?

The Digital Twin vs. diverse life cycle systems data

- By “digital twin”, we mean:
  - A (digital) data representation of a system of interest
  - Capturing all the aspects of interest over the system’s life cycle
  - Including time-based dynamic behavioral as well as static aspects
  - For use in the life cycle management of the system



The Digital Twin vs. diverse life cycle systems data

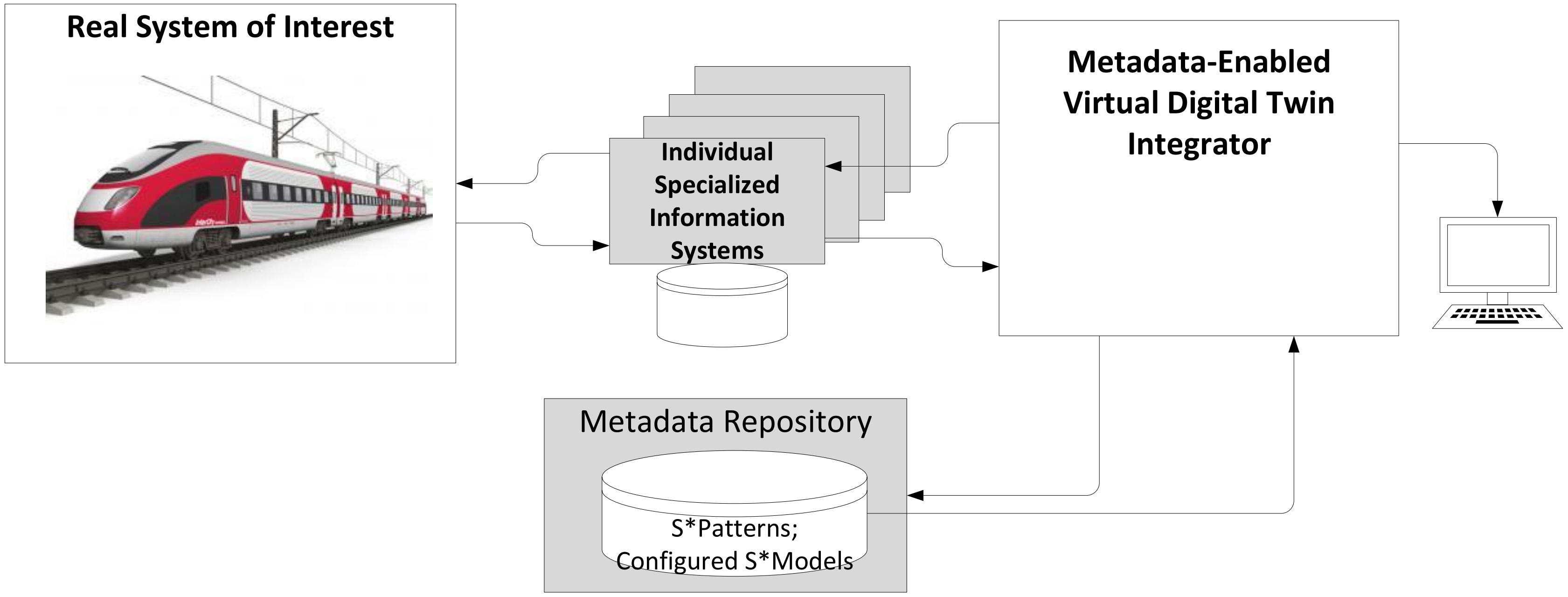
- For all the interest in it, the term “digital twin” is relatively new.
- There are many examples of engineering tools and life cycle management systems that cover partial “slices” of information:
  - Requirements tools
  - CAD tools and tool chains
  - Production recipes and databases
  - Maintenance and trouble recording systems
  - PDM and PLM systems
  - Performance data recording and analysis systems
  - Configuration management systems
  - Etc.
- It is arguable as to whether there are (so far) real industry examples of digital twins that cover the full range of dynamic and static historical data of all types covering life cycle management, in any single system.

The Digital Twin vs. diverse life cycle systems data

- A perceived obstacle to such a fully integrated Digital Twin is that the existing available data is spread across many dissimilar information systems:
  - The perceived challenge is not just that the systems are separate . . .
  - But that their information models (semantics) are also completely different.
  - So that a user wanting to navigate the life cycle data in an integrated way is frustrated by these apparent barriers.
- What to do?

The Digital Twin vs. diverse life cycle systems data

Solution: S*Metadata provides virtual integration of a Digital Twin across existing "incompatible" data systems:

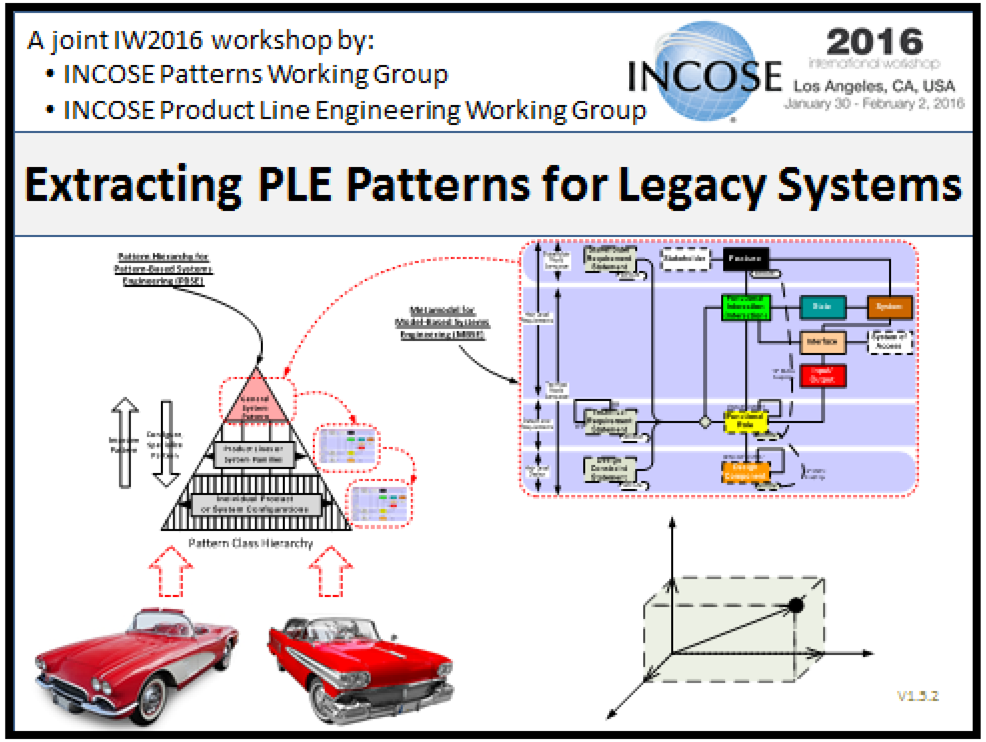


Harvesting product line simplicity from legacy complexity, through compression

- Legacy product lines grow up over years of efforts by multiple teams and acquisitions
  - These become difficult to manage as effectively as a platform-based product line that has been developed from the beginning using Product Line Engineering (PLE) methods
  - But, we don't always have the option of starting from scratch to create a new product line
  - How to harvest a pattern-based product line from a set of legacy products?
- The answer is to extract a common S*Pattern from a set of legacy products, by the Method of Projections . . .

Harvesting product line simplicity from legacy complexity, through compression

Projections, the INCOSE PLE WG – Patterns WG joint project:

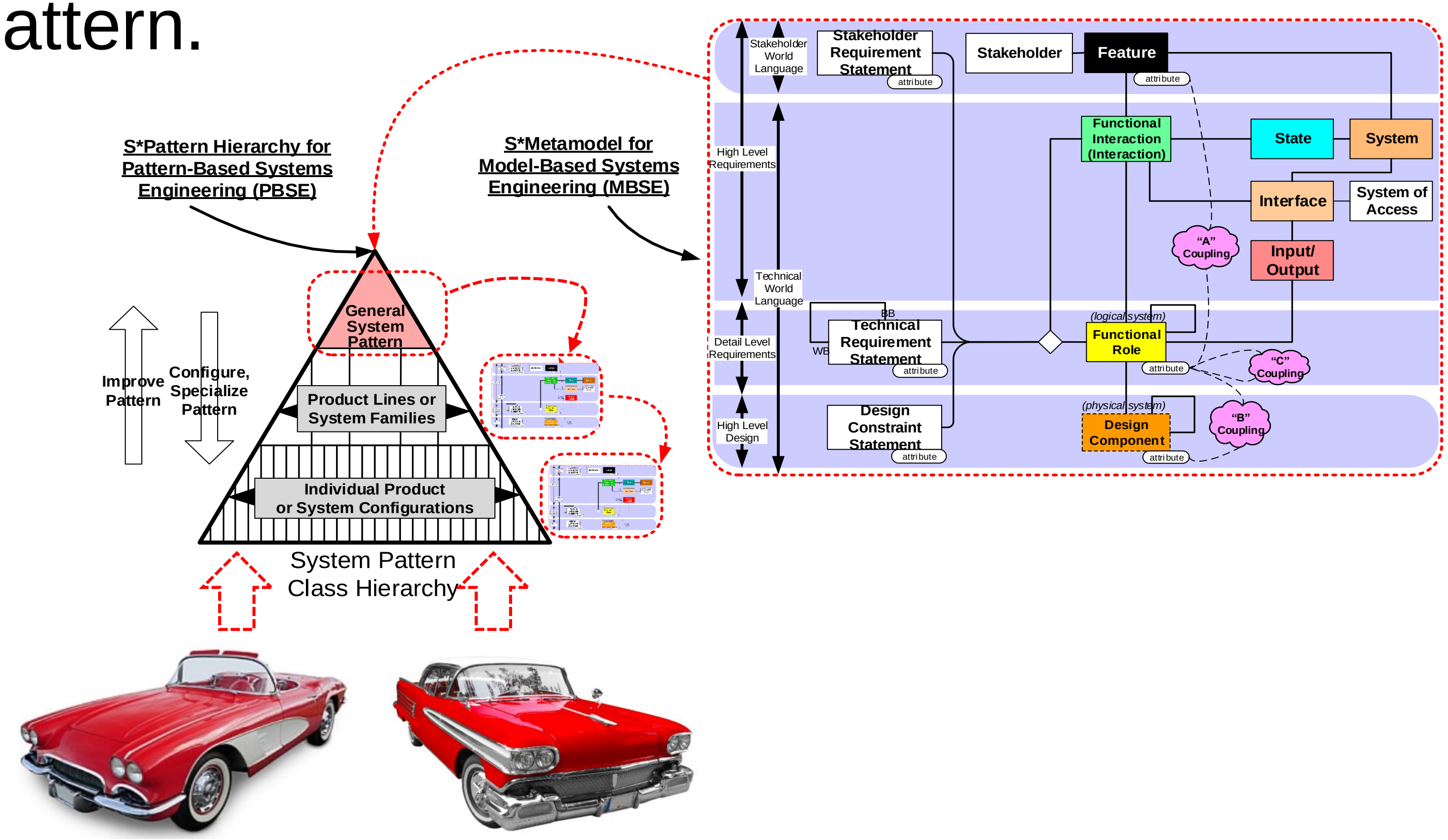


Knocking down organizational and marketplace walls, integrating stovepipes, and improving collaboration

- The following is the human equivalent problem of the previous IT systems problem . . .
- It is a well-known challenge that different organizations within an enterprise can become “silos” that don’t interact effectively, from the perspective of a customer or other stakeholder.
  - Likewise, different segments of a market or domain can become silo’ed in a ways that are disadvantageous to the global outcome.
  - One reason for this is the silo’ing of the languages and frameworks of communication and interaction of these segments.
- How to extract of common sheet of music when there are different frameworks in use by the different players?

Knocking down organizational and marketplace walls, integrating stovepipes, and improving collaboration

"Uncover the Pattern" is a S*Pattern-facilitated process for group discovery of the underlying common portion of a unifying pattern, as well as the legitimate differences that should be retained through the configurable part of the pattern.



Knocking down organizational and marketplace walls, integrating stovepipes, and improving collaboration

- Note especially that credibility of (trust in) the models/patterns is of particular importance:
  - Model credibility has become a key issue as models become more prominent.
  - V4I: the Virtual Verification, Validation, and Visualization Institute
- See also the session in this conference on collaboration across professional societies and regulators, facilitated by uncovered patterns

What you can do

- Join the activities of the INCOSE Patterns Working Group—bring a project of your own if you'd like.
- Participate in the S*Patterns Community and learn about the S*IP Landscape
- Conduct an Uncover the Pattern (UTP) project.
- Assess the credibility of your model with the V4I Institute.

Discussion, Questions

-
-
-
-
-

References

1. Schindel, W., “What Is the Smallest Model of a System?”, *Proc. of the INCOSE 2011 International Symposium*, International Council on Systems Engineering (2011).
2. Schindel, W., “Got Phenomena? Science-Based Disciplines for Emerging System Challenges”, in *Proc. of 2016 International Symposium*, International Council on Systems Engineering, Edinburgh, UK, 2016.
3. INCOSE Patterns Working Group web site, at http://www.omgwiki.org/MBSE/doku.php?id=mbse:patterns:patterns
4. INCOSE Patterns Working Group, “Pattern-Based Systems Engineering (PBSE), Based On S*MBSE Models”, INCOSE PBSE Working Group, 2015: http://www.omgwiki.org/MBSE/doku.php?id=mbse:patterns:patterns_challenge_team_mtg_06.16.15
5. V4 Institute Web Site: http://v4i.us/
6. Uncover the Pattern summary: https://www.ictt.com/UTP.html

GLRC 11

Speaker

Bill Schindel chairs the MBSE Patterns Working Group of the INCOSE/OMG MBSE Initiative. He is president of ICTT System Sciences, and has practiced systems engineering for over thirty years, across multiple industry domains. Bill serves as president of the INCOSE Crossroads of America Chapter, and is an INCOSE Fellow and Certified Systems Engineering Professional. An ASME member, he is part of the ASME VV50 standards team's effort to describe the verification, validation, and uncertainty quantification of models.