

Empirical Methods in Strategic Software Engineering

Daniel Port
Information Technology Management
University of Hawaii at Manoa

1

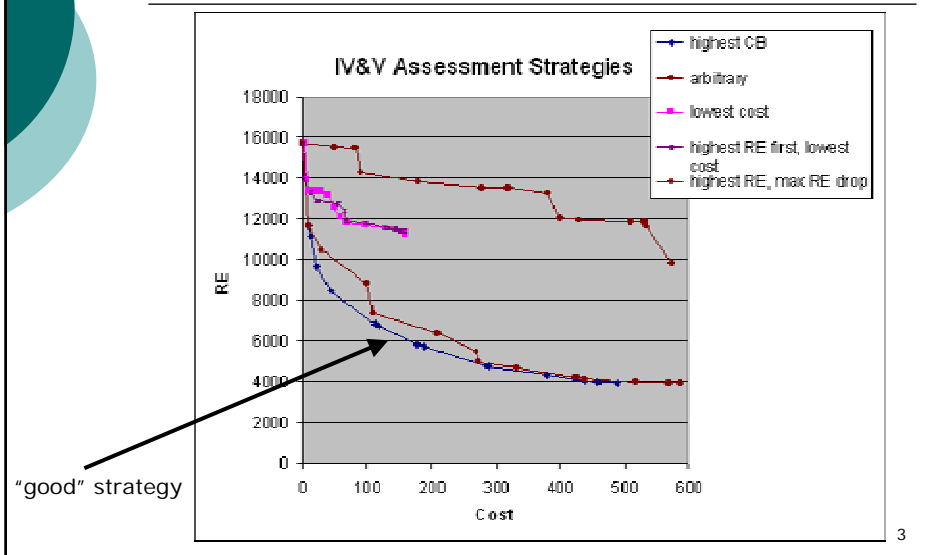
What problems am I tackling?

- Strategic decision making and planning of software engineering activities
 - Strategic
 - Activities explicitly planned to achieve desired goals *on average*
 - Examples: Maximize risk-reduction at minimum cost, "how much [testing, IV&V, architecture flexibility, COTS assessment, security, ...] is enough?", return on investment (ROI)
 - Tactical
 - Small-scale actions serving to *contain or respond*, made or carried out with only a limited or immediate end in view
 - Top-10 risk monitoring, Risk contingency plans, XP

We are seeking strategic models that help determine good strategies

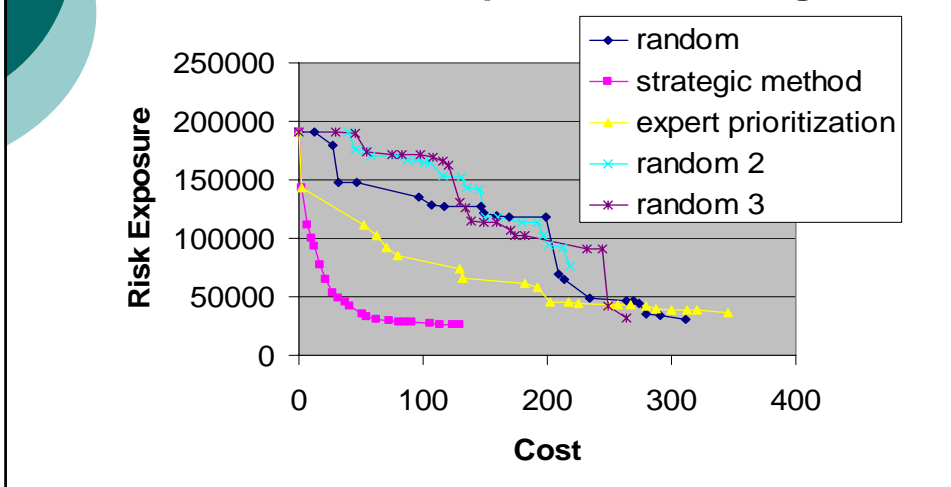
2

Strategy Matters!



Strategic Planning is not Easy for People

Comparison of Strategies



How am I working to solve them?

- “Scorecard” models to convert qualitative data to meaningful quantitative data

- $P(A_i) = c_1 * c_2 * \dots * c_{13} * 2^{-13}$

- $EF(T_j) = (1 + c_1)^2 * (1 + c_2)^3 * \dots * (1 + c_{13})^8 * d_1 * d_2 * \dots * d_8 * 2^{-13}$

Activity	Technique	System characteristics				NIV Content			Average effort	
		cc: Accuracy	cc: Peak release	cc: Control availability	cc: Target responsiveness	cc: Development time to contract	cc: Assessment of assessment tool	cc: Support	cc: Error	cc: Mean/ mode/ range/ page
R2: Requirement Precision	T2: SpectRM Modeling	Requirement 1	HT 1	Control controller 1	Requirement 1	Requirement 1	Software 1	1	1	
		Non-Automouse 1	HT 1	Real effect controller 1	Requirement 0.5	Design 1	Software 0.5			
T4: Network language modeling	Non-Automouse 1	Requirement 1	HT 1	Control controller 1	Requirement 1	Design 1	Software 1	0.5	2	
		HT 1	HT 1	Real effect controller 1	Requirement 1	Design 1	Software 1			

- Hackstat for automated data collection
 - www.hackstat.org
- The Strategic method

$$\min_{\tau, J} \left[\sum_{i=1}^k RE_{\text{after}}(A_i, T_{J(i)}) + RRCost(A_i, T_{J(i)}) + \sum_{i=k+1}^N RE_{\text{before}}(A_i) \right]$$

5

How does this apply or help Japanese software industry?

Three current relevant Strategic Software Engineering research projects

1. Strategic IV&V
2. Critical factors for adoption of custom developed software
3. Software telemetry inputs for strategic SWE tools

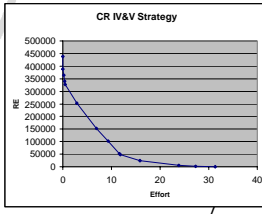
6

Current Strategic SWE projects (1)

Characteristic	Value (select box)
C1: 機能1 品質保証の考え方	逐次開発
C2: 機能2 要求される操作性	逐次開発
C3:	
C4:	
select	
Attribute	
CA: Y	A1: 適用ソフトウェアの操作性
CA: Y	A2: 適用ソフトウェアの要求仕様との適合性
CA: Y	A3: 開発者の経験と知識の豊富さ
Value (select box)	
Constraint	
CD: Y	D1: 開発期間 / プロジェクト / リソース / 予算
CD: Y	D2: IV&Vを実施する開発段階
CD: Y	D3: Access to source code
AD: Y	A4: 開発者の経験と知識の豊富さ
AD: Y	A5: 開発者の経験と知識の豊富さ
AD: Y	A6: 開発者の経験と知識の豊富さ
AD: Y	A7: 開発者の経験と知識の豊富さ
AD: Y	A8: 開発者の経験と知識の豊富さ
AD: Y	A9: 開発者の経験と知識の豊富さ
AD: Y	A10: 開発者の経験と知識の豊富さ
AD: Y	A11: 開発者の経験と知識の豊富さ
AD: Y	A12: 開発者の経験と知識の豊富さ
AD: Y	A13: 開発者の経験と知識の豊富さ
AD: Y	A14: 開発者の経験と知識の豊富さ
AD: Y	A15: 開発者の経験と知識の豊富さ
AD: Y	A16: 開発者の経験と知識の豊富さ
AD: Y	A17: 開発者の経験と知識の豊富さ
AD: Y	A18: 開発者の経験と知識の豊富さ
AD: Y	A19: 開発者の経験と知識の豊富さ
AD: Y	A20: 開発者の経験と知識の豊富さ
AD: Y	A21: 開発者の経験と知識の豊富さ
AD: Y	A22: 開発者の経験と知識の豊富さ
AD: Y	A23: 開発者の経験と知識の豊富さ
AD: Y	A24: 開発者の経験と知識の豊富さ
AD: Y	A25: 開発者の経験と知識の豊富さ
AD: Y	A26: 開発者の経験と知識の豊富さ
AD: Y	A27: 開発者の経験と知識の豊富さ
AD: Y	A28: 開発者の経験と知識の豊富さ
AD: Y	A29: 開発者の経験と知識の豊富さ
AD: Y	A30: 開発者の経験と知識の豊富さ
AD: Y	A31: 開発者の経験と知識の豊富さ
AD: Y	A32: 開発者の経験と知識の豊富さ
AD: Y	A33: 開発者の経験と知識の豊富さ
AD: Y	A34: 開発者の経験と知識の豊富さ
AD: Y	A35: 開発者の経験と知識の豊富さ
AD: Y	A36: 開発者の経験と知識の豊富さ
AD: Y	A37: 開発者の経験と知識の豊富さ
AD: Y	A38: 開発者の経験と知識の豊富さ
AD: Y	A39: 開発者の経験と知識の豊富さ
AD: Y	A40: 開発者の経験と知識の豊富さ
AD: Y	A41: 開発者の経験と知識の豊富さ
AD: Y	A42: 開発者の経験と知識の豊富さ
AD: Y	A43: 開発者の経験と知識の豊富さ
AD: Y	A44: 開発者の経験と知識の豊富さ
AD: Y	A45: 開発者の経験と知識の豊富さ
AD: Y	A46: 開発者の経験と知識の豊富さ
AD: Y	A47: 開発者の経験と知識の豊富さ
AD: Y	A48: 開発者の経験と知識の豊富さ
AD: Y	A49: 開発者の経験と知識の豊富さ
AD: Y	A50: 開発者の経験と知識の豊富さ
AD: Y	A51: 開発者の経験と知識の豊富さ
AD: Y	A52: 開発者の経験と知識の豊富さ
AD: Y	A53: 開発者の経験と知識の豊富さ
AD: Y	A54: 開発者の経験と知識の豊富さ
AD: Y	A55: 開発者の経験と知識の豊富さ
AD: Y	A56: 開発者の経験と知識の豊富さ
AD: Y	A57: 開発者の経験と知識の豊富さ
AD: Y	A58: 開発者の経験と知識の豊富さ
AD: Y	A59: 開発者の経験と知識の豊富さ
AD: Y	A60: 開発者の経験と知識の豊富さ
AD: Y	A61: 開発者の経験と知識の豊富さ
AD: Y	A62: 開発者の経験と知識の豊富さ
AD: Y	A63: 開発者の経験と知識の豊富さ
AD: Y	A64: 開発者の経験と知識の豊富さ
AD: Y	A65: 開発者の経験と知識の豊富さ
AD: Y	A66: 開発者の経験と知識の豊富さ
AD: Y	A67: 開発者の経験と知識の豊富さ
AD: Y	A68: 開発者の経験と知識の豊富さ
AD: Y	A69: 開発者の経験と知識の豊富さ
AD: Y	A70: 開発者の経験と知識の豊富さ
AD: Y	A71: 開発者の経験と知識の豊富さ
AD: Y	A72: 開発者の経験と知識の豊富さ
AD: Y	A73: 開発者の経験と知識の豊富さ
AD: Y	A74: 開発者の経験と知識の豊富さ
AD: Y	A75: 開発者の経験と知識の豊富さ
AD: Y	A76: 開発者の経験と知識の豊富さ
AD: Y	A77: 開発者の経験と知識の豊富さ
AD: Y	A78: 開発者の経験と知識の豊富さ
AD: Y	A79: 開発者の経験と知識の豊富さ
AD: Y	A80: 開発者の経験と知識の豊富さ
AD: Y	A81: 開発者の経験と知識の豊富さ
AD: Y	A82: 開発者の経験と知識の豊富さ
AD: Y	A83: 開発者の経験と知識の豊富さ
AD: Y	A84: 開発者の経験と知識の豊富さ
AD: Y	A85: 開発者の経験と知識の豊富さ
AD: Y	A86: 開発者の経験と知識の豊富さ
AD: Y	A87: 開発者の経験と知識の豊富さ
AD: Y	A88: 開発者の経験と知識の豊富さ
AD: Y	A89: 開発者の経験と知識の豊富さ
AD: Y	A90: 開発者の経験と知識の豊富さ
AD: Y	A91: 開発者の経験と知識の豊富さ
AD: Y	A92: 開発者の経験と知識の豊富さ
AD: Y	A93: 開発者の経験と知識の豊富さ
AD: Y	A94: 開発者の経験と知識の豊富さ
AD: Y	A95: 開発者の経験と知識の豊富さ
AD: Y	A96: 開発者の経験と知識の豊富さ
AD: Y	A97: 開発者の経験と知識の豊富さ
AD: Y	A98: 開発者の経験と知識の豊富さ
AD: Y	A99: 開発者の経験と知識の豊富さ
AD: Y	A100: 開発者の経験と知識の豊富さ

- Strategic IV&V planning (JAXA, NASA, ESA)
 - To build a practical and valuable tool to aid in the planning and rationalization of IV&V activities
 - Make best use of limited resources
 - What techniques to use, what priority, to what degree, how much is enough IV&V, what is the benefit of IV&V

Attribute	Technique	System characteristics			IV&V Constraint			cumulative risk reduction	cumulative cost					
		cf: Autonomy	cf: Fault tolerance	cf: Controller functionality	cf: Target Documentation	cf: Development phase to conduct	cf: Amount of development time							
A3: Requirement Precision	T3: SpecTRM Modeling	Autonomous 1	1FT 1	Central controller 1	T1	None	None	None	438373	0				
		Non-	1FT 1	End effect controller 1	T1	A6	T7	50000	0.02	49999.98	250000	388373	0.02	
		Autonomous 1	2FT 1	End effect controller 1	T2	A13	T21	25000	0.18	24999.84	156250	363373	0.18	
		Non-	2FT 1	Distributed controller 1	T3	A14	T22	25000	0.18	24999.84	156250	338373	0.34	
		Autonomous 1	3FT 1	Distributed controller 1	T4	A21	T31	12500	0.14	12499.86	86250.71	328873	0.48	
		Non-	3FT 1	Distributed controller 1	T5	A19	T27	75000	2.4	74997.6	31250	251873	2.88	
	T4: Natural language modeling	Autonomous 1	1FT 1	Central controller 1	T6	A3	T4	10000	4	9996	2900	151873	6.88	
			Non-	1FT 1	Central controller 1	T7	A12	T19	50000	2.4	4997.6	20833.33	101873	9.28
			Autonomous 1	2FT 1	Central controller 1	T8	A20	T30	50000	2.4	4997.6	20833.33	51873	11.68
			Non-	2FT 1	Distributed controller 1	T9	A7	T13	3125	0.18	3124.84	1953.25	4873	11.84
			Autonomous 1	3FT 1	Distributed controller 1	T10	A5	T6	25000	4	24996	6250	23750	15.84
			Non-	3FT 1	Distributed controller 1	T11	A18	T26	18750	8	18742	2343.75	5000	23.84
			Autonomous 1	4FT 1	Distributed controller 1	T12	A8	T15	2500	3.8	2496.2	714.286	2500	27.34
Non-	4FT 1	Distributed controller 1	T13	A4	T5	2500	4	2496	625	1	31.34			

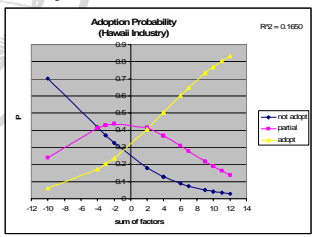


Current Strategic SWE projects (2)

- A "successfully completed" software development project *does not* guarantee that the resulting product will be adopted and used.
- Frequently factors outside the typical scope of the developers inhibit adoption of a new software product or system.

Our study: What are these factors and how do they affect adoption rate?

Application	Client Characteristics				Transition Preparation			Organizational Characteristics				Outcome		
	Focused	Representative	O & M Resources	Collaborative	Domain Knowledge	Software	Site	People	Business Case	Shared Vision	Win/Win SH/Infrs	Stable Envir.	Client Success	Adopted
org1														
org2	+	+	+	+	+	+	+	+	+	(+)	+	+	+	+
org3	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org4	+	+	+	+	+	(+)	+	+	+	+	+	+	+	+
org5	+	+	+	+	+	+	+	+	+	+	+	+	+	+
org6	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org7	+	+	+	+	+	+	+	+	+	+	+	+	+	+
org8	(+)	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org9	+	+	+	(+)	(+)	+	+	(+)	(+)	+	+	+	+	+
org10	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org11	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org12	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org13	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org14	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org15	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org16	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org17	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org18	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org19	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org20	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org21	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org22	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org23	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org24	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)
org25	+	+	+	(+)	(+)	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(+)

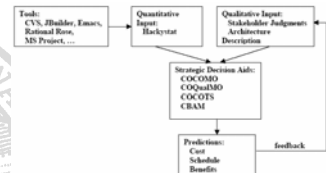


Current Strategic SWE projects (3)

Most strategic software engineering decision tools such as COCOMO II do not make use of automatically collected project data (software telemetry).

- The project has produced 100 KSLOC in 3.4 months, what is the expected remaining effort?
- The requirements have changed 12%, how does this change the schedule?

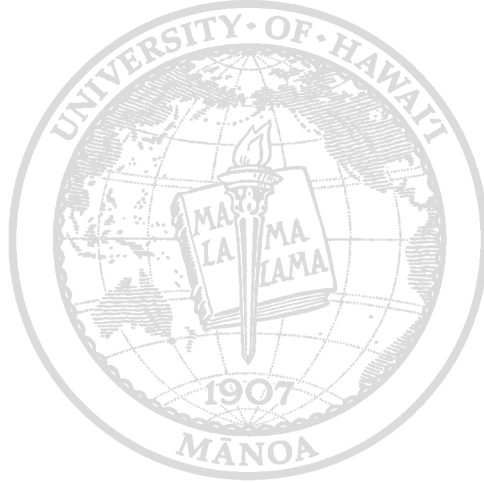
We are working on models to integrate software telemetry into strategic decision tools



How do empirical methods help with those problems and solutions?

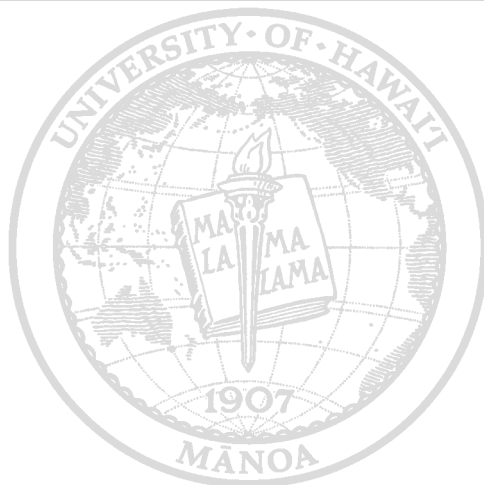
- Strategic methods depend critically on having good empirical data and sound empirical methods!
 - Scorecard models
 - Require qualitative input values and quantitative score estimates for scorecard
 - Calibration values for model from real-project data
 - Critical factors satisfaction and outcome data from real-projects for predictive model
 - Hackystat (obvious)
 - Strategic method depends on current risk, risk reduction, and cost data

How could EASE and EPM be used?



11

How can EASE or the Japanese software industry cooperate with your work?



12