

Valg av prosessmodeller, metoder, teknikker, praksiser, verktøy, språk etc.

- Organisasjoner, team og individer må stadig velge teknologi som støtter systemutvikling
- På hvilket grunnlag tar man slike valg?
 - Ofte basert på moter/"hype" og synsing fra "guruer"
 - Velger ofte det man kjenner best
- Ideelt:
 - Basert på empiriske undersøkelser (empirisk = erfaringsmessig, det som baserer seg på erfaring, dvs. observert i virkeligheten)

Er helikopter bedre enn sykkel?

Stor risiko ved vilkårlig innføring av teknologi

- Innføring av ny teknologi i private og offentlige organisasjoner kan være svært dyrt
- Trenger derfor skikkelig kost/nytte-evalueringer opp mot alternativer
- Men hvordan gjennomføre slike evalueringer? Når kan vi stole på resultatene?

Behov for kunnskap om empiriske metoder

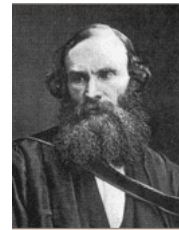
- Systemutvikling – menneskelig aktivitet i organisasjoner. Kan ikke bruke matematikk/logikk for å utlede hvordan utviklingen skal utføres
- Lytt til folk med erfaring, men den er
 - ofte begrenset til spesielle kontekster
 - ikke nødvendigvis gyldig i nye situasjoner
- Metode vektlegges mye i samfunnsfag og psykologi
- Nyttig kunnskap uansett hva man skal drive med
 - Masteroppgave
 - Rapporter

Sagt om måling

*To measure is to know.
If you cannot measure it, you cannot improve it.*
Lord Kelvin

*Not everything that counts can be measured.
Not everything that can be measured counts.*
Albert Einstein

In God we trust, all others bring data –
W. Edwards Deming



Lord Kelvin

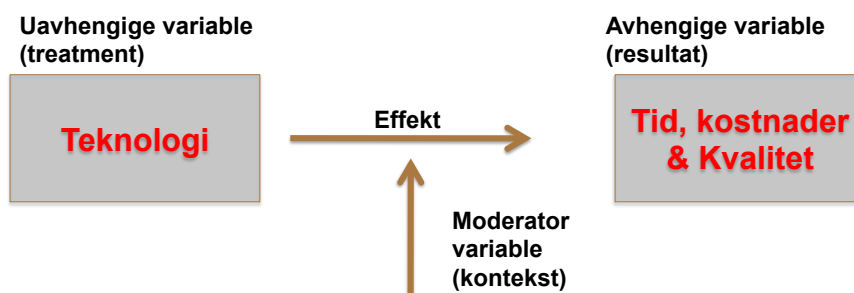
Empiriske forskningsmetoder

1. Eksperiment
2. Case-studie
 - Intervjuer
3. Etnografi
4. Spørreskjema-undersøkelse (survey)
5. Aksjonsforskning
6. Systematisk litteraturstudie
7. Meta-analyse

1. (Kontrollert) eksperiment

- Et eksperiment undersøker årsak-virkning – hva fører til hva?
- Direkte, presis og systematisk manipulering (endring) av det fenomenet man studerer, og så måler man virkningen
- Enkeltindivider eller team (deltakere) utfører systemutviklingsoppgaver der hensikten er å sammenligne ulike prosessmodeller, metoder, teknikker, språk eller verktøy (“treatments”)

Eksperiment



Statistisk styrke

- **Statistisk styrke:** sannsynligheten for at en statistisk test vil forkaste null-hypotesen hvis den faktisk er feil
- En test uten nok styrke vil ikke forkaste null-hypotesen selv om den er feil
- Må ofte gjøre et forsøk flere ganger eller ha mange deltakere i et eksperiment for å få nok styrke

Eksperiment om parprogrammering:

Nullhypotese

“Parprogrammering og solo-programmering gir like gode resultater mht tid, kostnader og kvalitet for ulike typer utviklere og oppgaver”

Effekten av å parprogrammering “kommer an på”

Programmerer- ekspertise	Oppgave- kompleksitet	Bruk PP?	Kommentarer
Junior	Enkel	Ja	Gitt at økt kvalitet er hovedmålet
	Kompleks	Ja	Gitt at økt kvalitet er hovedmålet
Mellomnivå	Enkel	Nei	
	Kompleks	Ja	Gitt at økt kvalitet er hovedmålet
Senior	Enkel	Nei	
	Kompleks	Nei*	

* Med mindre man vet at oppgaven er for vanskelig til å bli løst av selv en enkelt senior alene

Effekten vil avhenge av dyktigheten til utviklerne (som igjen avhenger av relevant intelligens, utdanning, arbeidserfaring, etc.) oppgave, system, utviklingsteknologi etc.

2. Case-studier

- Eksperimenter svarer på “hva”, case-studier mer på “hvordan” og “hvorfor”
- Case-studier legger vekt på å studere fenomener i sine naturlige omgivelser
- Mange flere variable enn datapunkter
- Teorier bør styre datainnsamling og analyse
- Generalisering ved bruk av teori, ikke statistikk som i eksperimenter

Typer av case-studier

- Enkel case-studie
- Multiple case-studier
 - Flere studier som studerer det samme fenomenet
- Komparativ case-studie
 - I multiple case-studier sammenlignes resultatene fra flere caser
 - I komparative case-studier sammenlignes data direkte på tvers av flere studier
- Longitudinal case-studie
 - En organisasjon eller prosjekt studeres over lengre tid og fokuserer på faktorer i omgivelsene som endrer seg og som påvirker fenomenet som studeres

Eksempel på komparative case-studie

Fire firmaer lagde samme system

Eks. på datakilder (fra studien av de fire prosjektene)

Data source	Description
Access logs	The logs from the web server were collected during the two years the systems were operational.
Bids	The companies' offered firm price was included in all the bids. The time schedule of the project, the planned development process, and the analysis and design of the product were included in many of the bids.
Customer time sheets	The customer team recorded the time they spent on the project.
CVS	At the completion of the development and testing, the researcher team received complete CVS bases from the projects.
E-mail	All the e-mail communication between the customer's project manager and the development projects was recorded.
Interviews	The projects' team members were interviewed weekly about their work on the project and about the possible effects of being the object of research. The interviews were semi-structured and based on an interview guide in which some questions were the same each week, while others varied depending on the status of their project.
Issue tracker	The companies registered their questions and needs for clarification in Bugzero; see http://www.websina.com/bugzero/ .
Log of defects	This is the log of the defects found after the systems became operational.
Project documents	These are documents related to overall project management, such as time schedules, design descriptions, acceptance test logs, and technical documentation.
Snapshots	Snapshots of all documents (including source code) were sent to the research team weekly.

INF1050/ 23.4.2013 / © Dag Sjøberg

Slide 29

Intervjuer – ofte brukt i case-studier

- **Strukturerte intervjuer**
 - Spørsmålene definert på forhånd, veldefinerte svaralternativer. Kan kvantifisere (telle opp) hvor mange som svarer hva på hvert spørsmål
- **Semistrukturerte intervjuer**
 - Intervjuerne baserer seg på stikkord og spørsmål som evt. kan droppes underveis, og nye spørsmål kan stilles avhengig av hvordan intervjuet forløper
- **Åpne (ustrukturerte) intervjuer**
 - Forløper seg mer som en samtale mellom intervjuer og intervjuobjekt

INF1050/ 23.4.2013 / © Dag Sjøberg

Slide 30

3. Etnografi/observasjon

- Forskeren observerer over lengre tid og hva folk gjør
- Dyp forståelse av hva folkene, organisasjonen og konteksten for arbeidet
- Personene trenger ikke å forklare hva de gjør
- Arbeidsoppgaver er ofte mer omfattende og komplekse enn hva system-modeller indikerer
- Sommerville nevner etnografi som mulig metode for å identifisere krav til et datasystem (avsnitt 4.5.5)

Etnografi versus case-studier

- Etnografi
 - forskeren mer involvert i gruppen som studeres
 - bruker gjerne “grounded theory”, dvs. lager teorier ut fra det som observeres
- Case-studier
 - bruker gjerne teorier for å bestemme hva slags data som samles inn (ikke så vanlig i systemutvikling)
 - teorien styrkes, forkastes eller justeres avhengig av det man finner
 - i eksplorative (utforskende) case-studier utvikler man også teorier ut fra dataene som samles inn

4. Spørreskjemaundersøkelser (Surveys)

- Vanlig i samfunnet – krever lite ressurser å nå ut til mange
- Lager statistikk og tester hypoteser over egenskaper ved gruppen som studeres (*målpopulasjon*)
- Man får svar på hva folk *mener* om hva, hvor mye, hvor mange, hvordan og hvorfor eller hva folk *sier* de gjør
 - I motsetning til eksperimenter kontrollerer man ikke de uavhengige og avhengige variable
 - I motsetning til case studier og etnografi observerer man ikke
- Vanligvis data fra et utvalg. Men hvis populasjonen er alle i et firma eller prosjekt, kan hele populasjonen undersøkes

Eksempler på temaer

- Identifisere problematiske områder i systemutvikling
 - Hyppighet av uklare kravspesifikasjoner
 - Gjennomsnittlige overskridelser i systemutviklingsprosjekter
- Identifisere faktorer som påvirker en egenskap eller en betingelse
 - Hvilke faktorer må være tilstede for vellykket innføring av UML i en organisasjon?

Når egner spørreskjemaundersøkelser seg?

- Har du definerte hypoteser?
- Er det klart hvilke variable du ønsker å måle?
- Er det klart hvordan disse variablene skal måles?
- Er det klart at man kan få pålitelige svar?
- Kan du få et utvalg som er representativt og stort nok (*response rate*) til å kunne si noe om målpopulasjonen?

Hvis "nei", er kanskje ikke en spørreskjemaundersøkelse egnet

Spørreskjemaer – type spørsmål

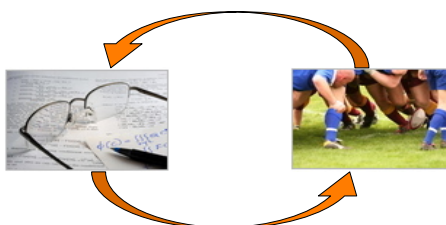
- Numeriske verdier, for eksempel alder
- Svarkategorier, for eksempel stillingstype
- Ja/Nei-svar
- Ordinalskala som vanligvis er bedre for holdninger og preferanser.
Tre typer
 - Enighetsskalaer, f.eks. 5-nivå Likert-skala med kategoriene: sterkt uenig, uenig, verken uenig eller enig, enig, sterkt enig
 - Frekvensskala, for eksempel aldri, sjelden, av og til, ofte, alltid
 - Evalueringsskalaer: svært dårlig, dårlig, passe, god, veldig god
- Åpne spørsmål

Fire typer skalaer

- Nominal
 - Verdiene tilhører kategorier uten rangering (rekkefølge), f.eks. farger, kjønn
- Ordinal
 - Som nominal skala, men kategoriene rangeres, f.eks. lite, middels, mye eller aldri, sjelden, av og til, ofte, alltid
- Intervall
 - Som ordinal, men avstanden mellom verdiene har betydning, f.eks. temperatur i Celsius
- Ratio (forholdstalls)
 - kontinuerlige skala med et absolutt nullpunkt. Forholdet mellom verdier kan beregnes, f.eks. temperatur i Kelvin

5. Aksjonsforskning

- Praksis og teori tett integrert slik at både praksis og forskning lærer av hverandre
- Forskere og ansatte i organisasjoner erfarer sammen problemer og utfordringer, og lærer hvordan man takler dem
- Sammen utfører de “aksjoner” (tiltak for endring) for å forbedre situasjonen



Aksjonsforskning prøver å oppnå
praktisk nytte
for organisasjonen samtidig med å fremskaffe
ny teoretisk kunnskap

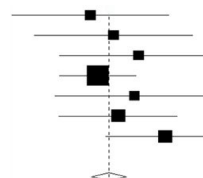
Sekundærstudier/sekundærforskning

Analyse av tidligere publiserte enkeltstudier for å frembringe mer generell kunnskap kalles

- Systematisk litteraturstudie
- Meta-analyse

6. Systematisk litteraturstudie

- Finner, evaluerer og setter sammen enkeltstudier innen et spesielt tema til en helhet (syntese)
- Bruker veldefinerte, godt dokumenterte teknikker slik at andre kan evaluere og replikere (gjenta) studien



INF1050/ 23.4.2013 / © Dag Sjøberg

Slide 41

Systematisk prosess

- Planlegg litteraturstudiet
 - Lag en protokoll for gjennomføringen
- Gjennomfør studien
 - Velg relevante (elektroniske) baser for søk etter primærstudier
 - Velg primærstudier
 - Vurder kvaliteten på disse
 - Samle inn dataene
 - Sammenlign og oppsummer dataene fra de ulike studiene
- Rapportert studien

INF1050/ 23.4.2013 / © Dag Sjøberg

Slide 42

Eksempel: Utfra eksisterende studier, hva er fordelene og ulempene ved smidige metoder?

- I hvilken grad er det evidens eller forskningsmessig belegg for påstander angående smidig systemutvikling?
- Hva sier eksisterende studier om konsekvensene for næringsliv, industri og forskning?



Available online at www.sciencedirect.com

ScienceDirect

Information and Software Technology 50 (2008) 833–859

INFORMATION
AND
SOFTWARE
TECHNOLOGY

www.elsevier.com/locate/infsof

Empirical studies of agile software development: A systematic review

Tore Dybå^a, Torgir Dingøy^a

^a SINTEF ICT, S.P. Andreassen, 15A, NO-7465 Trondheim, Norway

Received 22 October 2007; received in revised form 22 January 2008; accepted 24 January 2008
Available online 2 February 2008

Abstract

Agile software development represents a major departure from traditional, plan-based approaches to software engineering. A systematic review of empirical studies of agile software development up to and including 2005 was conducted. The search strategy identified 1996 studies, of which 36 were identified as empirical studies. The studies were grouped into four themes: introduction and adoption, human and social factors, perceptions on agile methods, and comparative studies. The review investigates what is currently known about the benefits and limitations of, and the strength of evidence for, agile methods. Implications for research and practice are presented. The main implication for research is a need for more and better empirical studies of agile software development within a common research agenda. For the industrial reader, the review provides a map of findings, according to topic, that can be compared for relevance to their own settings and situations.
© 2008 Elsevier B.V. All rights reserved.

Keywords: Empirical software engineering; Evidence-based software engineering; Systematic review; Research synthesis; Agile software development; XP; Extreme programming; Scrum

Contents

1. Introduction	834
2. Background – agile software development	834
2.1. The field of agile software development	834
2.2. Summary of previous reviews	836
2.3. Objectives of this review	837
3. Review method	837
3.1. Protocol development	837
3.2. Inclusion and exclusion criteria	838
3.3. Data sources and search strategy	838
3.4. Citation management, removal, and inclusion decision	839
3.5. Quality assessment	839
3.6. Data extraction	840
3.7. Synthesis of findings	840
4. Results	840
4.1. Overview of studies	840
4.2. Research methods	841
4.3. Methodological quality	841
4.4. Introduction and adoption of agile development methods	842

^a Corresponding author. Tel.: +47 73 59 29 43; fax: +47 73 59 29 77.
E-mail addresses: tore.dyba@sintef.no (T. Dybå), torgir.dingoy@sintef.no (T. Dingøy).

0950-5849/\$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.infsof.2008.01.006

7. Meta-analyse

- Hvis enkeltstudier er like mht uavhengige og avhengige variable (oftest eksperimenter), kan man bruke meta-analyse, dvs. statistiske metoder til å sammenligne og generalisere resultater fra flere studier
- I systemutvikling er ofte enkeltstudiene så ulike at det er vanskelig å kjøre statistikk over dem

Når bruke hvilken metode?

- Hvem ...
- Hva...
- Hvor...
- Hvordan ...
- Hvorfor ...



Type spørsmål, kompetanse, ressurser og mange andre forhold påvirker valg av metode

Ukeoppgave: Forskningsmetoder

- "Tenk deg at du er i et firma som utvikler programvare der utviklingssjefen har foreslått å innføre bruk av UML fordi han har lest at "forskning viser at det er nyttig." Samtidig innvender en av sjefsarkitektene at hun har hørt at "forskning viser at det er bortkastet å bruke tid på å lage UML-diagrammer".
- Hvordan kan det ha seg at forskning har kommet frem til tilsynelatende motstridende resultater?"