

IN2001: Software Engineering og prosjektarbeid
19. februar 2018

Forskningsmetoder / Evaluering av IT-systemer – med fokus på prosjektet



Professor Dag Sjøberg

Hvordan evaluerer man kvaliteten på et IT-system?

Anta at et firma vil utvikle en tjeneste for å finne filmer & serier

- Finnes fire versjoner av en prototype på en app med denne tjenesten (utviklet av studenter ved UiO)
- **Hvilken versjon bør firmaet velge?**

Hvordan evaluere kvaliteten på app'en?

Film & Serier

Finn den neste filmen eller TV-serien du vil se!

Appen skal ha støtte for å:

1. Liste populære filmer eller TV-serier
2. Liste sjangere
3. Søke i film eller TV-serie (søkeresultater skal vise tittel og poster / bilde)
4. Liste filmer eller TV-serier innen en sjanger (liste-elementer skal inneholde tittel og poster / bilde)
5. Vise detaljer av film eller TV-serie (detaljert skjerm viser stort bilde, tittel, beskrivelse (overview) og utgivelsesår)

Gruppeoppgave 1

Hvilke kriterier kan man vurdere tjenesten utfra?

- Tenk på det dere har lært om kravene til systemet
 - Funksjonelle krav
 - Ikke-funksjonelle krav
- Tenk over kriterier som kan variere mellom ulike interessenter (stakeholders)

Fra forrige forelesning:



Dimensjoner ved evaluering av IT-systemer: ISO 25010 har erstattet ISO 9126

- Functional **suitability***
- Reliability
- Usability
- **Performance** efficiency
- Maintainability
- Portability
- **Compatibility**
- **Security**

***Changes from ISO 9126 in red**

Funksjonelle krav

Hva systemet skal gjøre

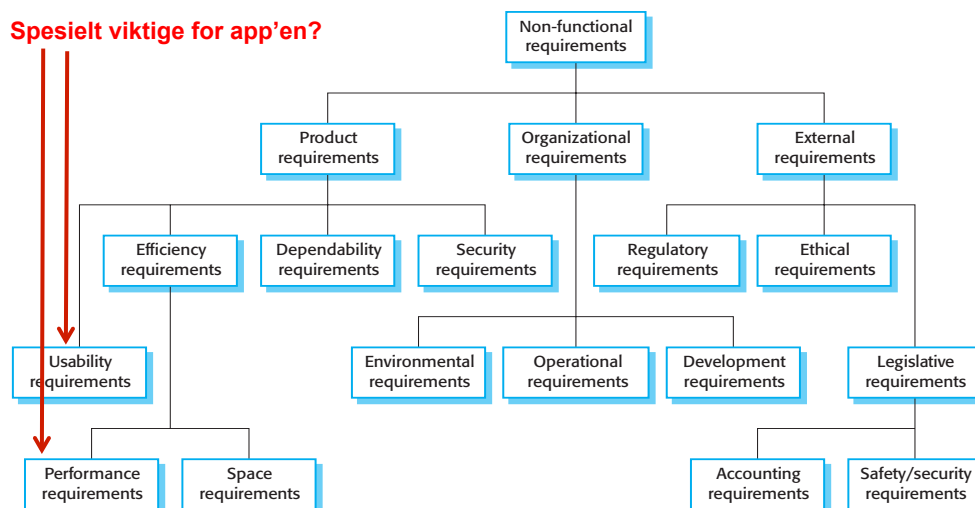
- Hvilke tjenester (funksjoner) skal systemet tilby?
- Hvordan skal det reagere på ulike typer input?
- For å avgrense systemet, vil man også kunne beskrive hva systemet *ikke* skal gjøre

Ikke-funksjonelle krav

- **Hvordan** systemet skal implementere de funksjonelle kravene

Typer av ikke-funksjonelle krav

Spesielt viktige for app'en?



Brukskvalitet (usability)

Tips fra PhD-student Raluca Florea:

- How to make a usability study:
<https://www.usability.gov/how-to-and-tools/methods/running-usability-tests.html>
- Guidelines on usability :
- Android: https://developer.android.com/guide/practices/ui_guidelines/index.html
- Advisable to do paper prototyping before development and test on users:
<https://www.youtube.com/watch?v=yafaGNFu8Eg>

Gruppearbeid

- Del inn i 4 grupper (samme som prosjektgruppene hvis mange nok tilstede i hver gruppe)
- Diskuter 15 minutter
- Oppsummering i plenum

Gruppe 1

Ikke funksjonelle:

- Nedetid på maks 15 minutter, tid mellom hver gang nede?
- God arkitektur for videreutvikling
- Mulig å drive versjonshåndtering
- Høy nok sikkerhet (hva betyr det konkret?)
- Lav feilrate (mål på dette?)
- Rask responstid (mål på dette?)
- Følger designplattformer
- Ikke vise offensive bilder (nakenhet/blasfemi osv) (følge rating filmprodusenter)

Gruppe 2

- Som bruker ønsker jeg å benytte appen på både mine Android og mine iOS devices. Jeg ønsker å ha samme opplevelse over alle plattformene.
- Som produkteier ønsker jeg at appen skal være enkel og billig å vedlikeholde i fremtiden (hvordan teste ut det? Hva betyr enkel og billig?)
- Som bruker ønsker jeg at appen skal være lett – altså ta lite plass både i minne og lagring
- Som produkteier ønsker jeg at appen skal se profesjonell ut slik at jeg kan fremme mine interesser (hva betyr profesjonell? Følge standarder som brukes i app'er som brukes i arbeidslivet/profesjonelt – seriøsitet)
- Som produkteier ønsker jeg at appen skal være lett å koble opp mot mine andre produkter (f.eks. felles innlogging, lenker til relaterte tjenester) – kompatibilitet
- Som bruker ønsker jeg å kunne benytte meg av tredjeparts verifisering gjennom f.eks google eller facebook.

Gruppe 3

Ikke-funksjonelle krav:

- Legal: ikke lekke søkehistorikk (konfidensialitet, personvern)
- Performance: 40 requests per 10. sekund
- Usability: easy to use. verdiløs dersom bruker ikke klarer å bruke appen (hvordan teste dette?)
- Development req: det skal være mulig å oppdatere og vedlikeholde appen (hva betyr "mulig"?)

Gruppe 4

- Videreutvikling (relativt enkelt å videreutvikle, kompatibilitet)
- Enkel kode som følger godtatte standarder
- God versjonskontroll og dokumentasjon
- Brukbarhet (usability testing)
- Fart på spørringer (ytelse/performance effisiency)
- Sikker (hva betyr det i praksis?) Dersom vi har brukerinformasjon så er dette ekstra viktig
- Etsiske hensyn, personvern, "Selge informasjon til tredjepart?"

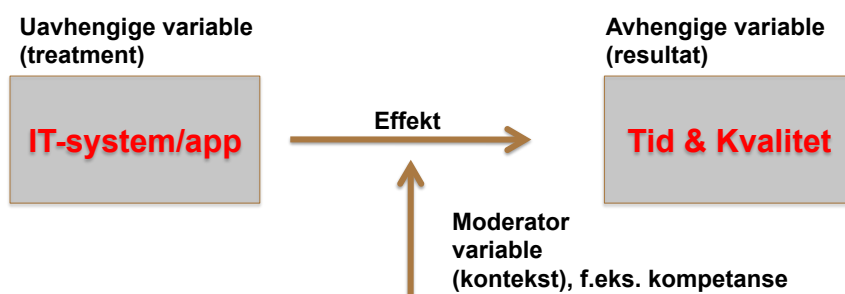
Når kriteriene for evaluering er klare, hvordan gjennomføre evalueringen?

- Hvilke (forsknings/undersøkelses)metoder egner seg?
 - Eksperiment
 - Case-studie
 - Etnografi (observasjon)
 - Spørreskjema-undersøkelse (survey)

(Kontrollert) eksperiment

- Et eksperiment undersøker årsak-virkning – hva fører til hva?
- Manipulerer (endrer) det fenomenet man studerer, og så måler man virkningen
- Enkeltindivider eller team (deltakere) utfører like oppgaver der hensikten er å sammenligne ulike “treatments”

Eksperiment



IN2001/ 19.2.2018 / © Dag Sjøberg

Slide 19

Case-studier

- Eksperimenter svarer på "hva" effekten er, case-studier mer på "hvordan" og "hvorfor"
- Case-studier legger vekt på å studere fenomener i sine naturlige omgivelser
- Case-studier har få datapunkter og mange variable. Derfor generalisering ved bruk av teori, ikke statistikk som i eksperimenter

IN2001/ 19.2.2018 / © Dag Sjøberg

Slide 20

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

IN2001/ 19.2.2018 / © Dag Sjøberg

Slide 21

Etnografi/observasjon



IN2001/ 19.2.2018 / © Dag Sjøberg

Slide 22

Metode

- Dyp forståelse av folk, organisasjon og konteksten for arbeidet
- Forskeren observerer over lengre tid hva folk gjør og er mer involvert i gruppen som studeres enn i case-studier
- Personene som studeres trenger ikke å forklare hva de gjør

Spørreskjemaundersøkelser (Surveys)

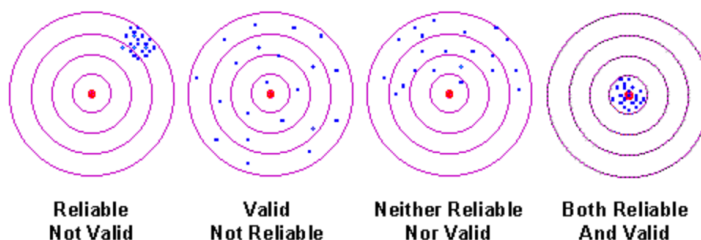


Spørreskjemaer – type spørsmål

- 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

Pålitelighet og gyldighet

- Se <https://www.socialresearchmethods.net/kb/relanval.php>



Gruppeoppgave 2

Hvilken metoder kan anvendes på hvilke kriterier?

(Innspill fra gruppene angitt i **rødt** på de neste slidene. Fikk bare 2,5 minutter til å registrere innspill fra hver gruppe i plenum, så ikke fullstendig oppsummering.)

Gruppe 1

Ikke funksjonelle:

- Nedetid på maks 15 minutter, tid mellom hver gang nede?
 - **Samle inn data over tid, evt. gjennomføre scenarier ved å pushe ny versjon eller oppdatering og teste om dette kan gjøres innen 15 minutter**
- God arkitektur for videreutvikling
 - **Ekstern ekspertvurdering / tredjepart (ulempe å bruke en av app'utviklerne fordi han/hun kjenner denne)**
- Mulig å drive versjonshåndtering
- Høy nok sikkerhet (hva betyr det konkret?)
- Lav feilrate (mål på dette?)
- Rask responstid (mål på dette?)
- Følger designplattformer
- Ikke vise offensive bilder (nakenhet/blasfemi osv) (følge rating filmprodusenter)

Gruppe 2

- Som bruker ønsker jeg å benytte appen på både mine Android og mine iOS devices. Jeg ønsker å ha samme opplevelse over alle plattformene.
 - Observasjon/etnografi evt. spørreundersøkelser for å se om app'en tilfredsstillende bruker
- Som produkteier ønsker jeg at appen skal være enkel og billig å vedlikeholde i fremtiden (hvordan teste ut det? Hva betyr enkel og billig?)
- Som bruker ønsker jeg at appen skal være lett – altså ta lite plass både i minne og lagring
 - La folk bruke den og observer, men også intervju brukerne (semistrukturert, men viktig å være åpen for at folk kan si det de har på hjertet).
- Som produkteier ønsker jeg at appen skal se profesjonell ut slik at jeg kan fremme mine interesser (hva betyr profesjonell? Følge standarder som brukes i app'er som brukes i arbeidslivet/profesjonelt – seriositet)
- Som produkteier ønsker jeg at appen skal være lett å koble opp mot mine andre produkter (f.eks. felles innlogging, lenker til relaterte tjenester) – kompatibilitet
- Som bruker ønsker jeg å kunne benytte meg av tredjeparts verifisering gjennom f.eks google eller facebook.

Gruppe 3

Ikke-funksjonelle krav:

- Legal: ikke lekke søkehistorikk (konfidensialitet, personvern)
 - La uavhengige sikkerhetsselskaper prøve å hente ut info som de ikke skal ha tilgang til
- Performance: 40 requests per 10. sekund
 - Stresstesting
- Usability: easy to use. verdiløs dersom bruker ikke klarer å bruke appen (hvordan teste dette?)
 - Observasjon
- Development req: det skal være mulig å oppdatere og vedlikeholde appen (hva betyr "mulig"?)
 - Statisk og dynamiske analyseverktøy, spørreundersøkelse blant utviklere (hva synes de om koden de har jobbet med)

Gruppe 4

- Videreutvikling (relativt enkelt å videreutvikle, kompatibilitet)
 - Eksperiment med et team som skal videreutvikle app'en (test en ny tjeneste på alle 4 appene)
 - Evt. ekstern ekspert som ser på koden til alle 4 app'ene
- Enkel kode som følger godtatte standarder
- God versjonskontroll og dokumentasjon
- Brukbarhet (usability testing)
- Fart på spørringer (ytelse/performance efficiency)
 - eksperimenter
- Sikker (hva betyr det i praksis?) Dersom vi har brukerinformasjon så er dette ekstra viktig
- Ethiske hensyn, personvern, "Selge informasjon til tredjepart?"