

TEKNIK EVALUASI

Tujuan

- **Mengevaluasi Pengujian Kegunaan dan Fungsi dari Sistem Interaktif**
- **Evaluasi dapat dilakukan pada :**
 - **Laboratorium**
 - **Lapangan Pekerjaan**
 - **Kerja Sama dengan User**
- **Beberapa Pendekatan Perancangan Evaluasi :**
 - **Metode Analitik**
 - **Metode Review**
 - **Metode Model Dasar**
- **Beberapa Pendekatan Evaluasi yang Telah diterapkan :**
 - **Metode Experimental**
 - **Metode Observasi**
 - **Metode Query**
- **Metode Evaluasi Yang Dipilih Secara Cermat Dan Harus Cocok Dengan Pekerjaan**

Evaluasi

- **Evaluasi digunakan untuk melihat apakah hasil rancangan dengan proses uji coba system yang telah dibuat sesuai dengan permintaan pengguna (user)**
- **Proses ini tidak dikerjakan dalam satu fase proses perancangan tetapi melalui perancangan dengan prinsip life cycle.**

Tujuan Evaluasi

Untuk melihat :

- **Seberapa jauh sistem berfungsi.**
- **Efek suatu interface ke pengguna**
- **Problem yang terjadi pada system**

Jenis- Jenis Evaluasi

- **Dalam kondisi percobaan (Laboratory)**
- **Dalam kondisi lokasi kerja sebenarnya**
- **Participatory Design**

- **Pada Kondisi Percobaan (Laboratory)**

Penggunaan pengujian system ini pada ruang percobaan mempunyai beberapa kondisi diantaranya

- Laboratorium yang bagus biasanya memiliki fasilitas perekaman audio/visual yang baik, komputer beserta perlengkapannya yang mungkin tidak ada pada lokasi kerja sebenarnya.
- Operator bebas dari gangguan yang menghambat pekerjaan.
- Sistem yang akan digunakan ditempatkan pada lokasi yang berbahaya atau lokasi yang terpencil, contoh stasiun ruang angkasa.
- Dapat memanipulasi situasi untuk memecahkan masalah dan melihat sedikit penggunaan prosedur atau membandingkan beberapa alternatif perancangan dengan situasi yang sebenarnya.
- Situasi pada laboratorium tidak dapat menggambarkan situasi ruang kerja sebenarnya dan terdapat beberapa orang yang tidak bisa bekerja pada kondisi di laboratorium

- **Pada Kondisi Lokasi Kerja Sebenarnya.**

Penggunaan pengujian system ini pada lokasi kerja sebenarnya mempunyai beberapa kondisi diantaranya :

- Tingkat gangguan yang melebihi ambang batas, tingkat-tingkat pergerakan yang besar dan interupsi yang tetap, seperti panggilan telepon menyebabkan observasi ini sulit dilakukan.
- Situasi yang lebih “terbuka” antara system dan pengguna, dimana kondisi ini tidak ditemukan pada kondisi di laboratorium
- Observasi pada lokasi kerja sebenarnya lebih baik dilakukan daripada di laboratorium dan gangguan-gangguan yang terjadi pada lokasi ini digunakan sebagai situasi yang mewakili situasi sebenarnya dan digunakan untuk proses penyimpanan dan pengambilan selama tugas.

- **Participatory Design**

- Adalah suatu pemikiran yang melibatkan keseluruhan alur perancangan dan tidak hanya proses evaluasi saja.
- Perancangan ini dilakukan pada ruang kerja yang melibatkan pengguna yang tidak hanya digunakan sebagai subyek percobaan tetapi juga sebagai anggota yang aktif dalam team perancangan.

- **Mempunyai tiga karakteristik :**

- Meningkatkan lingkungan kerja dan tugas
- Mempunyai sifat kerja sama, yakni pengguna dilibatkan dalam anggota team dan mempunyai kontribusi pada setiap tingkat perancangan.
- Mempunyai pendekatan iterative, perancangan adalah suatu subyek untuk evaluasi dan revisi pada setiap tingkatan.

- Beberapa metode untuk membantu mengkomunikasikan informasi antara pengguna dan perancang :
 - Brainstorming
 - Storyboarding
 - Workshops
 - Pensil dan kertas percobaan

Mengevaluasi Perancangan

- Cognitive Walkthrough
- Heuristic Evaluation

Cognitive Walkthrough

- Suatu usaha yang dilakukan untuk mengenalkan teori psikologi kedalam bentuk informal dan subyektif atau dengan kata lain mempunyai tujuan untuk mengevaluasi perancangan dengan melihat seberapa besar dukungan yang diberikan ke pengguna untuk mempelajari beberapa tugas yang diberikann. Pendekatan ini dikemukakan oleh Polson.
- Dalam pendekatan ini terdapat beberapa issue yang timbul seperti :
 - Pengaruh apa yang timbul setelah tugas ini diberikan ke pengguna?
 - Proses cognitive apa yang tersedia?
 - Masalah pembelajaran apa yang seharusnya timbul?
- Untuk melakukan Cognitive Walkthrough harus mempunyai informasi yang dibutuhkan :
 - Deskripsi dari suatu interface yang dibutuhkan itu sendiri
 - Deskripsi dari tugas termasuk usaha yang benar untuk melakukannya dan struktur tujuan untuk mendukungnya
- Dengan Informasi ini maka Evaluator dapat Melakukan langkah dari walkthrough :
 - Memilih Tugas
 - Mendiskrpsikan Tujuan dari user
 - Melakukan kegiatan yang tepat
 - Menaganalisa proses keputusan untuk setiap kegiatan
- Contoh Cognitive Walkthrough : Memprogram Video Dengan Remote Control
- Langkah selanjutnya adalah membandingkan struktur tujuan yang tersedia dengan struktur tujuan pada langkah awal. Hal ini dilakukan dengan empat pertanyaan:
 - Akankah user gagal untuk menambah tujuan yang diperlukan ?

- Akankah user gagal untuk memindahkan tujuan yang tidak diinginkan ?
- Akankah user mengadopsi beberapa dasar tujuan yang palsu dari interface ?
- Akankah user membatalkan tujuan yang masih dibutuhkan ?

Heuristic Evaluation

- Hampir sama dengan Cognitive Walkthrough tetapi sedikit terstruktur dan sedikit terarah. Dalam system ini terdapat beberapa criteria :
 - Perilaku Sistem dapat dipastikan.
 - Perilaku Sistem konsisten
 - Kemampuan memori user tidak melebihi batas
 - Dialog merupakan orientasi tugas
- Tujuan dari Heuristic Evaluation adalah untuk memperbaiki perancangan secara efektif. Orang yang melakukan evaluasi melalui kinerja dari serangkaian tugas dengan perancangan dan penilaiannya sesuai dengan kriteria setiap tingkatan. Jika ada kesalahan terdeteksi maka perancangan dapat ditinjau ulang untuk memperbaiki masalah ini sebelum pada tingkatan implementasi.

Review Based Evaluation

- Eksperimen antara psikologi dengan interaksi manusia dan komputer yang menghasilkan hasil-hasil eksperimen yang baik dan pengalaman yang nyata.
- Dalam kenyataannya hasil eksperimen ini tidak dapat dipastikan mempertahankan keadaan yang tetap. Orang yang melakukan evaluasi harus memilih data secara hati-hati, rancangan eksperimen yang dipilih, subyek masyarakat yang digunakan, analisa penyelenggaraan dan asumsi yang telah dibuat.

Model Based Evaluation

- Pendekatan terakhir untuk mengevaluasi perancangan dengan mengkombinasi spesifikasi perancangan dan evaluasi ke dalam kerangka kerja yang sama.
- Contoh GOMS model, keystroke level model dan design rationale.

Mengevaluasi model yang telah diterapkan

- Perbedaan yang besar dengan evaluasi perancangan adalah keberadaan implementasi system yang ada dalam berbagai bentuk. Hal ini dapat dimulai dari simulasi kemampuan interaktif dari suatu system, sebagai contoh Wizard of Oz, melalui fungsi prototype dasar sampai dengan system yang telah diimplementasi secara keseluruhan.

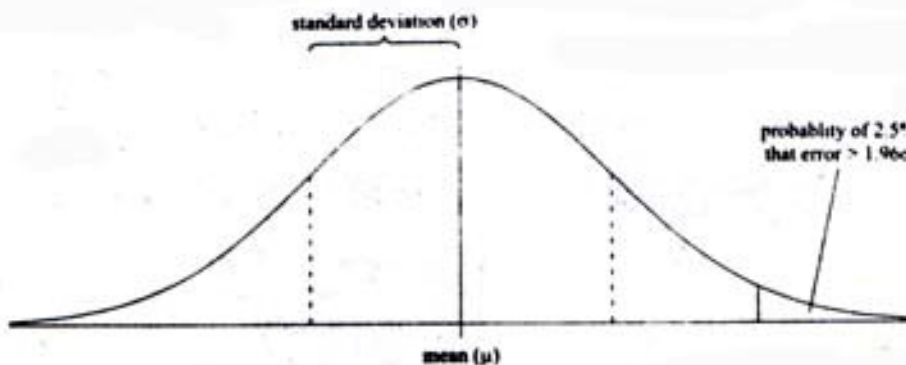
Metode Empirik : Evaluasi Eksperimen

- **Subyek**
 - Pemilihan subyek sangat penting dalam beberapa eksperimen.
 - Harus setepat mungkin dengan keinginan user.
 - Jika subyek bukan user sebenarnya maka subyek yang dipilih harus sama usia dan tingkatan pendidikan serta pengalaman menggunakan komputer secara umum dan keterhubungan system yang sedang dalam pengujian
- **Variabel**
 - **Independent Variabel**
Karakteristik suatu eksperimen yang memanipulasi untuk menghasilkan kondisi yang berbeda untuk perbandingan.
 - **Dependent Variabel**
Variabel yang dapat diukur dalam eksperimen.
- **Hypothesis**
 - Merupakan prediksi yang dihasilkan dalam eksperimen.
 - Masih menggunakan variable independent dan dependent, dimana variasi di dalam independent variable akan menyebabkan perbedaan pada dependent variable.
 - Tujuan dari eksperimen adalah untuk menunjukkan bahwa perkiraan benar dengan menyangkal adanya null hypothesis, yang menyatakan bahwa tidak ada perbedaan dalam dependent variable diantara tingkatan dari suatu independent variable.
- **Perancangan Eksperimen**
 - **Between-Groups (Randomized)**
 - Masing-masing subyek diberikan kondisi yang berbeda yakni kondisi eksperimen dan control.
 - Keuntungan perancangan ini adalah setiap user menghasilkan satu kondisi
 - Kerugiannya adalah dengan semakin banyak jumlah subyek yang tersedia akan menyebabkan hasilnya akan berkurang dan perbedaan antar setiap individu akan membuat bias hasil. Hal ini dapat diatasi dengan memilih dengan hati-hati subyek yang dipilih dan menjamin setiap kelompok di masyarakat terwakili.
 - **Within-Groups**
 - Setiap user akan menampilkan kondisi yang berbeda
 - Jumlah user yang tersedia lebih sedikit
 - Pengaruh dari subyek lebih sedikit
- **Pengukuran Statistik**
 - Dua aturan dalam analisa dengan statistik yakni : melihat dan menyimpan data
 - Variabel yang digunakan :
 - Discrete Variables

- Menggunakan jumlah yang terbatas dari suatu nilai atau tingkatan
Contoh: Warna layar monitor yakni : red, green atau blue.
- Merupakan suatu independent variable
- **Continuous Variable**
 - Menggunakan jumlah berapapun
Contoh : Ketinggian seseorang atau waktu yang diperlukan untuk menyelesaikan pekerjaan.
 - Merupakan suatu dependent variable

	Independent Variable	Dependent Variable	
Parametric			
	2 valued	Normal	Student's test on difference of means
	Discrete	Normal	ANOVA (Analysis Of Variance)
	Continuous	Normal	Linear (or non-linear) regression factor analys
Non-parametric			
	2 valued	Continuous	Wilcoxon (or Mann-Whitney) rank-sum test
	Discrete	Continuous	Rank-sum version of ANOVA
	Continuous	Continuous	Spearman's rank correlation
Contingency test			
	2 valued	Discrete	No special test, see next entry
	Discrete	Discrete	Contingency table and X squared test
	Continuous	Discrete	(rare) group independent variable and then as above

Contoh parametric test



Contoh non-parametric statistik

Kondisi A : 33,42,25,79,52
 Kondisi B : 87,65,92,93,91,55

Setelah diurut mulai 25,33,42,...,92,93 kemudian setelah di ranking maka 25 jadi 1, 33 jadi 2 dan seterusnya.

Kondisi A : 2,3,1,7,4
 Kondisi B : 8,6,10,11,9,5
 Dengan Wilcoxon test
 Kondisi A : 1+2+3+4+5 = 15
 Kondisi B : 1+2+3+4+5+6 = 21

	Rank sum	smallest	U
Kondisi A	: (2+3+1+7+4)	- 15	= 2
Kondisi B	: (8+6+10+11+9+5)-	21	= 28

Contoh mengevaluasi perancangan icon



Teknik Observasi

- **Think aloud**
 Menjelaskan : apa yang mereka percaya terjadi, mengapa mereka mengambil tindakan, apa yang mereka coba kerjakan.
 - Think aloud mempunyai keuntungan yakni sederhana, membutuhkan sedikit pengetahuan untuk menampilkannya dan menyediakan pengertian yang berguna dengan interface juga dapat digunakan untuk mengobservasi bagaimana system digunakan.
 - Variasi lain adalah cooperative evaluation dimana
 - Proses mempunyai hambatan yang sedikit sehingga lebih mudah dipelajari oleh orang yang melakukan evaluasi.
 - User mempunyai dorongan untuk menguji system
 - Orang yang melakukan evaluasi dapat menjelaskan kekacauan yang terjadi dan memaksimalkan pendekatan yang efektif untuk mengenali masalah

- **Analisa Protokol**
Terdapat beberapa metode untuk merekam kegiatan user :
 - Paper and pencil
 - Audio recording
 - Video recording
 - Computer logging
 - User notebook
- **Automatic protocol analysis tools**
 - EVA (Experimental Video Annotator)
System prototype yang berjalan pada multi-media workstation yang dihubungkan langsung ke video recorder
 - Workplace project pada Xerox PARC
- **Post-task walkthroughs**
 - Terdapat beberapa keadaan yang menyebabkan subyek tidak dapat berbicara selama observasi sebenarnya seperti subyek sedang mengerjakan tugas yang sangat penting dan banyak. Pada keadaan ini post-task walkthroughs dapat melihat secara subyektif perilaku user.

Teknik Query

- **Interview**
Menginterview user tentang pengalaman mereka dengan system interaktif yang menyediakan informasi secara langsung dan terstruktur
- **Questionnaire**
Metode alternatif yang agak kurang fleksibel dibandingkan dengan interview tetapi dapat meraih subyek yang banyak dan membutuhkan waktu yang tidak lama.
Terdapat beberapa jenis questionnaire :
 - General
 - Open-ended
 - Scalar
 - Multi-choice
 - Ranked

Memilih metode evaluasi

Faktor yang membedakan teknik evaluasi :

- **Tingkat siklus**
- **Jenis evaluasi**
- **Tingkat objektivitas dan subyektivitas**
- **Jenis ukuran yang tersedia**
- **Informasi yang tersedia**
- **Kesiapan dari suatu respon**
- **Tingkat gangguan yang tidak secara langsung**
- **Sumber yang tersedia**

Design VS. Implementation

Evaluasi pada tingkatan perancangan hanya membutuhkan ahlinya dan akan dianalisa sedangkan evaluasi pada tingkat implementasi membawa user sebagai subyek dari eksperimen.

Laboratory VS. Field studies

Secara ideal proses perancangan mencakup dua jenis evaluasi ini dimana pada laboratorium merupakan awal tingkatan sedangkan field studies ditempatkan pada tingkatan implementasi.

Subyektif VS. Obyektif

Pada teknik yang subyektif seperti pada cognitive walkthrough maupun think aloud mempercayakan pada evaluator atau orang yang melakukan evaluasi. Hal ini dapat menyebabkan penilaian evaluator menjadi bias sehingga diperlukan lebih dari satu orang evaluator. Sedangkan teknik obyektif seharusnya menghasilkan hasil yang berulang yang tidak tergantung pada evaluator. Eksperimen yang terkontrol merupakan contoh pengukuran obyektif. Secara ideal pengukuran secara subyektif dan obyektif seharusnya digunakan keduanya.

Qualitative VS Quantitative measures

Pengukuran menggunakan numeric lebih mudah digunakan menggunakan teknik statistik sedangkan non-numeric lebih sulit dilakukan akan tetapi menghasilkan sesuatu yang penting secara lengkap yang tidak dapat dilakukan secara numeric.

Informasi yang tersedia

Tingkatan informasi yang tersedia untuk evaluator pada setiap tingkat proses perancangan dimulai pada informasi yang rendah sampai dengan tingkatan yang tinggi.

Kesiapan suatu tindakan

Faktor lain yang membedakan teknik evaluasi adalah kesiapan mengambil tindakan yang tersedia. Seperti metode think aloud yang merekam perilaku user pada waktu interaksi. Sedangkan post task walkthrough mempercayakan pada pengumpulan kembali kegiatan user.

Intrusiveness

Merupakan teknik yang menghasilkan pengukuran segera yang dihasilkan user selama interaksi dan mempunyai resiko mempengaruhi perilaku user. Evaluator harus sensitive untuk mengurangi masalah ini tetapi tidak dapat megubah semuanya.

Sumber-sumber

Hal terakhir yang perlu diperhatikan dalam memilih teknik evaluasi yakni : peralatan, waktu, biaya, subyek serta evaluator yang ahli.

Klasifikasi teknik evaluasi

	Cognitive Walkthrough	Heuristic Evaluation	Review Based	Model Based
Stage	Throughput	Throughput	Design	Design
Style	Laboratory	Laboratory	Laboratory	Laboratory
Objective	No	No	As source	No
Measure	Qualitative	Qualitative	As source	Qualitative
Information	Low level	High level	As source	Low level
Immediacy	N/A	N/A	As source	N/A
Intrusive	No	No	No	No
Time	Medium	Low	Low-medium	Medium
Equipment	Low	Low	Low	Low
Expertise	High	Medium	Low	High

	Experiment	Interview	Questionnaire
Stage	Throughout	Throughout	Throughout
Style	Laboratory	Lab/Field	Lab/Field
Objective	Yes	No	No
Measure	Quantitative	Quantitative/Qualitative	Quantitative/Qualitative
Information	Low/High Level	High Level	High Level
Immediacy	Yes	No	No
Intrusive	Yes	No	No
Time	High	Low	Low
Equipment	Medium	Low	Low
Expertise	Medium	Low	Low

	Think-aloud	Protocol Analysis	Post-task Walkthrough
Stage	Implementation	Implementation	Implementation
Style	Lab/Field	Lab/Field	Lab/Field
Objective	No	No	No
Measure	Quantitative	Qualitative	Qualitative
Information	High/Low Level	High/Low Level	High/Low Level
Immediacy	Yes	Yes	No
Intrusive	Yes	Yes	No
Time	High	High	Medium
Equipment	Low	High	Low
Expertise	Medium	High	Medium