INTERAKSI MANUSIA DAN MESIN

Teknik Evaluasi

Budhi Irawan, S.Si, M.T

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 sistem*** 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 ***sistem***

JENIS - JENIS EVALUASI

1. Dalam kondisi *percobaan (Laboratory)*
2. Dalam kondisi *lokasi kerja sebenarnya*
3. *Participatory Design*

1. PADA KONDISI PERCOBAAN (LABORATORY)

Penggunaan pengujian sistem 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

2. PADA KONDISI LOKASI KERJA SEBENARNYA

Penggunaan pengujian sistem 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 sistem 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

3. 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.

KARAKTERISTIK EVALUASI

- *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.

METODE KOMUNIKASI USER - DESIGNER

- Brainstorming
- Storyboarding
- Workshops
- Pensil dan kertas percobaan

PROSES "EVALUASI PERANCANGAN"

a) Cognitive Walkthrough
b) Heuristic Evaluation
c) Review based evaluation
d) Model based evaluation

a) 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 diberikan***. 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?

a) COGNITIVE WALKTHROUGH (lanj.)

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

a) COGNITIVE WALKTHROUGH (lanj.)

- Dengan Informasi ini maka ***evaluator*** dapat melakukan langkah dari ***walkthrough*** :
  - Memilih Tugas
  - Mendeskripsikan Tujuan dari user
  - Melakukan kegiatan yang tepat
  - Menaganalisa proses keputusan untuk setiap kegiatan
- <u>Contoh Cognitive Walkthrough</u> : Memprogram Video Dengan Remote Control

a) COGNITIVE WALKTHROUGH (lanj.)

Langkah selanjutnya adalah membandingkan struktur tujuan yang tersedia dengan struktur tujuan pada langkah awal dan 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*** ?

b) 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 ***masuk pada tingkatan implementasi***.

c) 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 ekperimen yang dipilih, subyek masyarakat yang digunakan, analisa penyelenggaraan dan asumsi yang telah dibuat***.

d) 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.

PROSES "EVALUASI IMPLEMENTASI"

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

METODE EMPIRIK : EVALUASI EKSPERIMEN (lanj.)

- <u>Hypothesis</u>
  - 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 hyphotesis, yang menyatakan bahwa tidak ada perbedaan dalam dependent variable diantara tingkatan dari suatu independent variable.

METODE EMPIRIK : EVALUASI EKSPERIMEN (lanj.)

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

METODE EMPIRIK : EVALUASI EKSPERIMEN (lanj.)

Pengukuran Statistik

- Dua aturan dalam analisa dengan statistik yakni :
  melihat dan menyimpan data
- Variabel yang digunakan
  1. Discret Variables
     - Menggunakan jumlah yang terbatas dari suatu nilai atau tingkatan
       Contoh: Warna layar monitor yakni : red, green atau blue.
     - Merupakan suatu independent variable
  2. Continous Variable
     - Menggunakan jumlah berapapun
       Contoh : Ketinggian seseorang atau waktu yang diperlukan untuk menyelesaikan pekerjaan.
     - Merupakan suatu dependent variable

ANALISIS DATA

Data terdiri dari dua jenis, yaitu ***diskrit (jumlah terbatas dari nilai-nilai)*** dan ***kontinu (sembarang nilai)***.

Pada **sistem data** terdapat **tiga jenis tes**, yaitu :

1. ***Parametrik***

   Diasumsi ***berdistribudi normal*** data tetap dapat bekerja serta berdayaguna dalam berbagai kondisi

2. ***Non-parametrik***

   Distribusinya diasumsikan ***tidak berdistribusi normal***, kurang berdaya guna tetapi lebih handal.

3. ***Tabel kontingensi***

   klasifikasi data dengan atribut-atribut diskrit dan ***menghitung jumlah item data*** dalam setiap kelompok.

ANALISIS DATA (lanj.)

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

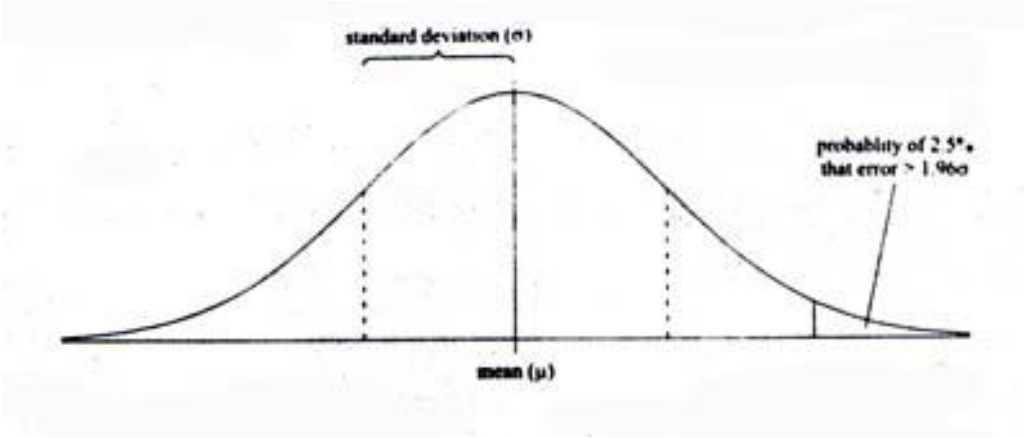
ANALISIS DATA (lanj.)

Tes parametrik dan non-parametrik ditunjukkan untuk hal-hal berikut :

- Informasi apa yang dibutuhkan untuk analisis data?
- Adakah perbedaan pada setiap data?
- Seberapa besar perbedaannya?
- Seberapa akurat taksirannya?

ANALISIS DATA (lanj.)

- Contoh parametric statistic



- Contoh non-parametric statistic
  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.

METODE EMPIRIK : EVALUASI EKSPERIMEN (lanj.)

- Contoh non-parametric statistic

  Kondisi A : 2,3,1,7,4 = 5 digit
  Kondisi B : 8,6,10,11,9,5 = 6 digit

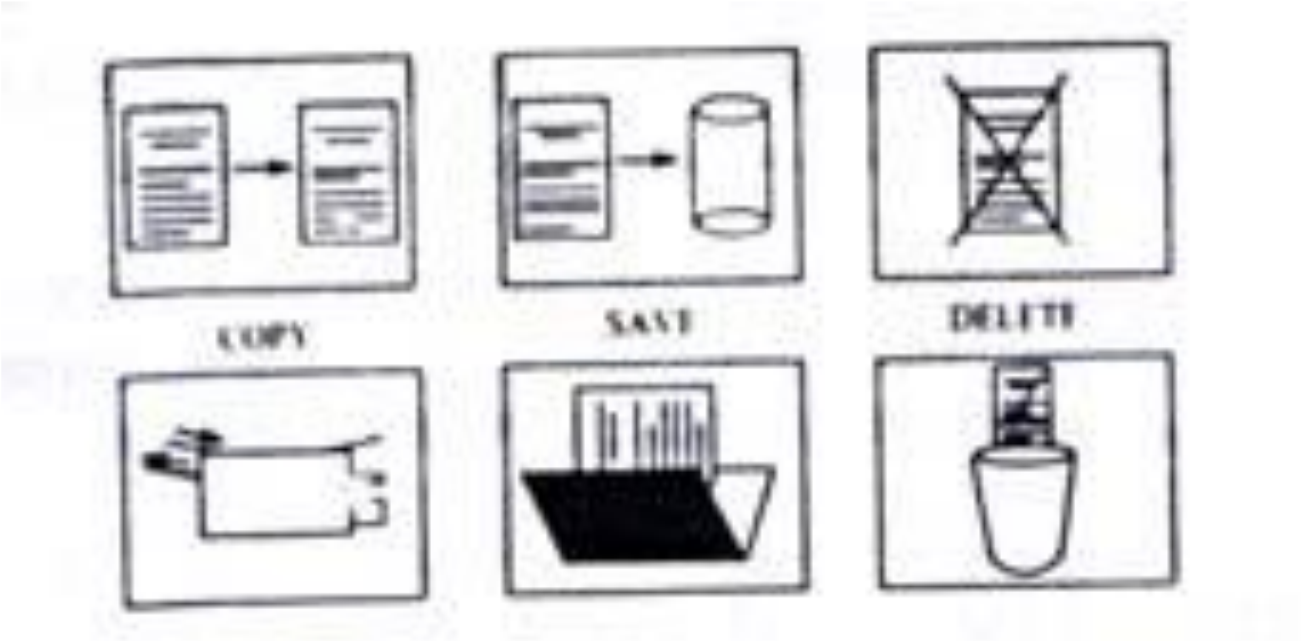
  Dengan tes Wilcoxon, setiap kondidi dijumlah rankingnya kemudian dikurangi dengan nilai terkecilnya sehingga didapatkan nilai statistik U-nya :
  Kondisi A : 1+2+3+4+5 = 15
  Kondisi B : 1+2+3+4+5+6 = 21

Rank sum	Smallest	U
Kondisi A : (2+3+1+7+4)	= 17 – 15	= 2
Kondisi B : (8+6+10+11+9+5)	= 49 – 21	= 28

  Ternyata jumlah dari kedua statistik U tersebut, 2 + 28 = 30, merupakan perkalian dari jumlah nilai dari masing-masing kondisi 5 X 6.

MENGEVALUASI IMPLEMENTASI

Contoh mengevaluasi perancangan icon



STUDI EKSPERIMENTAL PADA GRUP

Bebarapa masalah yang timbul yaitu :

- Subject group : Jumlah subjek lebih banyak, waktu yang dibutuhkan lebih lama karena munculnya banyak variasi dan sulit untuk menjadwal anggota grup yang ada.
- Task : Harus saking menyemangati dan kerjasama.
- Data gathering : Data yang dimiliki oleh setiap kelompok berbeda masalah seperti sinkronisasi, volume tipis dan sebagainya.
- Analisis : Waktu melakukan analisis dari suatu percobaan terdapat banyak variasi dari setiap anggota group.

TEKNIK OBSERVASI

1. Think Aloud
2. Analisa Protokol
3. Automatic protocol analysis tools
4. Post-task walkthroughs

1. 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

2. ANALISA PROTOKOL

Terdapat beberapa metode untuk merekam kegiatan user :

- Paper and pencil
- Audio recording
- Video recording
- Computer logging
- User notebook

3. 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

4. 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

TEKNIK QUERY (lanj.)

Questionnaire

- **Umum** : Pertanyaan untuk mengetahui latar belakang user dan memasukkannya ke dalam subjek populasi.
- **Open-ended** : pertanyaan yang membutuhkan jawaban atau opini sendiri, berguna untuk mengumpulkan informasi subjektif secara umum tetapi sulit untuk dianalisis.
- **Skala** : user diminta memutuskan pertanyaan tertentu berdasarkan skala angka, biasanya berhubungan dengan setuju atau tidak setuju terhadap pertanyaan.
- **Pilihan ganda** : Responden diberikan pilihan respon secara eksplisit dan hanya boleh memilih satu atau sebanyak yang diperlukan.
- **Ranking** : Mengurutkan item dalam daftar dan berguna untuk mengindikasikan preferensi user.

Untuk mengurangi kesalahan dari responden, sebaiknya digunakan pertanyaan tertutup seperti pertanyaan scalar, ranking atau pilihan ganda.

MEMILIH METODE EVALUASI

Faktor yang membedakan teknik evaluasi :

1) Tingkat siklus
2) Jenis evaluasi
3) Tingkat objektivitas dan subyektivitas
4) Jenis ukuran yang tersedia
5) Informasi yang tersedia
6) Kesiapan dari suatu respon
7) Tingkat gangguan yang tidak secara langsung
8) Sumber yang tersedia

1). TINGKAT SIKLUS

- **Design VS. Implementation**

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

2). JENIS EVALUASI

- 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.

3). TINGKAT OBJEKTIVITAS DAN SUBYEKTIVITAS

- 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.

4). JENIS UKURAN YANG TERSEDIA

- Qualitative VS. Quantitative measures

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

5). INFORMASI YANG TERSEDIA

- **Informasi yang tersedia**

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

6). KESIAPAN DARI SUATU RESPON

- **Kesiapan suatu tindakan**

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

7). TINGKAT GANGGUAN YANG TIDAK SECARA LANGSUNG

- 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 mengubah semuanya.

8). SUMBER YANG TERSEDIA

- 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

KLASIFIKASI TEKNIK EVALUASI (lanj.)

	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

KLASIFIKASI TEKNIK EVALUASI (lanj.)

	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

TUGAS

- Buatlah sebuah daftar pertanyaan bagi responden yang menunjukkan tingkat kepuasan dari suatu alat / web / aplikasi yang sering digunakan?