Modul Perkuliahan Sesi 5

Interaksi Manusia Komputer

EVALUASI

Evaluasi merupakan tes atas tingkat penggunaan dan fungsionalitas sistem. Evaluasi dilakukan untuk memastikan kecocokan dengan permintaan pengguna/tujuan pengguna dan untuk melihat apakah hasil rancangan dengan proses uji coba system yang telah dibuat sesuai dengan user. Proses evaluasi ini dikerjakan dalam satu fase proses perancangan tetapi melalui perancangan dengan prinsip life cycle dengan hasil dikembalikan untuk memodifikasi perancangan. Tujuan adanya evaluasi yaitu :

1. Melihat seberapa jauh sistem berfungsi. Desain sistem memungkinkan user melakukan tugas dengan lebih mudah. Desain tidak hanya membuat fungsionalitas yang sesuai dengan yang ada dalam sistem, tetapi juga memberikan kemudahan pada user untuk dapat melakukan aksi. Evaluasi pada tahap ini meliputi pengukuran unjuk kerja dari user untuk melihat keefektifan sistem dalam mendukung tugas.

2. Melihat efek interface bagi pengguna. Kemudahan untuk mempelajari sistem, usability dan perilaku user.

3. Mengidentifikasi problem khusus yang terjadi pada sistem, misal: ketika menggunakan konteks yang diinginkan tapi hasilnya tidak sesuai atau terjadi kekacauan diantara user. Ini tentunya berhubungan dengan usability dan fungsionalitas dari sistem (bergantung dari sebab masalah).

Jenis evaluasi dibedakan menjadi 2 yaitu :

1. Summative evaluation. Evaluasi yang dilakukan untuk menguji kesuksesan suatu produk yang sudah selesai yaitu mencapai standar yang ditentukan sebelumnya (ISO).

2. Formative evaluation. Evaluasi yang dilakukan selama desain untuk memastikan produk sesuai dengan yang diinginkan pengguna.

Tempat untuk melakukan evaluasi dilakukan di laboratorium dan di lapangan/ lokasi kerja pengguna sistem. Waktu evaluasi dilakukan ketika muncul produk baru atau perbaikan dari produk sebelumnya.

**Mengevaluasi Perancangan**

Evaluasi ini terjadi setelah proses perancangan. Secara ideal, evaluasi pertama dilakukan sebelum implementasi dimulai. Jika perancangan dievaluasi, kesalahan dapat dihindari karena perancangan dapat diperbaiki sebelumnya.

**Cognitive Walkthrough**

Tujuan evaluasi ini untuk melihat seberapa besar dukungan yang diberikan pada pengguna untuk mempelajari beberapa tugas yang diberikan. Pendekatan ini dikemukakan oleh Polson, dkk. Walkthrough dilaksanakan oleh perancang atau seorang ahli dalam psikologi kognitif. Ahli bekerja melalui perancangan tugas tertentu, tahap demi tahap, mengidentifikasi masalah yang berpotensi pada kriteria psikologi. Kemudian dibandingkan ke proses dimana perancang software bekerja dengan koding pada kondisi yang berbeda.

Dalam pendekatan ini ada beberapa issue yang timbul yaitu :

1. Pengaruh apa yang timbul setelah tugas ini diberikan ke pengguna?

2. Proses cognitive apa yang tersedia?

3. Masalah pembelajaran apa yang seharusnya timbul?

Analisis difokuskan pada tujuan user dan pengetahuan. Untuk melakukan evaluasi cognitive walkthrough membutuhkan informasi sebagai berikut :

1. Deskripsi dari suatu interface yang dibutuhkan itu sendiri

2. Deskripsi dari tugas termasuk usaha yang benar untuk melakukannya dan struktur tujuan untuk mendukungnya.

Dengan informasi yang sudah didapatkan maka evaluator dapat melakukan langkah dari walkthrough yaitu :

1. Pilih tugas

2. Deskripsikan tujuan awal dari user

3. Lakukan kegiatan/aksi yang tepat

4. Analisa proses keputusan untuk setiap kegiatan

Cognitive walkthrough berbasis formulir yang disediakan untuk merujuk evaluator melalui sekumpulan pertanyaan yang berhubungan dengan tugas dan tujuan user.

**Heuristic Evaluation**

Dikemukan oleh Neilsen dan Molich, hampir sama dengan cognitive walkthrough tetapi sedikit terarah dan terstruktur. Pada evaluasi ini, sekumpulan kriteria usability atau heuristic diidentifikasi dan perancangan dilaksanakan. Dalam sistem ini terdapat beberapa kriteria yaitu :

1. Perilaku sistem dapat dipastikan

2. Perilaku sistem konsisten

3. Feedback tersedia

4. Kemampuan memori user tidak melebihi batas

5. Dialog merupakan orientasi tugas

Tujuan heuristic evaluation adalah untuk memperbaiki rancangan secara efektif. Evaluator melakukan evaluasi melalui kinerja dari serangkaian tugas dengan perancangan dan disesuaikan dengan kriteria setiap tingkatan. Jika kesalahan terdeteksi dapat ditinjau ulang untuk diperbaiki sebelum tahap implementasi. Pendekatan ini tidak seluruhnya subyektif dengan menggunakan kriteria khusus untuk merujuk evaluasi. Selain itu membutuhkan level pengetahuan tertentu untuk mengaplikasikan evaluasi ini. Misalnya untuk menentukan jika suatu perancangan melebihi memori user, perancang perlu mengetahui kemampuan memori manusia.

**Review Based Evaluation**

Evaluasi antara psikologi eksperimen dengan interaksi manusia dan komputer menghasilkan hasil-hasil eksperimen yang baik dan pengalaman yang nyata. Misalnya pada usability dari tipe menu yang berbeda, pemanggilan nama perintah dan pemilihan icon. Dalam kenyataannya, hasil eksperimen ini tidak dapat dipastikan mempertahankan keadaan yang tetap. Evaluator harus memilih data secara hati-hati, menunjuk rancangan eksperimen yang dipilih, subyek masyarakat yang digunakan, analisa pelaksanaan dan asumsi yang dibuat. Misalnya: pengujian eksperimen, usability dari jenis sistem ‘bantu’ umum yang menggunakan subyek baru tidak menyediakan evaluasi yang tepat dari sistem ‘bantu’ yang dirancang bagi user ahli.

 **Model Based Evaluation**

Pendekatan evaluasi ini berguna untuk mengevaluasi perancangan dengan kombinasi spesifikasi perancangan dan evaluasi ke dalam kerangka kerja yang sama. Contohnya: GOMS Model, Keystroke level model dan design rationale.

**Mengevaluasi Implementasi**

**Evaluasi Eksperimen**

Evaluasi eksperimen merupakan evaluasi dengan pendekatan statistik dengan memilih hipotesis yang akan dicoba dan jumlah kondisi eksperimen. Setiap eksperimen memiliki dasar yang sama. Evaluator memilih hipotesa untuk diuji yang dapat ditentukan dengan mengukur beberapa atribut dari perilaku subyek. Sejumlah kondisi eksperimen ditentukan yang berbeda dalam hal nilai dari variabel tetap yang terkontrol. Setiap perubahan pengukuran perilaku diatributkan ke kondisi yang berbeda.

Faktor-faktor yang perlu dipertimbangkan dalam mendukung evaluasi eksperimen yaitu :

• Subjek, siapa yang merepresentasikan.

o Eksperimen sampai sesuai dengan keinginan user

o User yang dipilih sesuai dengan kelompok user yang dimaksud baik usia maupun tingkat pendidikannya jika subjek bukan user sebenarnya.

• Variabel, sesuatu untuk dimodifikasi dan ukuran

o Variabel independen, eksperimen untuk menghasilkan kondisi yang berbeda seperti: jenis interface, level dari help, jumlah item menu dan perancangan ikon.

o Variabel dependen, variabel yang dapat diukur dalam eksperimen seperti: waktu untuk menyelesaikan tugas, jumlah kesalahan yang dibuat dan kualitas dari unjuk kerja user.

• Hipotesis, apa yang ingin ditunjukkan.

o Merupakan prediksi yang dihasilkan dalam eksperimen.

o Hipotesis ini menggunakan variabel independen dan dependen, dimana variasi dalam variabel independen akan menyebabkan perbedaan pada variabel dependen.

o Tujuan eksperimen adalah untuk menyangkal keberadaan null hypothesis yang menyatakan tidak ada perbedaan dalam variabel dependen diantara tingkatan dari variabel independen.

o Pengukuran statistika menunjukkan bahwa nilai hasil dapat dibandingkan dengan tingkat kepercayaan. Jika nilai tertentu ada pada tingkat kepercayaan, hipotesis null tidak ada.

• Desain eksperimental, bagaimana cara melakukannya. Dalam perancangan eksperimen terdiri dari :

1. 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.

2. 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 statistic yaitu : melihat dan menyimpan data.

• Variabel yang digunakan yaitu :

O Discrete Variables

 Menggunakan jumlah yang terbatas dari suau nilai atau tingkatan, contoh : warna layar monitor yaitu red. green, dan blue

 Merupakan suatu variabel independen

O Continous Variable

 Menggunakan jumlah berapapun, contoh : tinggi seseorang atau waktu yang diperlukan untuk menyelesaikan suatu pekerjaan

 Merupakan variabel dependen

Contoh mengevaluasi perancangan icon



Dimisalkan kita akan membuat paket pengolahan kata baru yang menggunakan icon untuk presentasinya. Ada 2 jenis perancangan icon akan digunakan yang paling mudah bagi user. Jenis pertama menggunakan gambar natural (berdasarkan metapora dari dokumen kertas), jenis yang lain menggunakan gambar abstrak seperti terlihat pada gambar di atas. Yang pertama harus dilakukan adalah membentuk hipotesis : tampilan yang seperti apa yang diinginkan? Pada kasus ini kebanyakan user akan mengambil gambar natural sebagai icon yang lebih mudah dalam pemanggilan, karena lebih familiar bagi user. Hipotesis ini mengidentifikasikan variabel independent untuk eksperimen, natural dan abstrak. Pada pertimbangan variabel dependent, pengukuran ”lebih mudah” dilihat dari jumlah kesalahan dalam memilih dan waktu yang digunakan untuk memilih sebuah icon. Kontrol ekperimen dilakukan dengan membuat interface yang identik dengan perancangan icon dan tugas

pemilihan dibuat yang dapat diulang untuk setiap kondisi. Kemudian metode eksperimen dilakukan dengan menggunakan kelompok user yang dimaksud. Mengingat pembelajaran terlibat dalam eksperimen ini, maka perancangan within-groups digunakan dimana setiap user menggunakan kedua set icon tersebut. Kemudian hasil eksperimen secara detail dianalisis.

**Teknik Observasi**

Fungsi perangkat lunak yang digunakan pada teknik observasi yaitu :

• untuk memonitor test

• membuat catatan kejadian penting yang diobservasi dari video proses testing

• membantu membuat umpan balik diagnostic mengenai masalah penggunaan sistem.

 

Jenis-jenis teknik observasi yaitu :

• Think Aloud. Cara yang popular dalam mengumpulkan informasi dengan melihat interaksi user. Evaluator melihat dan merekam aksi user (dengan berbagai teknik).Konsekuensinya user ditanyai untuk mengelaborasikan aksinya dengan ”thinking aloud”, yaitu 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 sistem digunakan. Variasi lain adalah cooperative evaluation dimana:

O Proses mempunyai hambatan yang sedikit sehingga lebih mudah dipelajari oleh orang yang melakukan evaluasi.

O User mempunyai dorongan untuk menguji sistem

O 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 :

O Paper and pencil

O Audio recording

O Video recording

O Computer logging

O User notebook

• Automatic protocol analysis tools. EVA (Experimental Video Annotator) Sistem prototipe yang berjalan pada multi-media workstation yang dihubungkan langsung ke video recorder. Evaluator dapat menemukan sekumpulan tombol yang mengindikasikan event yang berbeda. Tombol digunakan dalam sesi perekaman untuk mencatat video dengan notes. Selama sesi user bekerja pada workstation dan direkam menggunakan video, logging system dan audio. Evaluator menggunakan tombol untuk mencatat event yang penting dari user dan merekam notes tambahan pada editor teks. Setelah sesi evaluator dapat menanyakan dan menggunakannya untuk mencari informasi. Contoh : 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

• Wawancara/interview. Wawancara merupakan suatu teknik yang dilakukan dengan menanyai user satu demi satu. Wawancara user mengenai pengalaman mereka dengan sistem interaktif yang menyediakan informasi secara langsung dan terstruktur. Wawancara efektif pada evaluasi tingkat tinggi, khususnya dalam memperoleh informasi tentang preferensi user, impresi dan perilaku. Jika digunakan bersama observasi, berarti mencari klarifikasi sebuah kejadian. Keunggulan metode ini pertanyaan yang diberikan dapat divariasikan sesuai dengan konteksnya dan pandangan user dapat diperoleh. Agar lebih efektif wawancara perlu direncanakan dan dibuat pertanyaannya. Kekurangan metode ini sangat subjektif dan memakan waktu yang lama.

• Kuesioner. Kuesioner merupakan suatu teknik dengan memberikan sejumlah pertanyaan yang sudah pasti diberikan dan dalam bentuk tercetak. Pertama yang harus dilakukan evaluator adalah menentukan tujuan dari kuesioner : informasi apa yang akan diambil ? Keunggulan metode ini yaitu cepat, dapat mencapai sasaran untuk kemompok pengguna yang besar dan dapat dianalisis dengan teliti, sedangkan kekurangannya kurang fleksibel. Beberapa jenis kuesioner yaitu :

O General, pertanyaan untuk mengetahui latar belakang pengguna dan memasukkanya ke dalam subyek populasi.

O Open-ended, pertanyaan yang membutuhkan jawaban dan opini sendiri yang berguna untuk mengumpulkan informasi subyektif secara umum tetapi sulit untuk dianalisis.

O Scalar, user diminta untuk memutuskan pernyataan tertentu berdasarkan skala angka, biasanya berhubungan dengan setuju atau tidak setuju terhadap pernyataan yang diberikan. Contoh : tidak setuju 1 2 3 4 5 setuju

O Multi-choice, responden diberikan pilihan respon secara eksplisit dan hanya boleh memilih satu saja atau sebanyak yang diperlukan. Misal : Software apa yang pernah Anda gunakan ?



• Ranked, mengurutkan item dalam daftar dan berguna untuk mengindikasikan preferensi user. Misal : rangkinglah kegunaan dari metode perintah



Tipe-tipe pertanyaan ini sangat berguna untuk tujuan yang berbeda-beda. Untuk mengurangi kesalahan dari responden, sebaiknya menggunakan pertanyaan tertutup seperti scalar, ranked atau multi-choice.

Klasifikasi teknik evaluasi dapat dilihat pada tabel-tabel berikut :

Tabel Teknis Analisis

 

Tabel Teknik Query

 

Tabel Teknik Observasi

 

**Pendekatan Evaluasi**

Pendekatan evaluasi dibedakan menjadi 3 yaitu :

• Usability testing

o Memastikan konsistensi struktur navigasi, penggunaan istilah, dan bagaimana sistem merespon pengguna

o Mengukur kinerja pengguna untuk menyelesaikan tugas tertentu

o Metode:

 Mengamati pengguna menggunakan sistem dan merekam kesalahan, waktu

 Interview/kuesioner untuk mendapatkan pendapat pengguna

o Dilakukan di laboratorium

o Contoh aplikasi : wordprocessor, spreadsheet, database.

• Field study

o Memahami apa yang dilakukan pengguna dan bagaimana produk terlibat/berperan dalam kegiatan mereka

o Identifikasi kesempatan adanya teknologi baru

o Membangun kebutuhan desain

o Memfasilitasi pengenalan teknologi baru atau meluncurkan teknologi dalam konteks yang berbeda

o Evaluasi teknologi

o Metode : interview dan observasi

• Analytical evaluation

• Evaluasi tanpa melibatkan end-user

o Metode:

 Inspeksi → heuristic evaluation, menggunakan standar untuk identifikasi masalah usability

 Walkthrough, melibatkan ahli untuk menjalankan skenario pada prototype aplikasi, dan menggunakan simulasi penyelesaian masalah.

Perbedaan ketiga pendekatan evaluasi ditunjukkan pada tabel berikut:

 

Kombinasi dari beberapa pendekatan seperti pada diagram berikut :

 

**Memilih Metode Evaluasi**

Ada 8 faktor yang membedakan teknik evaluasi yang berbeda yang membantu kita dalam memilih teknik evaluasi yang sesuai yaitu :

• Tingkat Siklus

Desain VS Implementasi

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.

• Tingkat obyektivitas 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.

• Jenis ukuran yang tersedia

Qualitative VS Quantitative measures

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

• 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 dari suatu respon

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.

• 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 sensitif untuk mengurangi masalah ini tetapi tidak dapat megubah semuanya.

• Sumber yang tersedia

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