

University of Toronto Department of Computer Science

Lecture 4: Requirements Elicitation

Last Week:
 Initiating an RE process:
 Stakeholders and Boundaries
 Goals and Scenarios
 Feasibility and Risk

This Week:
 Elicitation Techniques
 Interviews, Questionnaires, etc.
 Cognitive approaches
 Contextual approaches
 Ethnography as an RE technique

Next Week:
 Modeling and Analysis (I)
 Intro to RE Modeling
 Modeling Organisations
 Modeling soft systems

© 2000-2004, Steve Easterbrook 1

University of Toronto Department of Computer Science

Elicitation Techniques

→ **Traditional techniques**

- ↳ Introspection
- ↳ Reading existing documents
- ↳ Analyzing hard data
- ↳ Interviews
 - > Open-ended
 - > Structured
- ↳ Surveys / Questionnaires
- ↳ Meetings

→ **Collaborative techniques**

- ↳ Group techniques
 - > Focus Groups
 - > Brainstorming
- ↳ JAD/RAD workshops
- ↳ Prototyping
- ↳ Participatory Design

→ **Cognitive techniques**

- ↳ Task analysis
- ↳ Protocol analysis
- ↳ Knowledge Acquisition Techniques
 - > Card Sorting
 - > Laddering
 - > Repertory Grids
 - > Proximity Scaling Techniques

→ **Contextual approaches**

- ↳ Ethnographic techniques
 - > Participant Observation
 - > Ethnomethodology
- ↳ Discourse Analysis
 - > Conversation Analysis
 - > Speech Act Analysis
- ↳ Sociotechnical Methods
 - > Soft Systems Analysis

© 2000-2004, Steve Easterbrook 2

University of Toronto Department of Computer Science

Background Reading

→ **Sources of information:**

- ↳ company reports, organization charts, policy manuals, job descriptions, reports, documentation of existing systems, etc.

→ **Advantages:**

- ↳ Helps the analyst to get an understanding of the organization before meeting the people who work there.
- ↳ Helps to prepare for other types of fact finding, e.g. by being aware of the business objectives of the organization.
- ↳ may tell you the detailed requirements for the current system.

→ **Disadvantages:**

- ↳ written documents often do not match up to reality.
- ↳ Can be long-winded with much irrelevant detail

→ **Appropriate for**

- ↳ projects where analyst is not familiar with the organization being investigated.

© 2000-2004, Steve Easterbrook 3

University of Toronto Department of Computer Science

"Hard Data" Collection

→ **Identify Collections of Hard Data**

- ↳ Facts and figures, financial information,...
- ↳ Reports used for decision making,...
- ↳ Survey results, marketing data,...

→ **Sampling**

- ↳ Sampling used to select representative set from a population
 - > Purposive Sampling - choose the parts you think are relevant without worrying about statistical issues
 - > Simple Random Sampling - choose every kth element
 - > Stratified Random Sampling - identify strata and sample each
 - > Clustered Random Sampling - choose a representative subpopulation and sample it
- ↳ Sample Size is important
 - > balance between cost of data collection/analysis and required significance
- ↳ Process:
 - > Decide what data should be collected - e.g. banking transactions
 - > Determine the population to be sampled - e.g. all transactions at 5 local branches over one week
 - > Choose type of sample - e.g. simple random sampling
 - > Choose sample size - e.g. every 10th transaction

© 2000-2004, Steve Easterbrook 4

University of Toronto

Example of hard data

→ Questions:

- ↳ What does this data tell you?
- ↳ What would you do with this data?

Agate
Campaign Summary

Date 23rd February 1999

Client Yellow Partridge
Park Road Workshops
Park Road
Jewellery Quarter
Birmingham B2 3DT
U.K.

Campaign Spring Collection 1999

Billing Currency GB £

Item	Curr	Amount	Rate	Billing amount
Advert preparation: photography, artwork, layout etc.	GB £	15,000.00	1	15,000.00
Placement French Vogue	FFr.	47 000,00	11.35	4,140.97
Placement UK Vogue	GB £	5,000.00	1	5,000.00
Placement US Vogue	US \$	15,000.00	2.47	6,072.87
Total				30,213.84

This is not a VAT Invoice. A detailed VAT Invoice will be provided separately.

280-282 Carstairs Street, Birmingham B1 6TG
Tel 0121 711 2284 Fax 0121 711 0245
Email agate@agatetel.co.uk

© 2000-2004, Steve Easterbrook

University of Toronto Department of Computer Science

Interviews

→ Types:

- ↳ Structured - agenda of fairly open questions
- ↳ Open-ended - no pre-set agenda

→ Advantages

- ↳ Rich collection of information
 - Good for uncovering opinions, feelings, goals, as well as hard facts
- ↳ Can probe in depth, & adapt followup questions to what the person tells you

→ Disadvantages

- ↳ Large amount of qualitative data can be hard to analyze
- ↳ Hard to compare different respondents
- ↳ Interviewing is a difficult skill to master

→ Watch for

- ↳ Unanswerable questions ("how do you tie your shoelaces?")
- ↳ Tacit knowledge (and post-hoc rationalizations)
- ↳ Removal from context
- ↳ Interviewer's attitude may cause bias (e.g. variable attentiveness)

© 2000-2004, Steve Easterbrook Source: Adapted from Goguen and Linde, 1993, p154.

University of Toronto Department of Computer Science

Interviewing Tips

→ Starting off...

- ↳ Begin the interview with an innocuous topic to set people at ease
 - e.g. the weather, the score in last night's hockey game
 - e.g. comment on an object on the person's desk: "My, ... what a beautiful photograph! Did you take that?"

→ Ask if you can record the interview

- ↳ but put tape recorder in front of person
- ↳ say that they can turn it off any time.

→ Ask easy questions first

- ↳ perhaps personal information
 - e.g. "How long have you worked in your present position?"

→ Follow up interesting leads

- ↳ E.g. watch for things people say that indicate that your plan of action may be wrong,
 - e.g., "Could we pursue what you just said a little further?"

→ Ask open-ended questions last

- ↳ e.g. "Is there anything else you would like to add?"

© 2000-2004, Steve Easterbrook

University of Toronto Department of Computer Science

Questionnaires

→ Advantages

- ↳ Can quickly collect info from large numbers of people
- ↳ Can be administered remotely
- ↳ Can collect attitudes, beliefs, characteristics

→ Disadvantages

- ↳ Simplistic (presupposed) categories provide very little context
 - No room for users to convey their real needs

→ Watch for:

- ↳ Bias in sample selection
- ↳ Bias in self-selecting respondents
- ↳ Small sample size (lack of statistical significance)
- ↳ Open ended questions (very hard to analyze!)
- ↳ Leading questions ("have you stopped beating your wife?")
- ↳ Appropriation ("What is this a picture of?")
- ↳ Ambiguous questions (I.e. not everyone is answering the same question)

Questionnaires MUST be prototyped and tested!

© 2000-2004, Steve Easterbrook Source: Adapted from Goguen and Linde, 1993, p154.

University of Toronto Department of Computer Science

Meetings

- **Used for summarization and feedback**
 - ↳ E.g. meet with stakeholders towards the end of each stage:
 - > to discuss the results of the information gathering stage
 - > to conclude on a set of requirements
 - > to agree on a design etc.
 - ↳ Use the meeting to confirm what has been learned, talk about findings
- **Meetings are an important managerial tool**
 - ↳ Used to move a system development project forward.
 - ↳ Need to determine objectives for the meeting:
 - > E.g. presentation, problem solving, conflict resolution, progress analysis, gathering and merging of facts, training, planning....
 - ↳ Plan the meeting carefully:
 - > Schedule the meeting and arrange for facilities
 - > Prepare an agenda and distribute it well in advance
 - > The meeting itself may be structured or unstructured depending on objective;
 - > Keep track of time and agenda during the meeting
 - > Follow up with a written summary to be distributed to meeting participants
 - > Special rules apply for formal presentations (and how to prepare them), project walkthroughs, brainstorming,...

© 2000-2004, Steve Easterbrook 9

University of Toronto Department of Computer Science

Group Elicitation Techniques

- **Types:**
 - ↳ Focus Groups
 - ↳ Brainstorming
- **Advantages**
 - ↳ More natural interaction between people than formal interview
 - ↳ Can gauge reaction to stimulus materials (e.g. mock-ups, storyboards, etc)
- **Disadvantages**
 - ↳ May create unnatural groups (uncomfortable for participants)
 - ↳ Danger of *Groupthink*
 - ↳ May only provide superficial responses to technical questions
 - ↳ Requires a highly trained facilitator
- **Watch for**
 - ↳ sample bias
 - ↳ dominance and submission

© 2000-2004, Steve Easterbrook 10

University of Toronto Department of Computer Science

Joint/Rapid Application Development

- **JAD & RAD Principles:**
 - ↳ **Group Dynamics**
 - > one-to-one or group interview formats replaced with workshops
 - ↳ **Visual Aids**
 - > Use lots of visualization media, ranging from wall charts to large monitors or graphical interfaces
 - ↳ **Organized, Rational Process**
 - > Using techniques such as brainstorming and top-down analysis to structure the elicitation and analysis process
 - ↳ **WYSIWYG Documentation Approach**
 - > each JAD session results in a document which is easy to understand and is created and agreed upon during the session
- **Notes:**
 - ↳ Choose workshop participants carefully
 - > they should be the best people possible representing various stakeholder groups
 - ↳ Workshop should last 3-5 days.
 - > Must turn a group of participants into a team - this takes 1-2 days.
 - > Session leader makes sure each step has been completed thoroughly.
 - > Session leader steps in when there are differences of opinion - "open issues".
 - > Meeting room should be well-equipped for presentations, recording etc

© 2000-2004, Steve Easterbrook 11

University of Toronto Department of Computer Science

Knowledge Elicitation Techniques in RE

- **Background**
 - ↳ Knowledge elicitation is concerned with discovering 'expert' knowledge
 - ↳ Grew out of Expert Systems work in the 80's
 - ↳ Originally focussed on deriving expert's "rules" for Rule-based Systems
 - ↳ More recently, focussed on "problem solving methods"
- **But KE is hard**
 - ↳ Separation of domain knowledge from performance knowledge
 - ↳ Modeling problems
 - > Brittleness
 - > Assumption of rationality
 - ↳ Representational Problem
 - > epistemological inadequacy
 - > expressiveness vs. acquirability
 - ↳ Expert Bias

Example Techniques

- ↳ Eliciting domain knowledge
 - > Card Sorting
 - > Laddering
 - > Proximity Scaling Techniques
- ↳ Eliciting performance knowledge
 - > Protocol Analysis
- ↳ Using Multiple Experts
 - > Delphi Technique
 - > Focus Groups
 - > Repertory Grids
- ↳ Automated Techniques
 - > Machine Learning

© 2000-2004, Steve Easterbrook 12

University of Toronto Department of Computer Science

The Knowledge Level

→ View knowledge modelling as:

- Observe behaviour of an agent as black box
 - It acts as if it has some knowledge about its environment which it uses rationally
 - It takes actions to achieve ascribed goals
- Construct two models:
 - Symbol Level** - descriptions for mechanising behaviour
 - Knowledge Level** - descriptions of the agent's knowledge of the world

→ Two-step rationality:

- Agent applies its knowledge in two stages:
 - First creates a task specific model from the KL model based on features of the task.
- Hence, we actually need 3 models:
 - Domain model** - a systematic way of talking about a domain, with a coherent ontology.
 - Task model** - models goals, what it means to achieve a goal, and how goals are related.
 - Problem-solving method** - a way of relating task and domain models to accomplish goals.

© 2000-2004, Steve Easterbrook 13

University of Toronto Department of Computer Science

Knowledge Elicitation Techniques

→ Protocol Analysis

- based on vocalising behaviour
 - Think aloud vs. retrospective protocols
- Advantages
 - Direct verbalisation of cognitive activities
 - Embedded in the work context
 - Good at revealing interaction problems with existing systems
- Disadvantages
 - Essentially based on introspection, hence unreliable
 - No social dimension

→ Proximity Scaling Techniques

- Given some domain objects, derive a set of dimensions for classifying them:
 - step 1: pairwise proximity assessment among domain elements
 - step 2: automated analysis to build multi-dimensional space to classify the objects
- Advantages
 - help to elicit mental models, where complex multivariate data is concerned
 - good for eliciting tacit knowledge
- Disadvantages
 - Requires an agreed on set of objects
 - Only models classification knowledge (no performance knowledge)

© 2000-2004, Steve Easterbrook Source: Adapted from Hudlicka, 1996. 14

University of Toronto Department of Computer Science

more KE techniques

→ Card Sorting

- For a given set of domain objects, written on cards:
 - Expert sorts the cards into groups...
 - ...then says what the criterion was for sorting, and what the groups were.
- Advantages
 - simple, amenable to automation
 - elicits classification knowledge
- Problems
 - suitable entities need to be identified with suitable semantic spread across domain.
 - No performance knowledge

→ Laddering

- Uses a set of probes (types of question) to acquire structure and content of stakeholders' knowledge.
 - Interview the expert.
 - Use questions to move up and down a conceptual hierarchy
- Advantages
 - deals with hierarchical knowledge, including poly-hierarchies (e.g., goal trees, "is-a" taxonomies).
 - knowledge is represented in standardised format
 - can elicit structural knowledge
 - suitable for automation.
- Disadvantages
 - assumes hierarchically arranged knowledge.

© 2000-2004, Steve Easterbrook 15

University of Toronto Department of Computer Science

KA from Multiple Experts

→ Delphi technique

- Used where contact between experts is difficult:
 - Each expert submits their judgement
 - All judgements are circulated anonymously to all experts
 - Each expert then submits a revised judgement
 - Iterate until judgements converge

→ Focus Groups

- A technique derived from marketing:
 - Assemble experts together and discuss the problem
 - Discussion may be structured (e.g. debate) or unstructured

→ Repertory Grids (based on Kelly's Personal Construct Theory)

- Used to detect terminological differences
 - Get the experts to agree a set of entities
 - Each expert provides attributes and values
 - For each attribute in expert A's grid, find the closest match in expert B's grid. (i.e. are there attributes which have the same discriminatory function?)
 - Experts then rate the entities using each other's attributes

© 2000-2004, Steve Easterbrook 16

University of Toronto Department of Computer Science

Abstractionism vs. Contextualism

→ **Abstractionism**

- ↳ Builds models abstracted from a domain; the model is used to answer questions
 - (1) Decide on the ontology of the phenomena we wish to describe
 - (2) Use this ontology to represent the domain of discourse
- ↳ Assumes knowledge and understanding are independent from context
- ↳ Used by natural scientists and engineers.
 - > ...although many scientists don't realize that step 1 involves choice
 - > logical positivism vs. theory-driven observation

→ **Contextualism**

- ↳ Emphasizes the details and idiosyncrasies of the domain
 - (1) Collect naturalistic data from the domain of study (Rich descriptions)
 - (2) Use the data to support explanations (but don't build abstract models)
- ↳ Assumes it is impossible to build models that have meaning when removed from their context
- ↳ Used by many social scientists
 - > but generally limits them to the descriptive rather than predictive/prescriptive

© 2000-2004, Steve Easterbrook 17

University of Toronto Department of Computer Science

Participant Observation

→ **Approach**

- ↳ longitudinal studies:
 - > Observer spends time with the subjects, joining in, long enough to become a member of the group

→ **Advantages**

- ↳ Contextualized;
- ↳ Reveals details that other methods cannot

→ **Disadvantages**

- ↳ Extremely time consuming!
- ↳ Resulting 'rich picture' is hard to analyze
- ↳ Cannot say much about the results of proposed changes

→ **Watch for**

- ↳ going native!

© 2000-2004, Steve Easterbrook 18

University of Toronto Department of Computer Science

Ethnomethodology

→ **Basis**

- ↳ Social world is ordered
 - > The social order may not be obvious, nor describable from common sense
- ↳ The social order cannot be assumed to have an a priori structure
 - > Social order is accomplished on a moment-to-moment basis through participants' collective actions (rather than through any pre-existing structures)
 - > i.e. social order only observable when an observer immerses herself in it.
- ↳ Observation should be done in a natural setting
- ↳ Need to consider how meanings develop and evolve within context

→ **"Use the members' own Categories"**

- ↳ Most conventional approaches assume preexisting categories
 - > This may mislead the observer (e.g. appropriation)
- ↳ Ethnomethodology attempts to use the subjects' own categories
 - > What categories (concepts) do they use themselves to order the social world?
- ↳ What methods do people use to make sense of the world around them?
 - > During observation, use the same methods members use, eg by developing a legitimate role within the community under observation.

→ **Measurement**

- ↳ No scientific objectivity, so use the subjects' own measurement theory

© 2000-2004, Steve Easterbrook Source: Adapted from Goguen and Linde, 1993, p158. 19

University of Toronto Department of Computer Science

Ethnomethodological approach

→ **Ethnomethodology is a subarea of Anthropology**

- ↳ Looks for behaviours that may be different in a specific culture but which have the same underlying purpose or meaning.
 - > E.g. how do people go about gaining status in different cultures:
 - > Frenchmen brag about sexual conquests to gain status;
 - > Americans brag about money to gain status.
 - > Each of these topics is taboo in the other culture

→ **Uses a very tightly controlled set of methods:**

- > Conversational analysis
- > Measurement of body system functions - e.g. heartbeat
- > Studies of Non-verbal behaviour (e.g. gestures, body language)
- > Detailed video analysis
- ↳ These techniques are useful in capturing information about a social setting.

→ **Other observation techniques can be applied:**

- ↳ Time-motion study
 - > who is where, when?
- ↳ Communication audit
 - > who talks to whom about what?
- ↳ Use of tools - status symbols plus sharing rules

© 2000-2004, Steve Easterbrook 20



Postscript: Postmodernism

→ Modernism

- ↳ **Rationality is the highest form of mental functioning**
- ↳ **Modern science produces universal truths**
 - > ...independent from the context and status of the scientist who produced them
- ↳ **Rationality will always lead to progress and perfection**
 - > All human institutions can be scientifically analyzed and improved
- ↳ **Reason is the ultimate judge of what is right (true, legal, ethical,...)**
- ↳ **Language must be rational**
 - > it only exists to represent the real world;
 - > there must be a firm, objective connection between the "signifier" and the "signified"
 - > the meaning cannot depend on the audience

→ Postmodernism (PoMo)

- ↳ **Questioning the grand narrative**
 - > A grand narrative is a story that a culture/society tells itself about its practices and beliefs
 - > E.g. in the US: "democracy is the most enlightened/rational form of government"
 - > E.g. in science: "scientific truths are universal and eternal"
 - > Postmodernism identifies and critiques such narratives
- ↳ **Instead, look for mini-narratives**
 - > Stories that explain small practices, local events, situated, contingent behaviour
 - > ...and don't make any claims about universality, truth, or stability
- ↳ **E.g. Literary Deconstruction**
 - > Examine what a text does not say, what it represses
 - > Reveal internal arbitrary hierarchies and dichotomies
- ↳ **E.g. Semiotics**
 - > The study of the relationship between signs and the things they signify



What has PoMo got to do with RE?

→ logical positivist view:

- > "there is an objective world that can be modeled by building a consistent body of knowledge grounded in empirical observation"
- ↳ **In RE: "there is an objective problem that exists in the world"**
 - > Build a consistent model; make sufficient empirical observations to check validity
 - > Use tools that test consistency and completeness of the model
 - > Use reviews, prototyping, etc to demonstrate the model is "valid"

→ Popper's modification to logical positivism:

- > "theories can't be proven correct, they can only be refuted by finding exceptions"
- ↳ **In RE: "requirements models must be refutable"**
 - > Look for evidence that the model is wrong
 - > E.g. collect scenarios and check the model supports them

→ post-modern view:

- > "there is no privileged viewpoint; all observation is value-laden; scientific investigation is culturally embedded"
- ↳ **In RE: "validation is always subjective and contextualised"**
 - > Use stakeholder involvement so that they 'own' the requirements models
 - > Use ethnographic techniques to understand the participant's worldviews