



Evaluating Web Site Quality

Gordan Gledec

Department of Telecommunications, FER, University of
Zagreb

CUC 2005, November 21-23, 2005



Outline

- introduction – why quality matters?
- description of the Web site quality model
- characteristics and subcharacteristics
- evaluation process
- verification and results
- conclusion

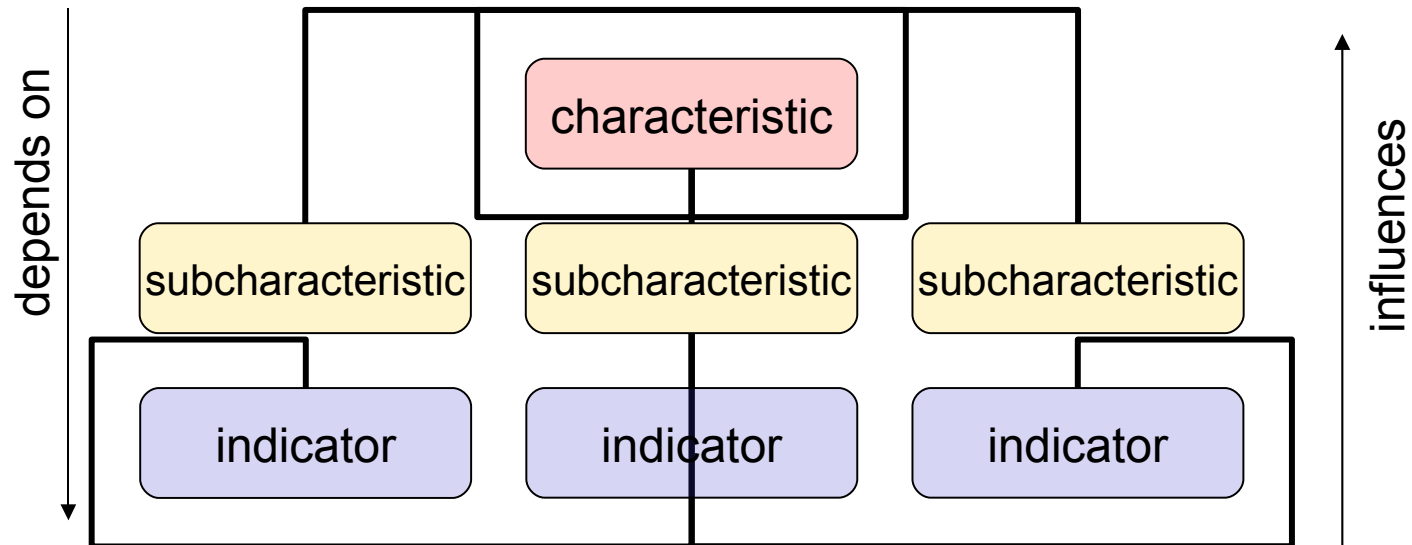


Introduction

- success of the Web site is based on its quality
 - quality perceived by the user – subjective
 - how to quantify quality?
- quality requirements → measurable attributes
→ metrics
- quality model simplifies the evaluation process
 - no need to reinvent the wheel – ISO standards
 - Web is equivalent to software product
- this work describes the Web site quality model

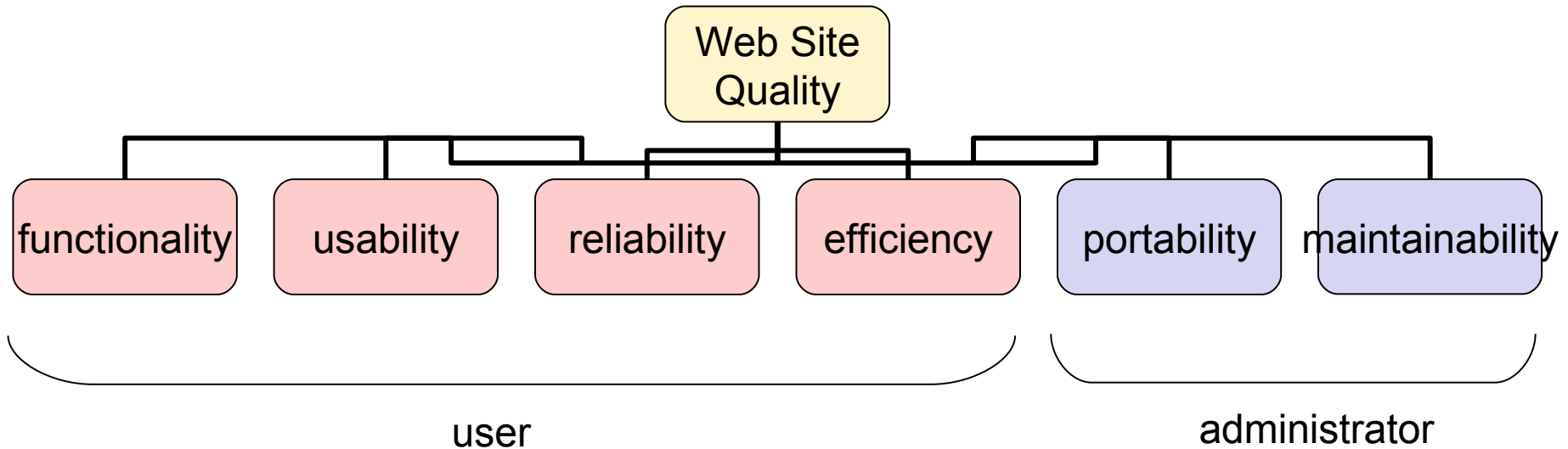
Web site Quality Model

- model based on ISO 9126 standard
 - three levels of hierarchy





Quality Model Characteristics





Functionality

- capability of the Web site to provide functions and properties which meet stated and implied needs when the site is used under specified conditions.
 - **suitability** - capability of the site to provide an appropriate set of functions for specified tasks and user objectives
 - searchability, navigability, relevance
 - **accuracy** - capability of the site to provide the right or agreed results or effects
 - credibility, freshness.



Functionality (cont.)

- **interoperability** - capability of the site to interact with one or more specified systems
 - compliance to standards and guidelines, format suitability.
- **confidentiality** - capability of the site to prevent accidental or deliberate unauthorised access and allow access to authorised persons or systems
 - protection suitability, access control.



Usability

- capability of the Web site to be understood, learned and liked by the user, when used under specified conditions.
 - **ease of use** - capability of the site to be used with ease while performing a specified task
 - ease of navigation, quality of links, quality of search and quality of window management
 - **comprehensibility** - capability of the site to present the content in a way that is suitable for performing a specified task
 - content clarity, content readability, presentation quality



Usability (cont.)

- **level of communication** - capability of the site to enable active or passive communication with users
 - quality of forms, quality of help system, feedback quality
- **attractiveness** - capability of the site to be attractive to the user
 - user satisfaction



Reliability

- capability of the site to maintain a specified level of performance when used under specified conditions.
 - **availability** - capability of the site to be in a state to perform a required function at a given point in time, under stated conditions of use
 - availability of site elements, support for browsers
 - **fault tolerance** - capability of the software product to maintain a specified level of performance in cases of faults
 - link reliability, informativeness of error messages, code quality
 - **security** - capability of the site to prevent unauthorised access which might lead to performance degradation
 - site vulnerability



Efficiency

- capability of the site to provide appropriate performance, relative to the amount of resources used, under stated conditions
 - **time behaviour** - capability of the site to provide appropriate response and processing times and throughput rates when performing its function, under stated conditions
 - page loading time, response time
 - **resource utilization** - capability of the site to use appropriate amounts and types of resources when the site performs its function under stated conditions
 - user's resources, site resources



Efficiency (cont.)

- **scalability** - capability of the site to maintain the expected level of performance in the event of increased server load
 - load management, adaptability of presentation.
- **visibility** - capability of the site to be available with respect to URL and domain recognizability
 - address recognizability, quality of META tags
- **flexibility** - capability of the site to ensure different methods of access to the content
 - adaptability of presentation



Maintainability

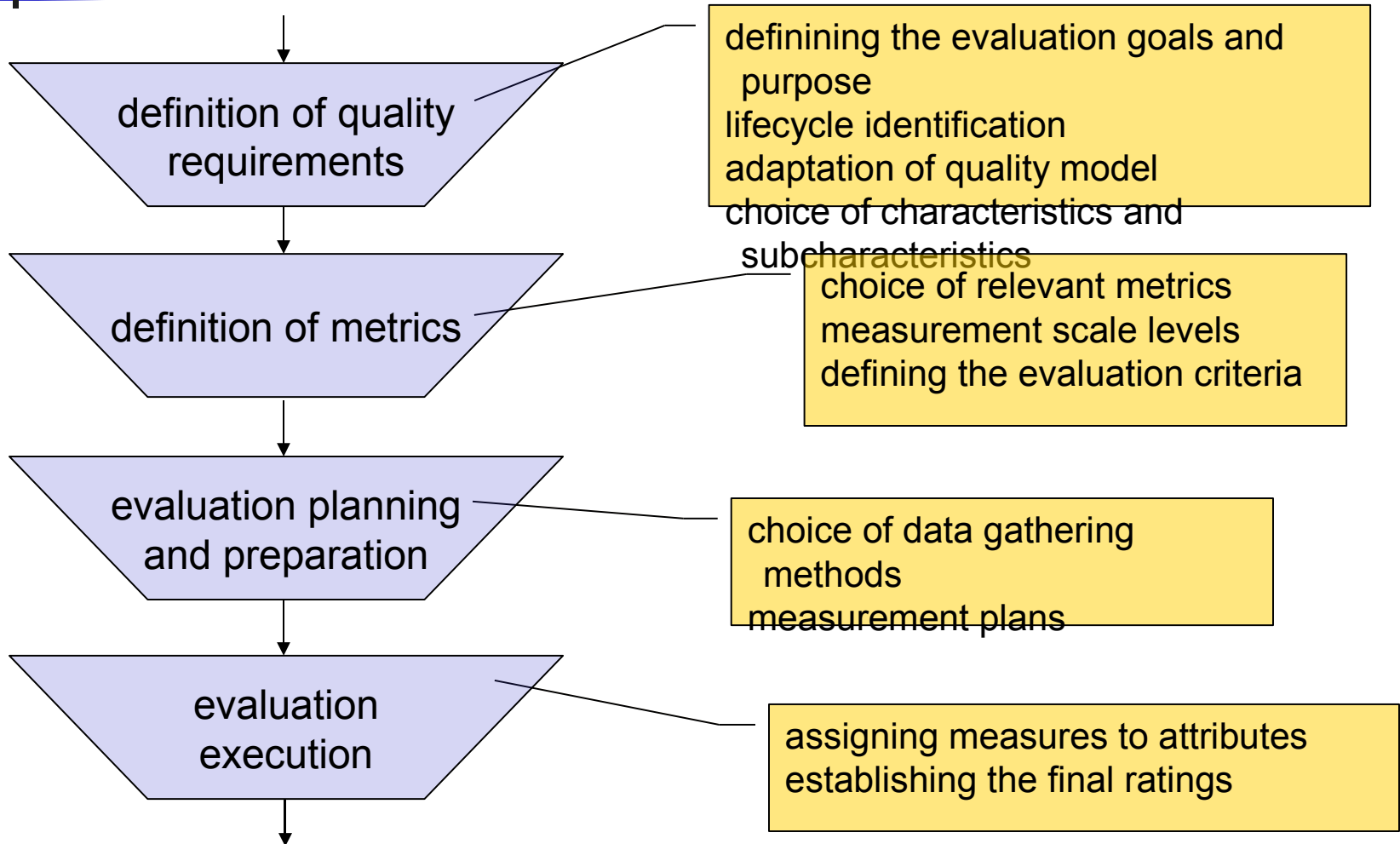
- capability of the site to be modified
 - **analysability** - capability of the site to be diagnosed for deficiencies or causes of failures, or for the parts to be modified to be identified
 - content generation, quality of code
 - **changeability** - capability of the site to enable a specified modification to be implemented
 - ease of change, administrator's privileges



Portability

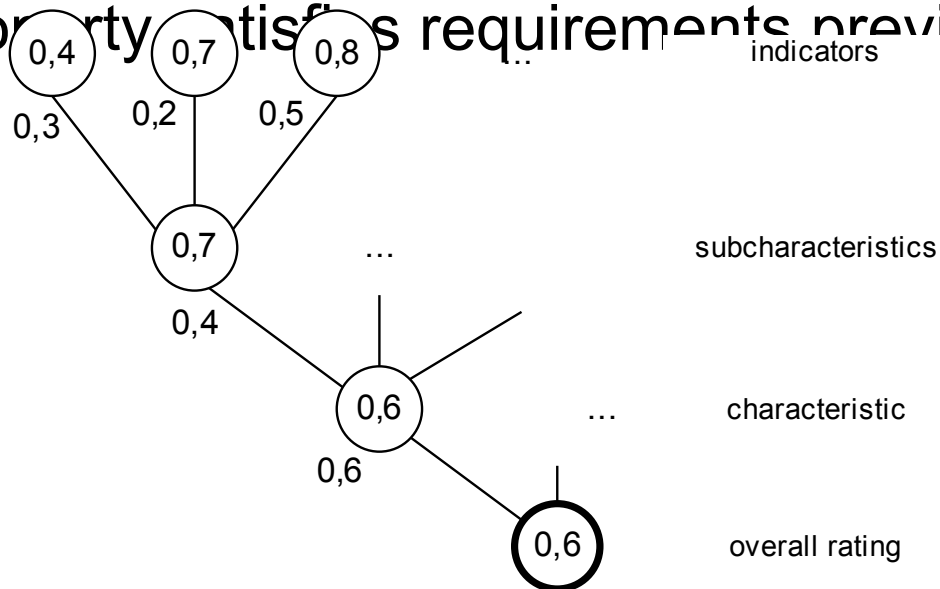
- capability of the site to be transferred from one environment to another
 - **adaptability** - capability of the site to be adapted for different specified environments
 - type of links, compliance to standards and guidelines.
 - **installability** - capability of the site to be installed in a specified environment
 - server's programming modules and components
 - **coexistence** - capability of the site to co-exist with other independent site components in a common environment sharing common resources
 - compliance to standards

Evaluation process



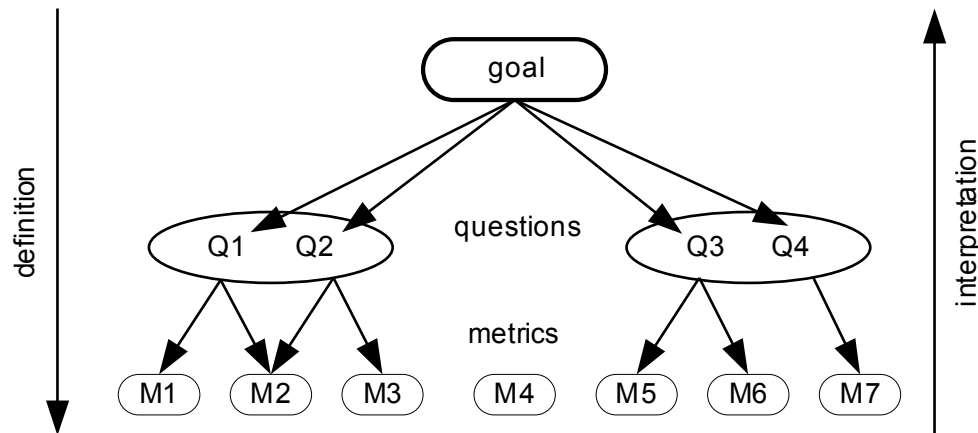
Metrics

- ISO definition
 - measurement method and its measurement scale which is used in measurement process to assign numerical values from the measurement scale to the measured attributes
 - metrics define the extent to which the measured property satisfies requirements previously set



Goal Question Metric Model

- the framework that guides the analysis which involves data collection
 - identification of goals of the analysis
 - specifying the relevant questions that make the goals more concrete
 - definition of metrics to give answers to the questions





GQM example

Characteristic: Usability

Subcharacteristic: Comprehensibility

Indicator: Content readability

4. What is the contrast of between the text and the background?
 - M1.1. brightness difference between text and background
 - M1.2. color difference between text and background
 - M1.3. solid text background color
 - M1.4. static text
5. What character set is used?
 - M2.1. suitability of the character set
 - M2.2. text size
 - M2.3. capital letters used
 - M2.4. margin alignment
6. Is there a need to scroll the text?
 - M3.1. horizontal scroll
 - M3.2. vertical scroll



Scoring techniques

$$A = w_1 A_1 + w_2 A_2 + \dots + w_n A_n, \quad 0 \leq A_i \leq 1; \quad 0 < w_i < 1, \quad \sum_{i=1}^n w_i = 1$$

- not appropriate for complex relationships
 - not possible to assess the effect of obligatory attributes
 - the impact of the attribute is limited by the weight factor
 - limited number of attributes



Logic Scoring of Preference

$$E = \sqrt[r]{w_1 E_1^r + w_2 E_2^r + \dots + w_n E_n^r}, \quad \prod_{i=1}^n w_i = 1, w_i > 0, \quad -A \leq r \leq +A$$

- takes into account more complicated relations among elements
 - E_i are elementary priorities
 - parameter r reflects the relation between elements
 - $r \rightarrow +\infty$ - disjunction
 - $r \rightarrow -\infty$ - conjunction
 - $r = 1$



Metrics example

M 3.1. horizontal scroll

data collection method: visual inspection

reason: More than 40% of users still uses resolutions of no more than 800x600 pixels. If the content of the page is too long, the vertical scroll bar on the right side will appear. If the content is too wide, the horizontal scroll bar on the bottom of the browser window will appear. Vertical scroll is acceptable and almost inevitable, but horizontal scroll should be avoided, because it irritates the user who has to move the scroll bar in order to read the page content.

measured value: The presence of horizontal scroll bar is monitored on the screen resolution of 800x600 pixels. It is acceptable if the scroll bar is not present on any of the pages, and it is not acceptable if the it appears on most of the pages. It is partly acceptable if it appears on the small number of pages.

elementary priority: E=1 for acceptable values, E=0 for unacceptable values, E=0,4 partly acceptable values.

sources: usability guidelines



Verification

- selected subset of Departmental sites at FER
- GQM model was used
 - 36 questions and 130 metrics were generated
 - additional 13 questions and 48 metrics defined for functionality characteristic
- evaluation performed using LSP method

		Web1		Web2		Web3		Web4	
	w_i	E_i	E	E_i	E	E_i	E	E_i	E
content readability	0,31		0,56		0,92		0,92		0,70
Q 1	0,37		1,00		1,00		1,00		1,00
M1.1	0,28	1,0	0,28	1,0	0,28	1,0	0,28	1,0	0,28
M1.2	0,28	1,0	0,28	1,0	0,28	1,0	0,28	1,0	0,28
M1.3	0,25	1,0	0,25	1,0	0,25	1,0	0,25	1,0	0,25
M1.4	0,19	1,0	0,19	1,0	0,19	1,0	0,19	1,0	0,19
Q 2	0,30		0,62		1,00		1,00		0,89
M2.1	0,22	1,0	0,22	1,0	0,22	1,0	0,22	1,0	0,22
M2.2	0,28	0,0	0,00	1,0	0,28	1,0	0,28	0,6	0,17
M2.3	0,25	0,6	0,15	1,0	0,25	1,0	0,25	1,0	0,25
M2.4	0,25	1,0	0,25	1,0	0,25	1,0	0,25	1,0	0,25
Q 3	0,33		0,44		0,82		0,82		0,55
M3.1	0,45	0,4	0,18	1,0	0,45	1,0	0,45	0,4	0,18
M3.2	0,36	0,7	0,25	1,0	0,36	1,0	0,36	1,0	0,36
M3.3	0,18	0,0	0,00	0,0	0,00	0,0	0,00	0,0	0,00



Evaluation results

	swing	w_i	Web1	Web2	Web3	Web4
functionality	100	0,29	0,75	0,64	0,78	0,57
usability	100	0,29	0,65	0,47	0,60	0,54
reliability	70	0,21	0,85	0,84	0,93	0,85
efficiency	70	0,21	0,65	0,59	0,64	0,64
overall quality			0,70	0,58	0,69	0,60



Conclusion

- the paper presents the methodology of evaluating Web site quality based on the quality model
- quality model:
 - three levels of hierarchy
 - 6 characteristics, 21 subcharacteristics, 43 basic indicators
- methodology independent of the model
 - Goal Question Metrics used
 - different users' perspectives
- the model was verified on the set of university sites