

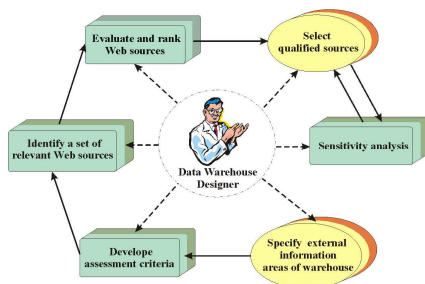
Motivation

- **Web** – independent platform for providing and accessing information
- **Data Warehousing** – supports OLAP and decision making in an enterprise
- An enterprise's internal data is insufficient for improving OLAP and making reasonable decisions
- Systematically integrating relevant external data from the Web with internal data in a data warehouse for reasonable decision making

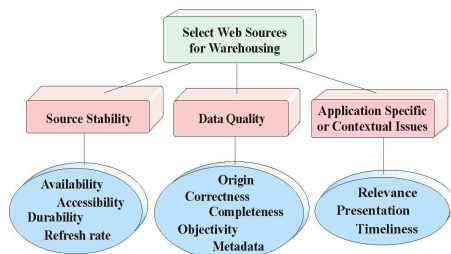
Issues of Warehousing Web Data

- Web source stability
 - autonomy and dynamics
- Web data quality
 - freely published on the Web
 - not carefully edited and reviewed
- Application specifics
 - Relevance, ease of extraction, and metadata

Steps of Information Source Evaluation and Selection



Evaluation Criteria

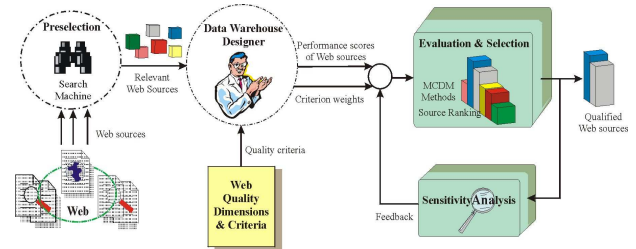


Evaluation Approaches

- Multi-Criteria Decision Making (**MCDM**) Approaches
- Compensatory MCDM methods
 - a decline in one attribute can be compensated by an enhancement in one or more other attributes
 - Scoring (e.g., SAW and AHP)
 - Compromising (e.g., TOPSIS)
 - Concordance (e.g., DEA)

Assessment Procedure

- The preselected Web sources are ranked in terms of criteria by using a MCDM method
- The higher its ranking score, the more qualified the Web source



Sensitivity Analysis

- How stable is the final rank of sources if critical measures (criterion weight, source performance score) are changed ?
- Which criterion or Web source is most sensitive ?
- How much must a measure change (threshold value) to cause the final rank reversion ?
- The most sensitive measure has the smallest threshold in all minimum relative changes

Comparison of MCDM Approaches

- SAW
 - ✓ simple additive weighting, synthetically consideration of the impact of all measures
 - ✗ evaluation measures are assigned subjectively
- AHP
 - ✓ consists of several techniques - decomposition, comparative judgment, and priorities synthesis
 - ✗ possible man-made inconsistency and time consuming in comparison
- TOPSIS
 - ✓ calculates the Euclidean Distance of alternatives
 - ✗ criterion with the highest weight has disproportionate influence on the ranking process
- DEA
 - ✓ Linear Programming-based, no need to subjectively assign weight values to criteria
 - ✗ needs to assume a zero value for some variables in order to make the number of variables meet the number of available constraints

Summary

- MCDM approaches are useful for a systematical and comprehensive evaluation process
- but need to subjectively assign weight values to criteria and to rate performance scores of alternatives - limitations
- Sensitivity analysis gains an insight of the impact of critical measures on the final decision
- SAW and AHP are simple to use, discriminative, robust, and suitable for Web source evaluation and selection
- TOPSIS and DEA are less suitable.