2nd Corporate Householding Workshop

Introduction

Richard Wang (rwang@mit.edu)
TDQM Program, M.I.T
Progress Report...

- What we suggested to pursue on April 26, 2002
- What we have done since the last meeting on April 26, 2002
- What we plan to do…
What have we accomplished?

1. Developed a TDQM sponsor only URL at TDQM home page: “Corporate Householding (password required)” containing
   - CHH Brochure: MIT TDQM research Initiative on CHH data for sponsors to pass out to clients
   - GAC filled Interview
   - All the presentation slides for 1st CHH Workshop
   - 2nd CHH Workshop materials after July 18
What have we accomplished?

2. Conducted a literature review of ROI and costs of data quality
   - Thanks due to Tony Nguyen and Xin-Ping Chen
   - 20 pages of report – see handout (for sponsors only)
   - A preliminary analysis – see handout (for sponsors only)
ROI and Cost of Data Quality II

- Search the ABI/INFORM online databases and search the following engines: Google, Yahoo, and MSN.
- Using “data quality”, Google, returned approximately 4 million matched values.
- Using “data quality costs” or “data quality ROI” yielded over 100,000 matched values, with many repeated and many no value, still a large number.
Due to time constraint, some 100 articles were read, 21 were selected for inclusion in this paper because they offer insights into the ROI and costs of DQ.

Findings include quotes, anecdotes, industry examples, and white papers published by companies.
ROI and Cost of Data Quality (III)

- Mostly Vendor White papers.
- Only one article was from academia.
- There is no standard measurement for data quality issues and problems and few enterprises routinely measure it.
- The costs for data quality initiatives are not hard to quantify, but the benefits are soft.
ROI and Cost of Data Quality (IV)

- Costs of data quality initiatives
  - Human resources, software, hardware…

- Benefits (or costs of poor data quality)
  - Estimate costs incurred and revenue lost in the critical business areas, such as customer services, product developments, billing, marketing, etc.
  - Estimate costs for prevention, correction and rework due to poor data quality
  - Information chain and data flow analysis (see paper)
2nd Corporate Householding Workshop

Current Research Plan

• Track I: CHH ROI and DQ Costs - Jesse, Tony and Xinping
  – a direct response to sponsors' voice of perceived value, value proposition, ROI, and Press release quotes.

• Track II: CHH Theories - Yair Wand and Yang Lee
  – Whoever can contribute to the establishment of theoretical foundations for corporate householding go to that category.

• Track III: CHH Technologies
  – Business Rules and CHH engine in Context - Wei
  – CHH technologies include applying COIN to mediate purposes, or a stand-alone research prototype that can be value-added to COIN mediator, and so on.

• Track IV: CHH Practices - by (potential) sponsors
  – case studies, interviews, etc. that will provide efficacy tests and monetary rewards to all involved.
2nd Corporate Householding Workshop

Voice of Sponsor – 1L

- What research directions are there?
- Do we focus on industry verticals, or functional horizontals?
- Do we focus on improving the context of householding data, and new data acquisition methods?
- What is the role of COIN in the future of householding?
- Are we going to catalog the list of industry-specific challenges to householding?
- What is the priority of the directions?
- We need to plan a progression of white papers that follow the research directions
- What MIT resources will be applied to the project and how?
- What synergies are there amongst the group members that can be leveraged?
- What deliverables against the research plan should each member be assigned?