Informed Participation
and
The National Information Infrastructure

Thomas A. Kalil
National Economic Council
The White House

John C. Mallery
Artificial Intelligence Laboratory
and Political Science Department
Massachusetts Institute of Technology

with commentary by

Roger Hurwitz
Artificial Intelligence Laboratory
Massachusetts Institute of Technology

How can digital computer networks be used to improve the policy-making process in government? Thomas Kalil will review opportunities for public input into decision processes. Then, John Mallery will discuss several technologies that might be applied to public access. Finally, Roger Hurwitz will comment and guide the discussion, focusing on the need to ensure fair access to public discussion and to avoid the imposition of new barriers to entry.

Menu

1. The Fan-in Problem
2. Communications Integration
3. Modes of Communications Integration
4. Interpretive Issues
5. Conclusions

The Digital Network Revolution

- Over 40 million people worldwide have access to email.
- The number of internet hosts is growing at 17 percent per month, and there are already over 2 million connected.
- In March 1994, NSFnet traffic grew by 20 percent.
- Distributed hypermedia (WWW) traffic growing at 1 percent per day.
- 8 terabytes of data cross the US internet every month, whereas the figure is 1 teraflop for Europe.
- Telecommunications giants and computer companies are positioning for what many believe has become a single multimedia market.
- The White House is publishing press releases and major documents electronically and also accepting email to the President and Vice-President.
- The US Congress is experimenting with email and other internet connectivity.
- The OMB A-130 Circular directs Federal agencies to devise plans for interfacing electronically with the public and each other.
- The Government Information Locator System and agency email plans.
- Legislation in support of the National Information Infrastructure.

Challenges for Wide-Area Communications

- Focused information distribution and retrieval.
- Distributed data-integration and analysis.
- Automatic management of information bottlenecks.
- Enabling large-scale collaborative and interactive problem-solving.
- Distributed (inconsistent!) knowledge representation and cooperative computing.
- Integration of organization design with computational infrastructure.
- Representing political processes for communication.

Modes of Communications Integration

- **Electronic Surveys:** Electronic surveys make it possible to determine what a large number of people think.
- **Self-Framing Surveys:** Electronic surveys can be improved by allowing the respondents to define the survey and frame the problem before responding to the actual survey.
- **Policy Hypertexts:** Distributed hypermedia can enable a structured conversation about arguments for or against means and goals (norms).

Communication as Representation

- **Routing:** The granularity of routing (retrieval) depends on the degree to which the communication is understood, or represented.
- **Precision:** The precision with which communications can reach recipients depends on the granularity of indexation or representation.
- **Accuracy:** The accuracy of routing depends on the match between the sender's criteria for routing and the recipient's understanding of those criteria.
- **Data Ontology:** Various paradigms of communication follow from the characteristics of different representations.

Paradigms of Information Routing

1. **Static Routing:**
   - **Mailing Lists:** fixed pointers that change slowly.
   - **Hierarchical Mailing Lists:** Taxonomically arranged mailing lists.
2. **Dynamic Routing:**
   - **Taxonomic:** Routing on the basis of boolean combinations of keywords associated with categories in a multidimensional taxonomy (Top-down)
   - **Filtering:** Routing on the basis of user defined statistical filters on messages. (Bottom-up)
   - **Mixed-Strategy:** Routing combines methods taxonomic and filtering approaches with some syntactic parsing (*e.g.*, noun phrase parsing) and possibly some limited knowledge (*e.g.* sense disambiguation).
3. **Knowledge-based Routing:**
   - **Human-Assisted:** Indexation and routing based on categories or links supplied by people (e.g., distributed argument hypertext).
   - **Semantics-based:** Routing based on syntactically parsing and semantically representing message so that recognizers at the semantic level can assign relevance to parts of messages.
   - **Understanding-based:** AI-complete system does not make errors in representing the message so that it can do the right thing.

1992 Presidential Campaign

- Implemented databased mail agent system
- Implemented automatic form processing
- Implemented automatic survey research
- Conducted three Automatic surveys
- Direct participation by roughly 2500 people
- Small, discernible impact of participants on campaign materials
- Large indirect participation (c. 100,000+ ?)
- Malicious hacks and wacky errors showed importance of dynamic linking.
- Acquired data for future study

Current Research

1. **Artificial Intelligence**
   - Email Server Substrate - Mallery, Shrobe
   - Knowledge-based Postmaster - Naha, Shrobe
   - White House Publications System - Mallery, Hurwitz, Shrobe
   - Pollster's Assistant - Hurwitz, Renaud, Mallery
   - Data Exploration Environment - Renaud, Mallery
   - Intelligent Receptionist - Katz
   - Correspondence System - Shrobe, Davis, Hurwitz, Mallery
   - Transactional Stereotypes - Hurwitz, Malone, Mallery
   - Natural Language Processing - Berwick, Katz, de Marken, Mallery
2. **Social Science**
   - Political Communication - Hurwitz
   - Policy Hypertexts - Hurwitz, Mallery
   - The Digital Organization - Tom Malone (MIT Sloan School)
   - Sociology of Communication - Lee Sproull (Boston University)

White House Publications System

1. Automatic Subscription Maintenance.
2. Taxonomic Document Routing.
   - Top-down by boolean combinations of categories.
   - *Bottom-up by words and phrases.*
3. Document Retrieval.
   - Top-down by boolean combination of categories.
   - Bottom-up by WAIS-style queries.
4. Standing Usage Survey.
5. Publications Interface.
6. Failed Mail Processing.
7. *Automatic Document Categorization.*

Email Server Substrate

- Subject line command processing.
- Form processing based on CLIM.
- Reliable task execution.
- Persistent object database.
- Myriad issues concerning networks and mailers.
- Specification/Protocol for a remote GUI for forms.
- GUI for form definition and release.
- *Scaling for high volume.*

Correspondence System

1. **Robomail:** Returns form letter responses from a stock of predefined responses.
   - **Statistics:** Use statistical methods to match documents to replies. (Problem with negation.)
   - **Form Processing:** Use forms to discriminate and return the correct response.
   - *Generation: Generate custom responses from a knowledge base of answers (the representation of the response forms).*
   - *NLP: Parse and represent inbound correspondence to determine the appropriate response form.*
2. **Service Processing:** Use a form-processing email server to act as an self-service information desk, doling out information on request and providing a range of services.
3. **Opinion Sampling:** Use automatic surveys to elicit more than whether someone is "for or against" an issue or policy.
4. **Argument Hypertext:** Allow people to collectively construct hypertext reasons *for* or *against* issues or policies. People could also endorse existing links.
5. **Reason Sampling:** Use form processing to acquire relational models of people's opinions.

Pollster's Assistant

- Survey research building on automatic form processing.
- Hierarchical and adaptive questions.
- User friendly follow-up to failing or omitted queries.
- Portable Interface for rapid survey design, release, and analysis.
- Machine learning algorithms to explore returns.
- *Meta-Surveys that allow respondents to write the survey (delphi style).*
- *Argument Hypertexts that allow participants to flesh out rationales for and against policies in a distributed cooperative manner.*
- *Acquiring relational models of people's beliefs and norms.*

Policy Hypertexts

- **Transport:** Utilize email and WWW over networks for comment on policy.
- **Working Groups:** Internal applications to improve bureaucratic coordination (**e.g.**, organizational restructuring, complex negotiations).
- **Public comment:** External applications for to structure and widen public participation (**e.g.**, NII, GII).
- **Argument Connectives:** Support, oppose, qualify, extend, reframe, positive and negative examples.
- **Problem-Solving Connectives:** Goal, function, enabling structure.
- **Equality:** Methods will be needed for determining if opaque text is the same or not.
- **Substructure:** Methods will be need to decompose text so that comment can be made on parts of linked policy texts.

Limitations in Communications Integration

- **Electronic Surveys:** Electronic surveys require people to communicate through frameworks prestructured by survey designers who may not anticipate all possible viewpoints.
- **Self-Framing Surveys:** Surveys are normally associated with an outcome that is aggregated statistically.
- **Policy Hypertexts:** Opacity limits fine-tuned linked as the link repertoire restricts expressibility.

Interpretive Issues

- **Broader Communities:** Technical specialization, educational level, region, sectoral affiliation, class, ethnicity, or nationality.
- **Educational Level:** Technical specialization or years of schooling.
- **Educational Role:** The infrastructure needs to help people build maps across different belief systems.

Consequences

- How can we make complex systems accessible to everyday people?
- How can organizations become more effective through better information processing and intelligent agents?
- How will tighter information integration change organizational dynamics? Heightened conflict or oppressive homogenization?
- What will be the consequences for political systems, especially democratic ones such as our own?