

## Our Focus ...

- ***Information Resource Discovery:***

Where do I find what I am interested in ?

- ***Query Expressiveness:***

How do I tell the Global Information System what I am interested in ?

Little new in the  
"old areas" of  
consistency mgmt  
traditional  
query optimization

## Semantics-based approach

### Based metadata approach

*roach:*

*ase:*

e **context** of **user query**  
and display ontologies for  
mains for the above

Domain  
Ontology

Query  
Context

COMPARE

Information  
Resource  
Context

*base:*  
**stadata** from information source  
**mappings** between metadata  
lation  
**information resource context**  
Metadata  
Metadata

# *Challenges with current techniques for Information Resource Discovery*

---

---

- ***Content-based access:***

Consider a paper titled “Heterogeneous Information” to be submitted to a conference in Singapore.

The paper ***may not contain*** the word ***Singapore***

Consider the following ***keyword-based*** query:

***“document submitted Singapore”***

The paper mentioned above shall not be retrieved !!

- Need a descriptive approach to associate this information with document meta-data.

## But ....

ograms were great for browsing, they aren't for finding a specific peice of information on fact, we couldn't even use them to find any nentation on HTML, though we found many ns.

Mosaic: Two Free Tickets Around t," PC Magazine, April 26, 1994.]

# Challenges with current techniques for Information Resource Discovery

- **Lack of controlled use of terminology:**

To: kashyap@paul.rutgers.edu (Vipul Kashyap)  
From: publications@whitehouse.gov (WhiteHouse.Gov publications server)  
Subject: RE: your request

## Regarding your request:  
topic INDIA

Matching Filenames:

File-#Name  
25629 Office-of-Nava-jo-and-Hopi-Indian-Relocation (397 bytes)  
25654 National-Commission-on-American-Indian,-Alaska-Native,-and-Native-Hawaiian-Housing (436 bytes)  
25668 Institute-of-American-Indian-and-Alaska-Native-Culture-and-Arts-Development (430 bytes)  
148626 National-Indian-Gaming-Commission (1556 bytes) 148632 Bureau-of-Indian-Affairs (12368 bytes)  
153622 Public-and-Indian-Housing-Programs (3426 bytes) 158930 Indian-Health-Services (2552 bytes)  
133206 1994-04-29-Babbitt-and-Deer-Briefing-on-Indian-Affairs (27192 bytes)  
133305 1994-04-29-President-in-Meeting-with-Indian-Tribal-Leaders (17939 bytes)  
30837 1994-05-19-President-and-India-PM-RAO-in-Press-Availability (26732 bytes)  
30925 1994-05-11-President-Names-Frank-Wisner-as-Ambassador-to-India (1204 bytes)  
81960 1994-08-02-Eight-Named-National-Advisors-on-Indian-Education (2721 bytes)  
81962 1994-08-03-Four-on-American-Indian-Culture-Development-Board (2289 bytes)  
179395 Aid-to-India-10-01-93 (1855 bytes)  
194579 Transcript-of-Presidents-Press-Conference-w-Indiana-Press (36729 bytes)  
322589 National-Commission-on-American-Indian,-Alaska-Native,-and-Native-Hawaiian-Housing (931 bytes)  
322609 Institute-of-American-Indian-and-Alaska-Native-Culture-and-Arts-Development (925 bytes)  
66569 Indian-Affairs (13258 bytes)  
66564 National-Indian-Gaming-Commission (2393 bytes)  
66575 Public-and-Indian-Housing-Programs (4869 bytes)  
102417 Indian-Health-Services (3453 bytes)

- Context for the word INDIA lacking (Asia v/s America)

Country vs Ethnic  
Group

## *!S with current techniques*

### *ation Resource Discovery*

#### *s of Navigation and Browsing*

) if the number of links are more than 3 or 4  
ate links with contextual information in  
duce the “link-chasing”

#### *ems in Indexing information:*

all the information on the internet !!  
ex heterogeneous but related information  
sults obtained by using independent/  
s

#### *'pre-determined relationships:*

ght make some hyper-links meaningless !!  
rganizations might prove expensive to

cified criteria for search is different from  
organization

## Some approaches ....

- **User centered approach:**
  - menu-based browsing
  - hypertext browsing
- **Current Emphasis**
  - **Syntactic/Structural approach:**
    - information retrieval, indexing techniques
      - WWW/Mosaic
      - WAIS
    - name and attribute-based search, pattern matching
      - Archie
  - **Descriptive (symbolic) semantics-based approach:**
    - making design assumptions explicit
    - capturing the semantics of the query
  - **Cognitive (sub-symbolic) semantics-based approach:**
    - Pattern/Speech Recognition Algorithms ?
    - Neural Networks ?

# the environment Execution Requirements sues and Requirements inforable specifications

- )lation of intertask dependencies  
• , may be not)  
• states)
  - eduler)
  - ot known at design time
  - added dynamically
  - t invocation of the same task type

## *Conclusions*

- Better relevance and wider applicability than heterogeneous DDBMS and multidatabase transactions.
- It is the way to provide glue to handle legacy systems, to support migration/evolution.

# *Research Projects and Prototypes on Relaxed Transaction, Workflow and Related Issues*

*(a partial list)*

- APRICOT (Germany), ETM (DEC), Interbase (Purdue),
- Carnot (MCC), ASSET (Bell Labs),
- TSME/DOM (GTE Lab),
- Pegasus (HP Lab), METEOR (Bellcore, UofH),
- TriGSflow (Austria), INCA (MITL),
- ...

## Summary of technical issues

Issue 3:

Information Manipulation  
[multisystem applications and  
transactional work-flows/activities]

Issue 1:

Information  
modeling, discovery,  
resolving heterogeneity

Issue 2:

Information integrity/consistency  
[multidatabase dependencies]