

Knowledge warehousing with WebKB-2

A knowledge server for structuring, sharing and publishing information

Dr Philippe MARTIN

philippe.martin@gu.edu.au



The on-line slides are accessible at <http://www.webkb.org/doc/slides/dstc0302.html>

Plan

- Documents, databases, knowledge bases
- Data/knowledge warehouses
- The advantages of WebKB-2
- Examples with generated interfaces
- Examples with a tailored interface
- More examples of knowledge representation/sharing

Documents, databases, knowledge bases

- The semantic content of documents cannot be extracted automatically
- Lexical techniques (e.g. keyword-based techniques) do not support question answering (\rightarrow precision/recall ratio, redundancies, no structure/inference)
- Data indexation and retrieval in databases and structured documents must follow *predefined* schemas
- Knowledge bases (KBs) permit to store and interconnect logical representations of facts, rules or definitions according to dynamically updatable hierarchy of concepts and relations (ontologies), and thus may support knowledge sharing, comparison and retrieval via conceptual browsing or querying

But knowledge representation is a difficult manual task that is not much facilitated by current knowledge base systems/servers.

Data/knowledge warehouses

Merging data from distributed databases require specially-designed wrappers (+ middlewares such as CORBA) and may be inefficient. In such a case, data warehousing is the solution.

Correctly merging knowledge representations from independently developed KBs is much more difficult, even manually.

Genuine knowledge sharing implies knowledge warehousing.

The advantages of WebKB-2

- Can be used by Web users/agents (www.webkb.org)
- Supports very large KBs
- Initialized with the biggest general KB after CYC
- Exploits the KB to generate forms for guiding and easing knowledge entering
- Has intuitive expressive and high-level input/output formats
- Permits users to update a shared KB without lexical/semantic conflicts nor redundancies, *and* without obliging the users to agree with each other.

Applications:

- Yellow-Pages, Auctions, Classifieds, Shopping, Jobs, Personals, ...
- Corporate memories, cooperative ontology building, state of the art

Examples with generated interfaces

- Looking for the price of used Toyota Corolla manual cars in Australia
- Looking for used cars for sell in Southport for at most 7000 AUD
- Looking for new cars for sell by a car dealer

Netscape: Query result

File Edit View Go Communicator Help

Location: <http://meganesia.int.gu.edu.au/~phmartin/WebKB2/bin/termDisplay.cgi?term=wn%23car> What's Related

Category #car (or: auto, automobile, machine, motorcar) *4-wheeled motor vehicle; usually propelled by*
[generalization: motor_vehicle](#)
[specializations: ambulance, station_wagon, taxi, compact_car, convertible, coupe, patrol_car, electric](#)
[equals: horseless_carriage](#)
[parts: accelerator_pedal, air_bag, auto_accessory, automobile_engine, automobile_horn, buffer.device,](#)

1 schema is about **car**
[graph1_on_car](#)

```
[any car (^General characteristics^),
  may have for kind: car_of_a_certain_make (^$(explore 2)$^),
  may have for kind: commercial_or_passenger_car (^$(explore)$^),
  may have for kind: new_or_used_car (^$(explore 1)$^),
  may have for kind: automatic_or_manual_car (^$(explore)$^),
  may have for kind: hatchback_or_sedan_car (^$(explore)$^),
  may have for kind: car_with_a_certain_energy_source (^$(explore)$^),
  may have for kind: car_with_a_special_shape (^$(explore)$^),
  may have for kind: car_with_a_function (^$(explore)$^),
  may have for kind: car_with_a_role (^$(explore)$^),
  may have for kind: car_with_a_modification (^$(explore)$^)]
```

];

1 statement is about a direct instance of **car**: [pm#phmIsOwnerOfSomeRedCar](#)

4 statements are about indirect instances of **car**: [GlassGuide#graph1_on_Corolla_AE93_CSI_manual_hatchba](#)
[click here to display them](#) or click [here for a search form](#) or [here to add a statement](#)

100%

Fill the "relation: object" pairs you wish, then submit. You may add new pairs in the large text edit. For each object, you may specify intervals, collections, quantifiers and relations. The sub-menus "+", "proposes relations". Object examples are given, prefixed by "//". [Click here for details.](#)

Search statements about some car: most #car

General characteristics

kind: Toyota pm#Toyota +

kind: commercial_or_passenger_car +

kind: used_car pm#used_car +

kind: manual_car pm>manual_car +

kind: hatchback_or_sedan_car +

kind: car_with_a_certain_energy_source +

kind: car_with_a_special_shape +

kind: car_with_a_function +

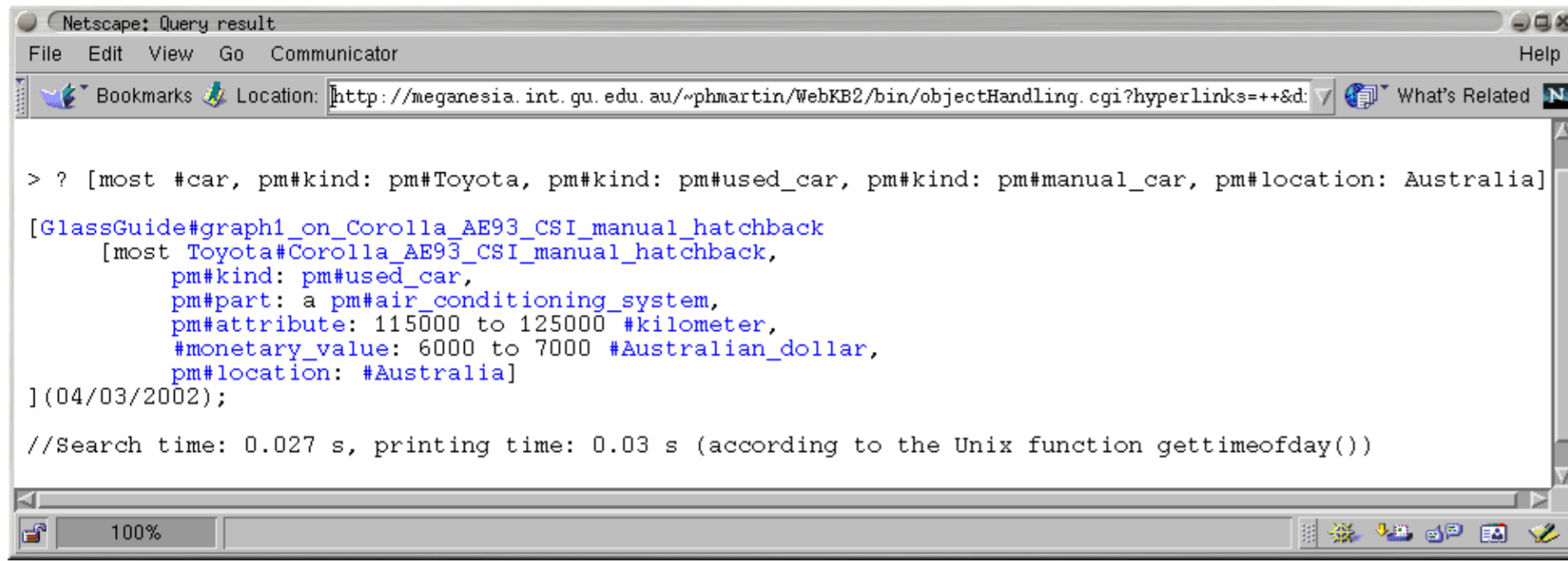
kind: car_with_a_role +

kind: car_with_a_modification +

relation: a_value, another_relation: some other_value //clear this example if yo

physical_part: //a pm#physical_ent + matter: //a #substanc

physical_part_of: //a pm#physical_ent + location: Australia



The screenshot shows a Netscape browser window with the title "Netscape: Query result". The address bar contains the URL "http://meganesia.int.gu.edu.au/~phmartin/WebKB2/bin/objectHandling.cgi?hyperlinks=++&d". The main content area displays the following text:

```
> ? [most #car, pm#kind: pm#Toyota, pm#kind: pm#used_car, pm#kind: pm#manual_car, pm#location: Australia]
[GlassGuide#graph1_on_Corolla_AE93_CSI_manual_hatchback
 [most Toyota#Corolla_AE93_CSI_manual_hatchback,
  pm#kind: pm#used_car,
  pm#part: a pm#air_conditioning_system,
  pm#attribute: 115000 to 125000 #kilometer,
  #monetary_value: 6000 to 7000 #Australian_dollar,
  pm#location: #Australia]
](04/03/2002);
//Search time: 0.027 s, printing time: 0.03 s (according to the Unix function gettimeofday())
```

The browser's status bar at the bottom shows a zoom level of 100% and various system icons.

Searching information about most used Toyota Corolla manual cars in Australia.

The screenshot shows a Netscape browser window with the title "Netscape: Query result". The address bar contains the URL "http://meganesia.int.gu.edu.au/~phmartin/WebKE2/bin/objectHandling.cgi?h". The main content area displays a query result for a car listing. The query is: "> ? [a #car, pm#kind: pm#used_car, pm#object of: (a #sell, #asking_price: at most 7000 AUD)]". The result is a list of objects, starting with "[pm#graph2_on_Corolla_AE93_CSI_manual_hatchback". The first object is a Toyota Corolla AE93 CSI manual hatchback, with details: pm#kind: pm#used_car, pm#part: a pm#air_conditioning_system, pm#attribute: 120000 #kilometer, #color: a #red, pm#owner: philippe.martin@gu.edu.au, #monetary_value: 7000 #Australian_dollar, pm#location: QLD#Southport, pm#object of: (a #sell, #asking_price: 6500 #Australian_dollar, pm#time: 9/3/2002, pm#place: QLD#Southport), pm#object of: (a #trade, pm#time: 21/4/1998, pm#place: QLD#Southport, pm#seller: pm#Car_Carnival, pm#customer: philippe.martin@gu.edu.au, #purchase_price: 11400 #Australian_dollar), pm#object of: (an #insurance, pm#from_time: 28/4/1999, pm#until_time: 28/4/2000, pm#seller: pm#Suncorp_Metway, #customer: philippe.martin@gu.edu.au, #monetary_value: 476.55 #Australian_dollar)]. The result ends with "](04/03/2002);". At the bottom, there is a comment: "//Search time: 0.027 s, printing time: 0.08 s (according to the Unix function gettimeofday())". The browser window also shows a status bar at the bottom with "100%" zoom and various icons.

```

> ? [a #car, pm#kind: pm#used_car, pm#object of: (a #sell, #asking_price: at most 7000 AUD)]

[pm#graph2_on_Corolla_AE93_CSI_manual_hatchback
 [some Toyota#Corolla_AE93_CSI_manual_hatchback,
  pm#kind: pm#used_car,
  pm#part: a pm#air_conditioning_system,
  pm#attribute: 120000 #kilometer,
  #color: a #red,
  pm#owner: philippe.martin@gu.edu.au,
  #monetary_value: 7000 #Australian_dollar,
  pm#location: QLD#Southport,
  pm#object of: (a #sell,
    #asking_price: 6500 #Australian_dollar,
    pm#time: 9/3/2002,
    pm#place: QLD#Southport),
  pm#object of: (a #trade,
    pm#time: 21/4/1998,
    pm#place: QLD#Southport,
    pm#seller: pm#Car_Carnival,
    pm#customer: philippe.martin@gu.edu.au,
    #purchase_price: 11400 #Australian_dollar),
  pm#object of: (an #insurance,
    pm#from_time: 28/4/1999,
    pm#until_time: 28/4/2000,
    pm#seller: pm#Suncorp_Metway,
    #customer: philippe.martin@gu.edu.au,
    #monetary_value: 476.55 #Australian_dollar)]
](04/03/2002);

//Search time: 0.027 s, printing time: 0.08 s (according to the Unix function gettimeofday())

```

Searching information about used cars for sell in Southport for at most 7000 AUD.

Examples with a tailored interface

- Locating short-term accommodations and attractions on the Sunshine Coast
- Locating hotels or bed-and-breakfast charging less than 100 AUD a night, where check-out is no sooner than 11:00, without a minimum stay of 3 nights, near a Chinese restaurant
- Locating Eurasian restaurants near the selected accommodation
- Locating Chinese restaurants serving seafood

Netscape: search accommodations
File Edit View Go Communicator Help
Location: http://meganesia.int.gu.edu.au/~phaartin/WebKB/ib/wotif/s.html What's Related

Locate and compare accommodations on the Sunshine Coast

Apartment Hotel Bed&Breakfast Backpackers Caravan/Camping/Park From Wotif list only

Show nearby Tourist attraction Activity / Other

Room price (per day). Min: 10 AUD Max: 900 AUD Filter out if information unknown

Filter out accommodations with a minimum stay of 4 nights, [More price info](#) [Tours](#) [Surf spots](#)

where check-in has to be later than 14:00 and check-out sooner than 10:00

100%

Netscape: Object Search Form

File Edit View Go Communicator Help

Location: <http://meganesia.int.gu.edu.au/~phmartin/WebKB/bin/objectForm.cgi?term=%23resta> What's Related

Sub-menu for some restaurant

Fill the "relation: object" pairs you wish, then submit. You may add new pairs in the large text edit. For each object, you may specify intervals, collections, quantifiers and relations. The sub-menus '+' proposes relations. Object examples are given, prefixed by "///". [Click here for details.](#)

address: + phone_No:

menu: + cuisine_origin: +

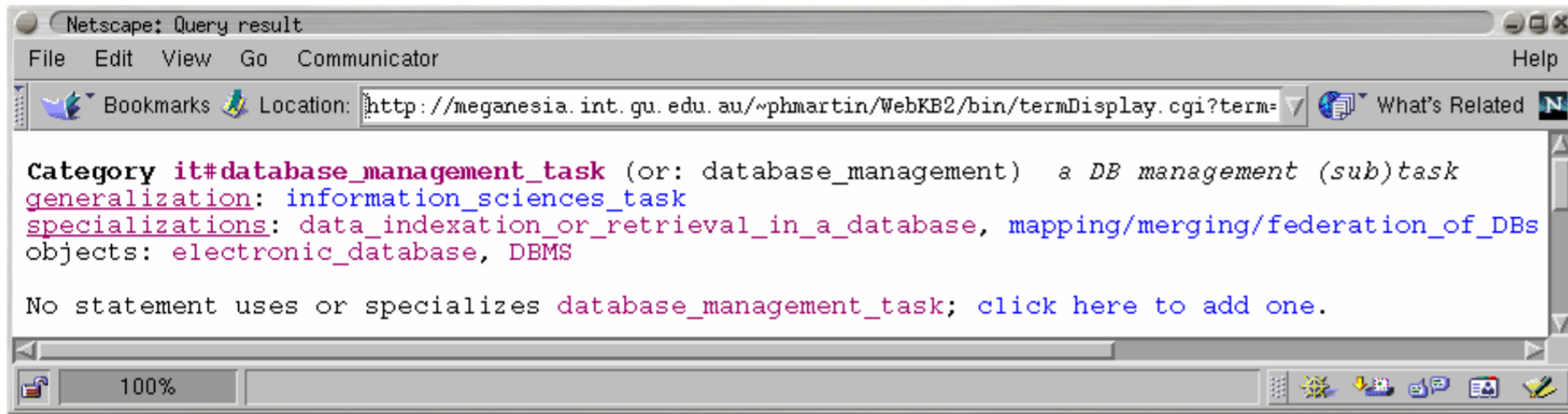
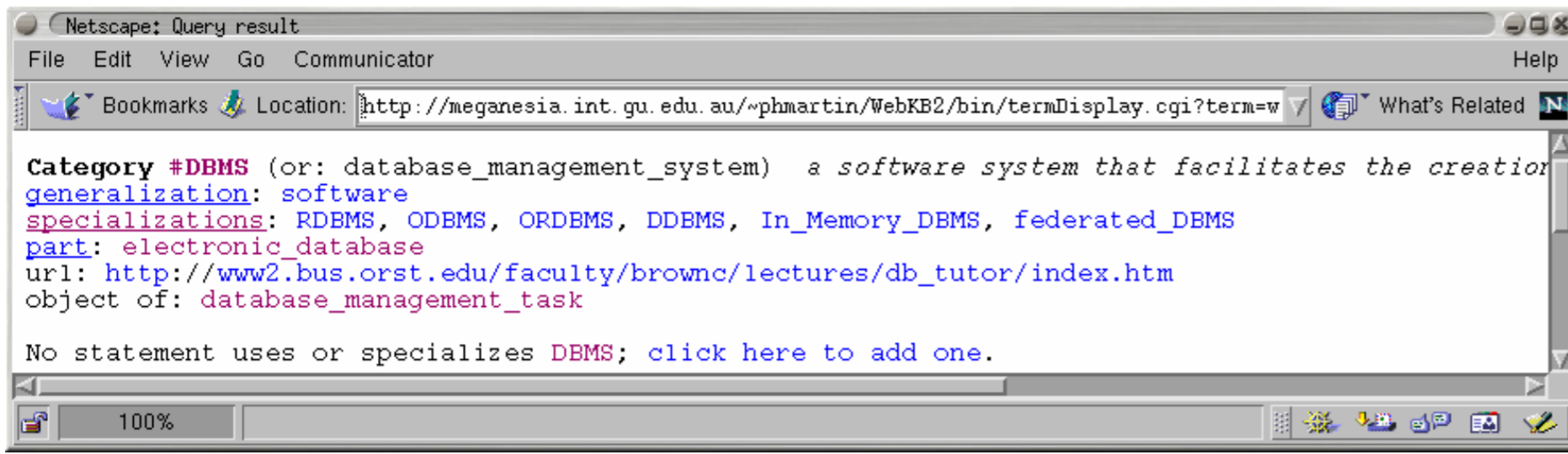
facility: + owner: +

100%

Generated submenus can be used when the tailored menu is insufficient.

More examples of knowledge representation/sharing

- About databases
- About information technology
- Other examples



The small part of the current Information Technology ontology related to databases.