Open

CC/PP and Beyond

Talk at the W3C Delivery Context Workshop
Sophia Antipolis, France, March 2002

Johan Hjelm
Senior Specialist
Ericsson Research

Main Points

- Ontologies, not vocabularies
- Vocabularies are a social problem (this is why content negotiation does not scale)
- The mechanisms have to be in the ontology mechanisms
- Device independence, content adaption, and location-based services are just special cases of ontology processing
- It’s all a matter of scoping

What is CC/PP?

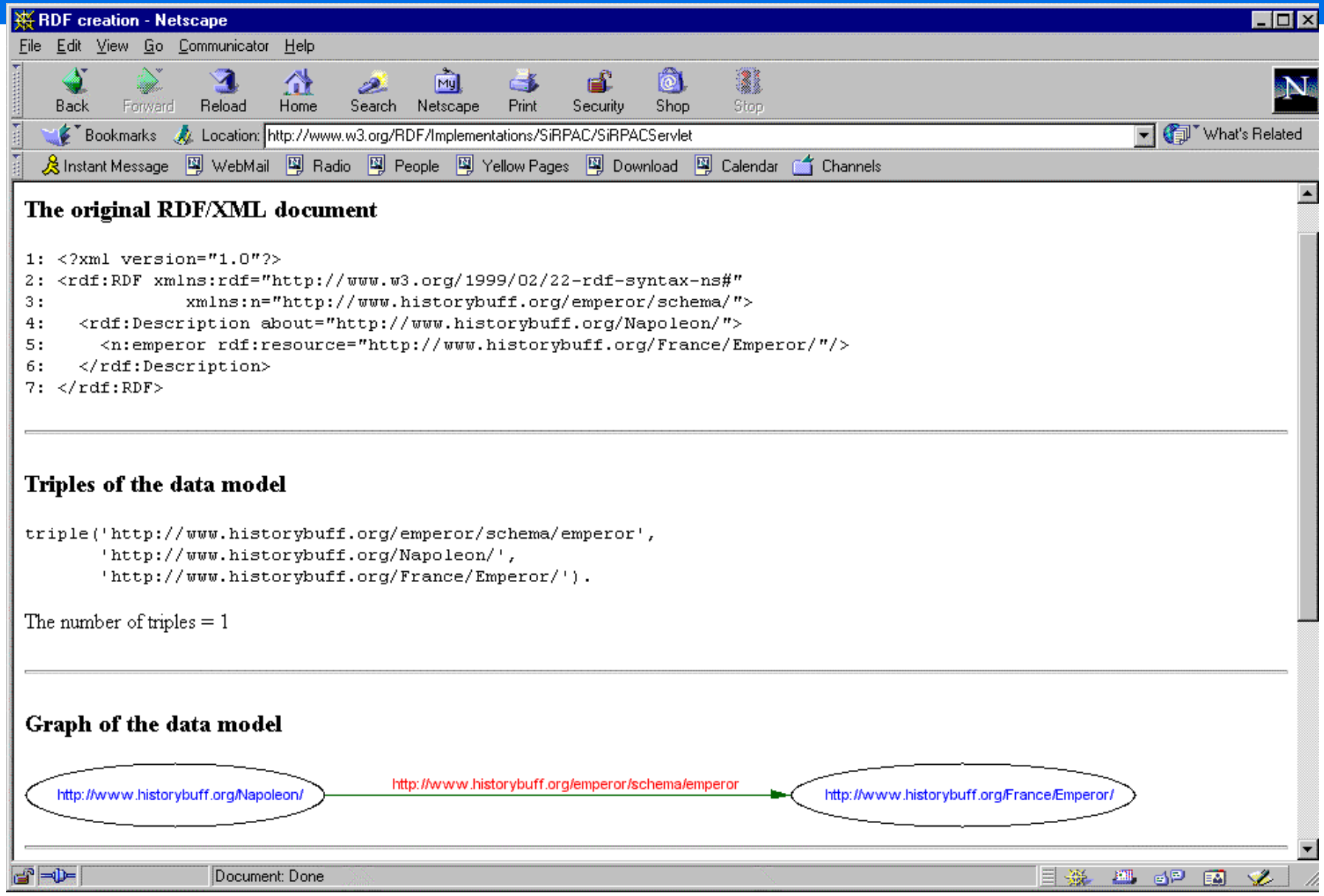
- Composite Capabilities/Preferences Profile
- A framework for contextualization
  - Adopted by the WAP Forum for terminal capabilities communication
  - But it is not tied to WAP, and can be used for other things as well
- A profile framework
  - Profiles are “structured assertions about an object within a framework”
- An attempt to minimize the number of transactions in server-based content negotiation

CC/PP design

- The profile is organized in components
  - The only predefined components in the specification are hardware, software, and proxy
  - Anyone can define their own components (the WAP Forum did)
- Components contain default values
  - Overrides can replace defaults
- The design leverages RDF for namespaces, assertion structure, and XML encoding

Why did you choose RDF, and not XML?

- RDF is XML, but it adds features
  - Triples give elements relations
  - RDF is a graph format, so encodings are equivalent - which means an application can work with triples, and humans write graphs
  - The XML encoding brings transportability and all the advantages of XML
- Because RDF enables us to use the user profile as part of the Semantic Web
  - RDF has roots in metadata, knowledge management, databases, and AI
- We wanted to understand if you could use a framework like RDF to represent a physical entity

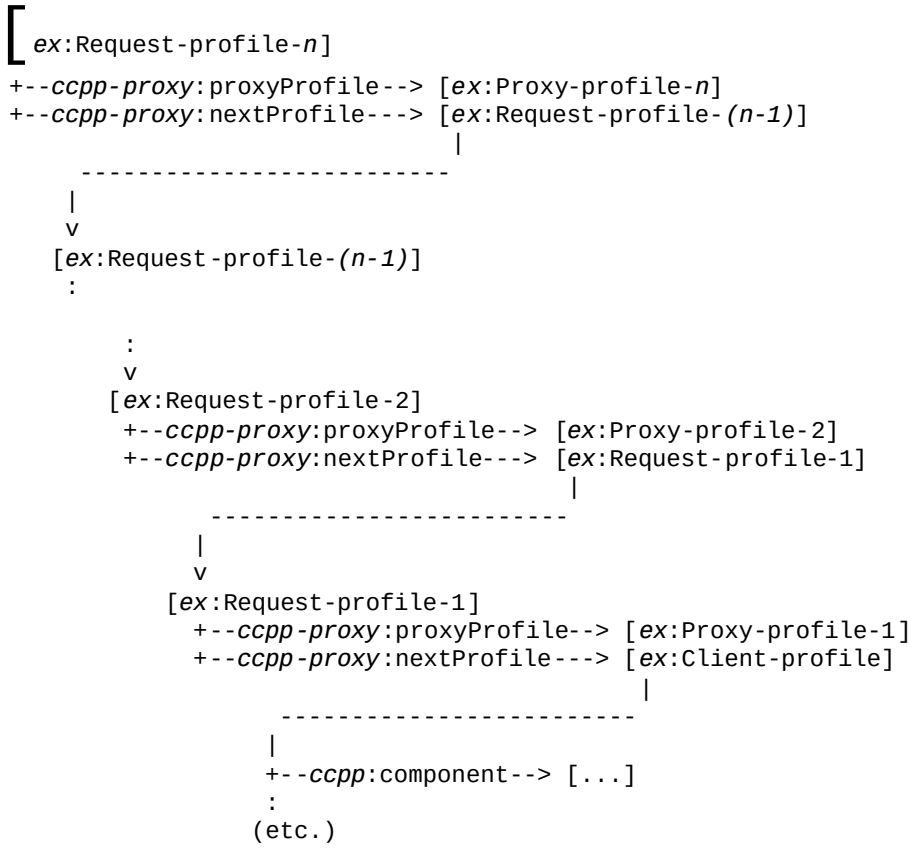


An example of graph, triple, and document from the SiRPAC parser/validator

Overrides, Defaults, and Caching

- CC/PP has a concept of defaults, which can be overridden.
- The profile is organized in components
- Overrides apply at element level
- Components (I.e. defaults) can reference a URI, thus use the HTTP cache mechanisms to speed up retireval

Capability Chaining



- Profiles can be chained (several proxies can add their capabilities to the client profile)
- The end result represents the capabilities of the chain of components from the viewpoint of the origin server
- Proxies can add capabilities (e.g. transcoders) or block capabilities (e.g. firewalls)
- This part of the specification is not mandatory

Where are we now?

- Soon to be Candidate Recommendation
  - Only thing missing is getting Philipp Hoschka happy with Disposition of Comments, then we go to the director for CR approval
- After that, short step to Proposed Recommendation
  - We need information about (commercial) implementations, to show that all features are implemented - please provide them to Kaz (kaz@w3.org)

How can anyone make their own vocabulary?

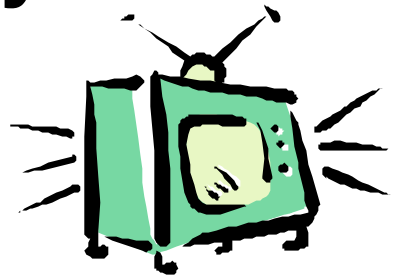
- RDF uses XML namespaces
  - Namespaces are based on URIs (which are based on DNS) and makes element names unique
- Multiple namespaces can be used in a RDF document
- RDF Schemas (which describe the elements in the vocabulary) are accessible over the Internet
- Anyone can create a vocabulary, as long as you follow the rules
  - This is a function of the use of RDF
  - Parsers today are either RDF parsers or tied to the specific application (e.g. WAP Forum UAPROF)

ERICSSON

We will have lots of device types in the future - and lots of display formats

XHTML

Java



HTML

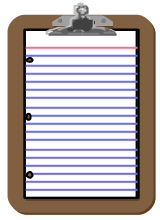
Vectormaps or bitmaps

PDF

Multimedia

Voice

WML



CC/PP is the enabler for optimized presentation - and more

Why did you not do a vocabulary?

- We could not agree on the definition of a web browsing device
  - So we provided some examples
  - Intended only as that
  - Then, of course, there are the structural attributes
- Vocabularies are about much more than display
  - Preferences, Context information….
  - Vocabularies have to be agreed by the constituents
- A vocabulary where the structure allows for filtering and/or comparison is probably more efficient than a dedicated filtering algorithm
  - This would be an interesting research topic
  - Privacy, security, etc would be handled by the framework

Extending the reach of CC/PP

- Windows Registry to CC/PP Converter

CC/PP can carry any contextualization information

- What is context?
  - My terminal
  - My preferences for the use of the terminal (e.g. r380 lid open or closed)
  - My environment
    - My position
    - The weather
    - The time
    - My agenda
  - What is context is up to you - anyone can define their own vocabularies

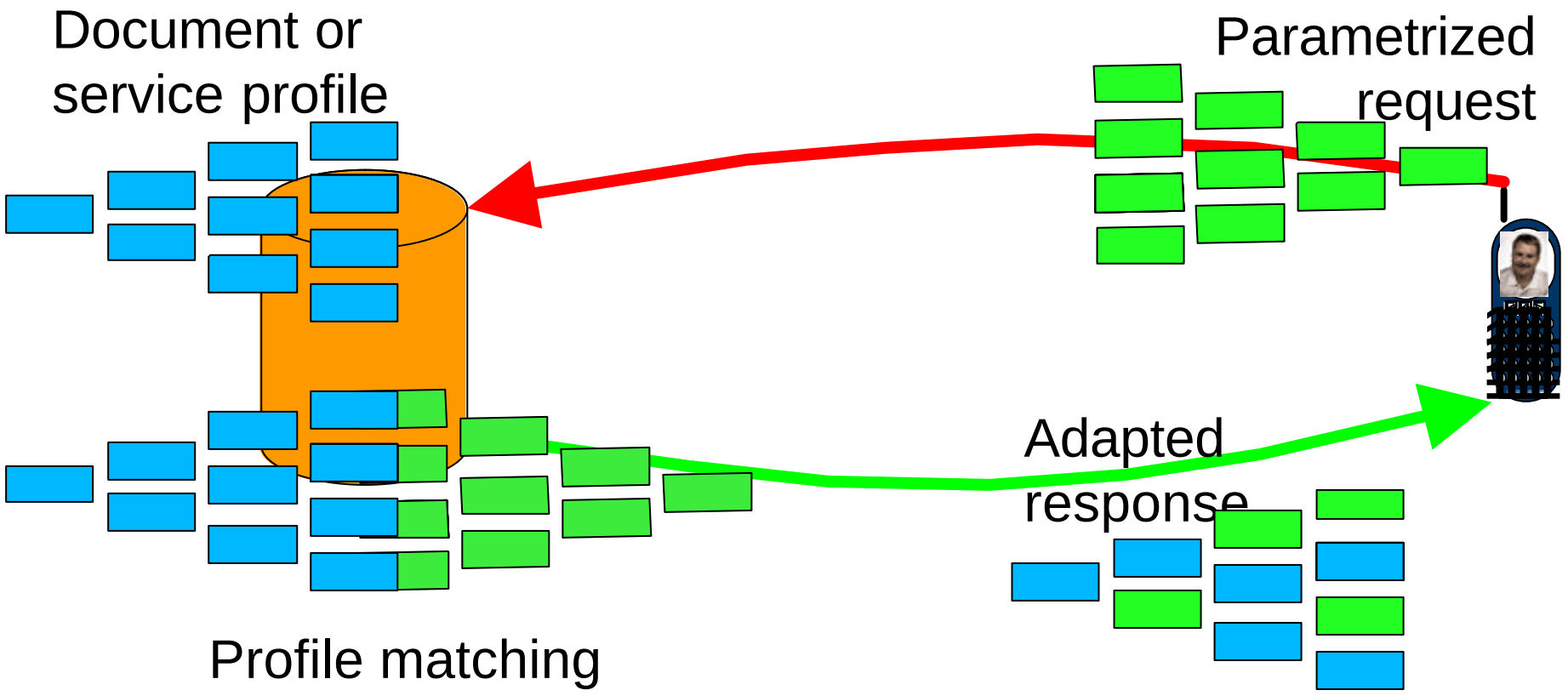
An aside: The ignorance in the industry is surprising

- Web designers think that web pages behave like printed pages
  - The web has always been a best effort medium
  - The user was always able to reset the preferences
  - The intelligent webmaster makes sure the adaption is graceful, instead of unenforceable
  - How do you handle users who are not using visual browsers if you design pages with fixed pixel widths?
- But the truth is: Users want content, not design. And this is especially true for wireless services.

Who needs to make inferences based on profiles?

- Anyone who wants to generate personalized content
  - e.g. providers of wireless information services
  - e.g. web sites that run database-backed systems
- Anyone who does not want their web pages to look like shit on other computers than where they were made
  - i.e. any webmaster who understands how the technology works

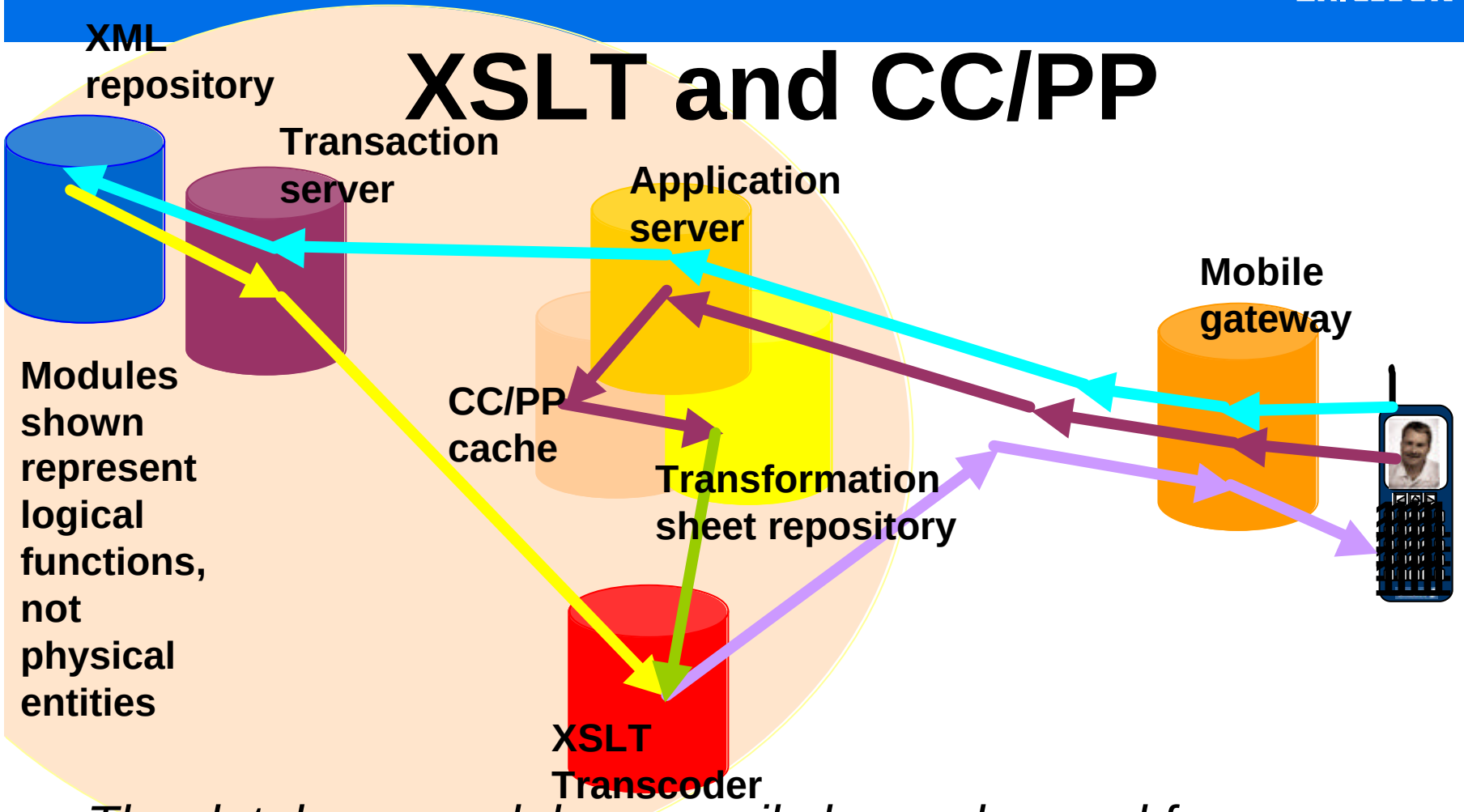
Metadata profile matching: Client and server



The technology is an enabler for future systems

ERICSSON

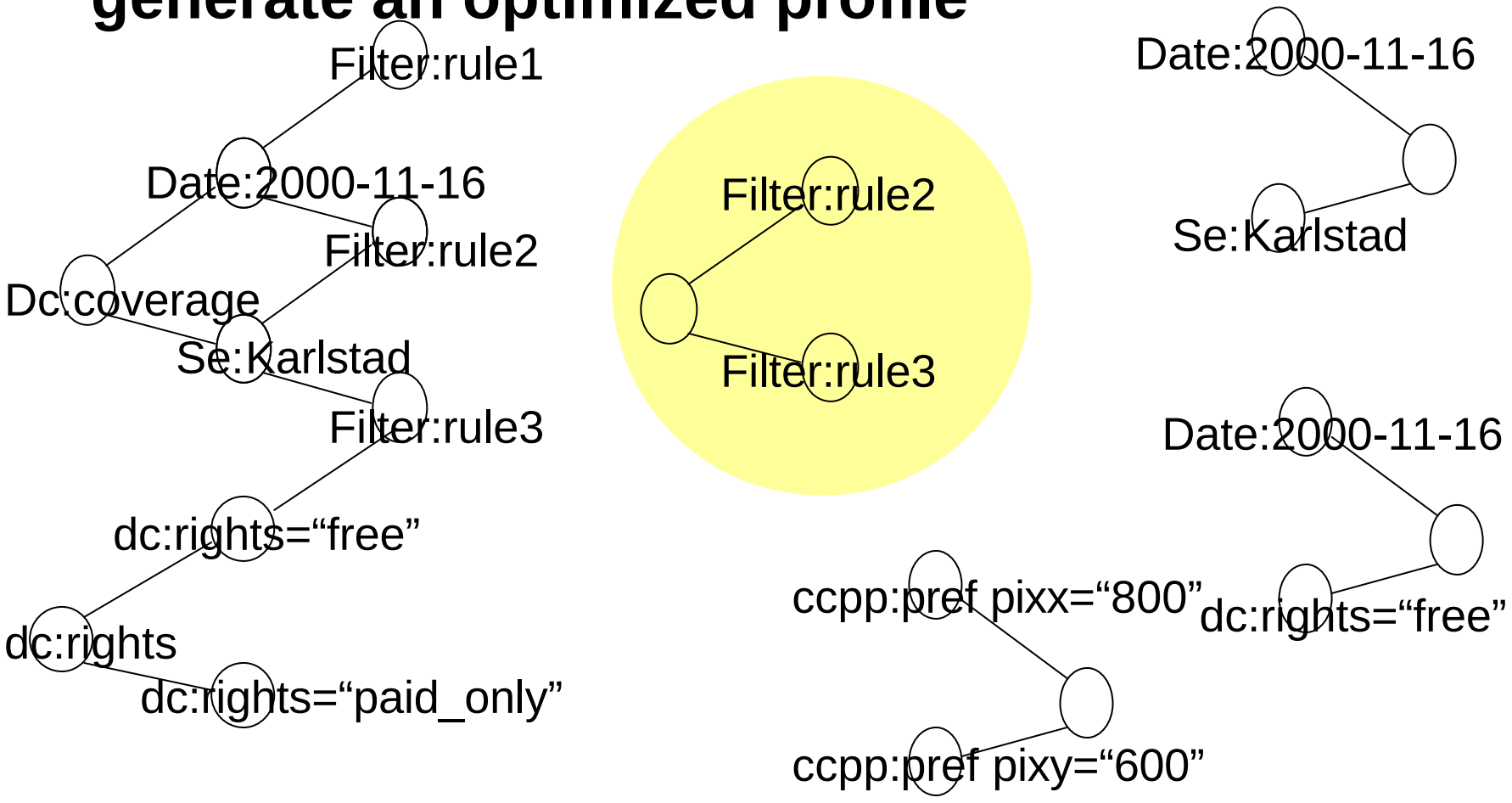
XSLT and CC/PP



- *The database model can easily be exchanged for a standardized model*
- *This works for XML content - it can not handle the errors in HTML (60 % of all HTML deployed has errors!)*

ERICSSON

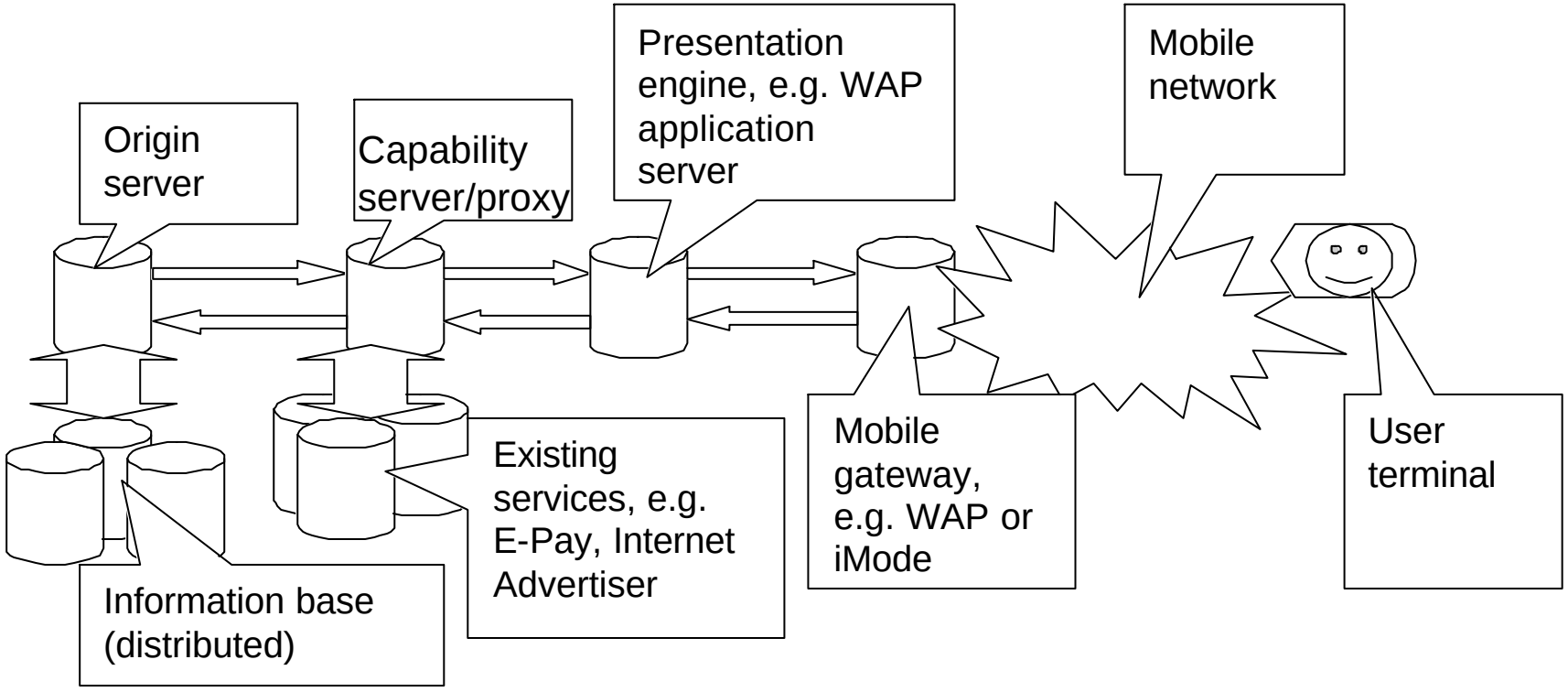
If content generation or selection is based on profiles, many profiles can be matched to generate an optimized profile



Generic filtering heuristics

- You could get by with only one attribute
  - Remove
  - i.e. the filtering starts with a full document, removes everything that is marked up
- But if you want to do more to the content, you need more heuristics
  - Add to summary
  - If attribute is
    - e.g. if vendor=FunFon
- Further work needed
  - What are generic heuristics for content filtering?

The future information system: Interacting XML applications



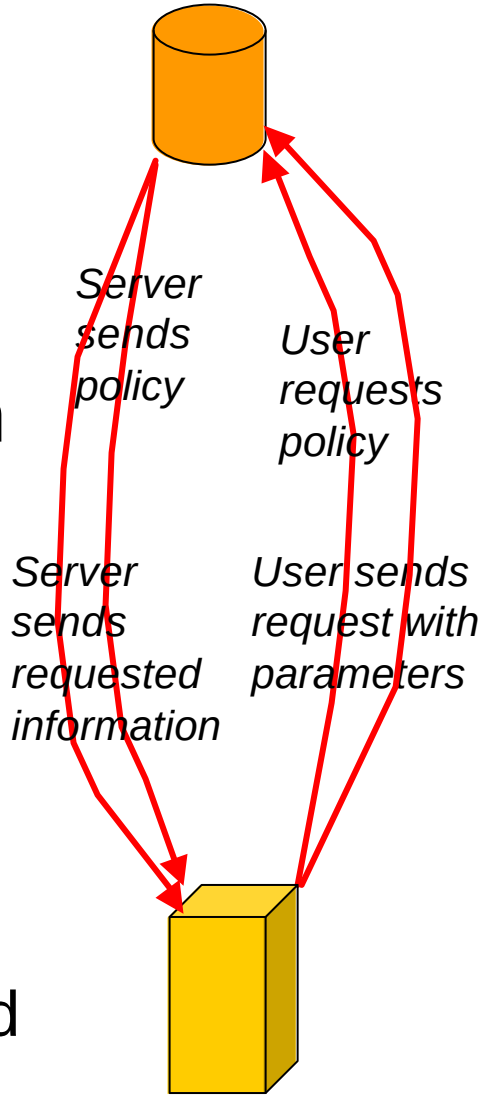
What are the profile problems?

- Matching algorithms
  - Database matching works well in a single-node environment, but maybe not in multinode environments
- Communicating and exchanging profiles
  - Wich protocol: SOAP, BEEP, i-CAP, HTTP? And how?
  - In a trusted way?
- Profile filtering heuristics
- Occlusion of part of profiles, and policies for this
  - For instance, using P3P to keep part of a profile private
- Formats (frameworks) and vocabularies
  - Not only do we need to define them, we also need to publish them, preferrably as standards
- Education, education, education

ERICSSON

Privacy management and Contractual management

- These aspects go hand in hand with interchange of profiles
- Only workable solution so far: P3P from the W3C
  - The user gets the ability to automatically deny the release of parts of his profile
- But still, there are lots of things unclear
  - Will P3P be deployed?
  - How does it work with CC/PP?
  - How does it work with other profile formats?
- The problem can be handled using out-of-band negotiated contracts



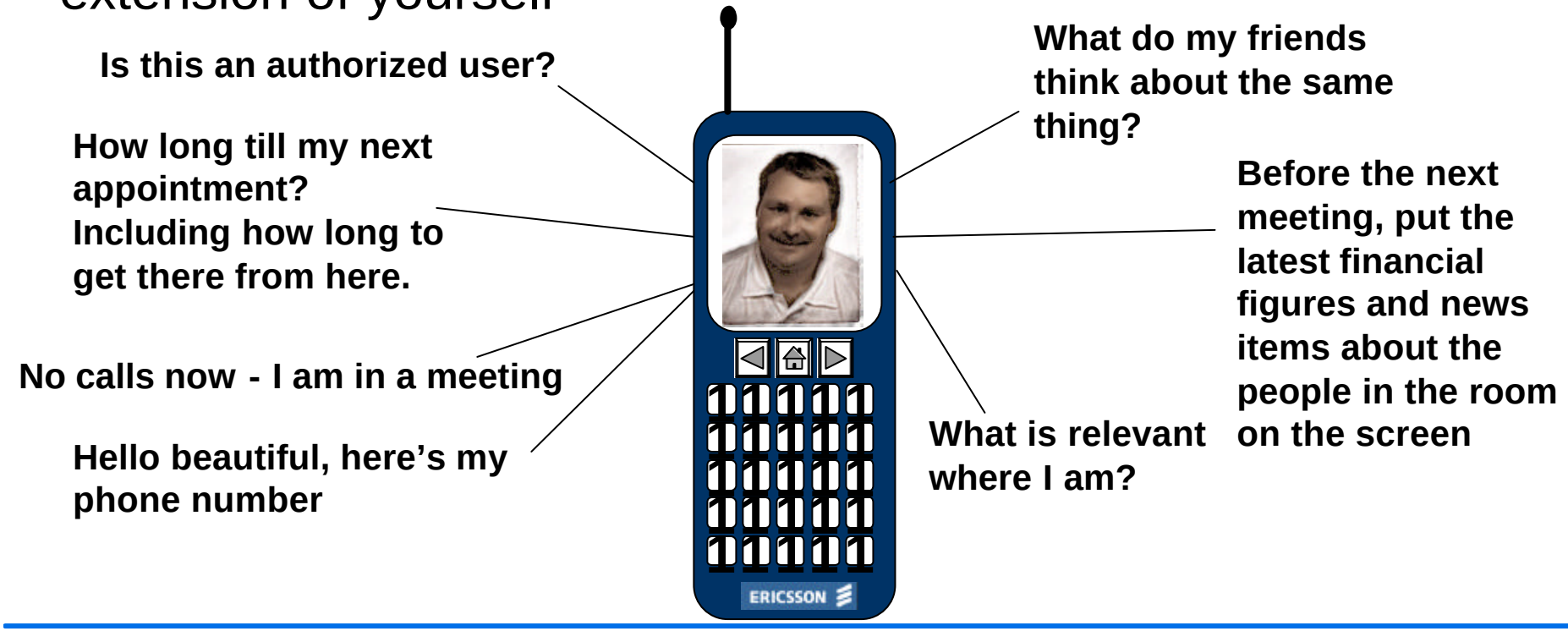
Interchange of profile information (i.e. protocols)

- If profile information is to be useful to more than one party, it has to be interchanged between parties
- This implies the use of a ***standard protocol*** and a ***standard data format***
- HTTP accept-profile header under work in the IETF (semantics reminescent of the user agent header)
- But there is a bunch of consortia and working groups
  - SOAP (Microsoft et al)
  - BEEP (IETF)
  - XML Protocol (W3C)

ERICSSON

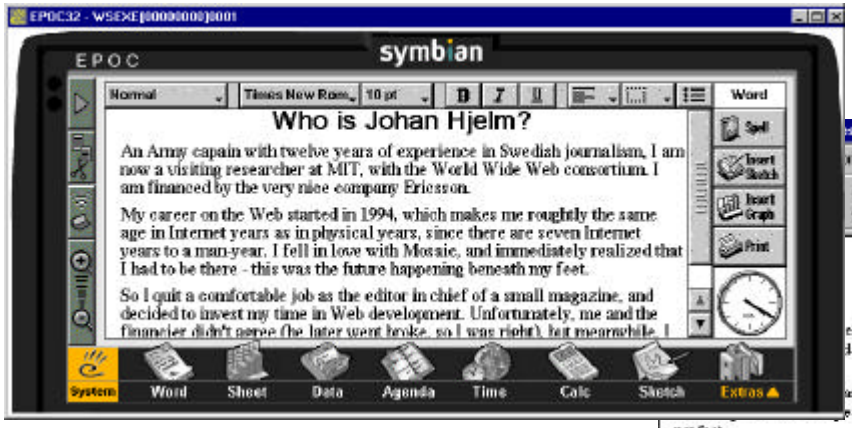
Beyond Presentation

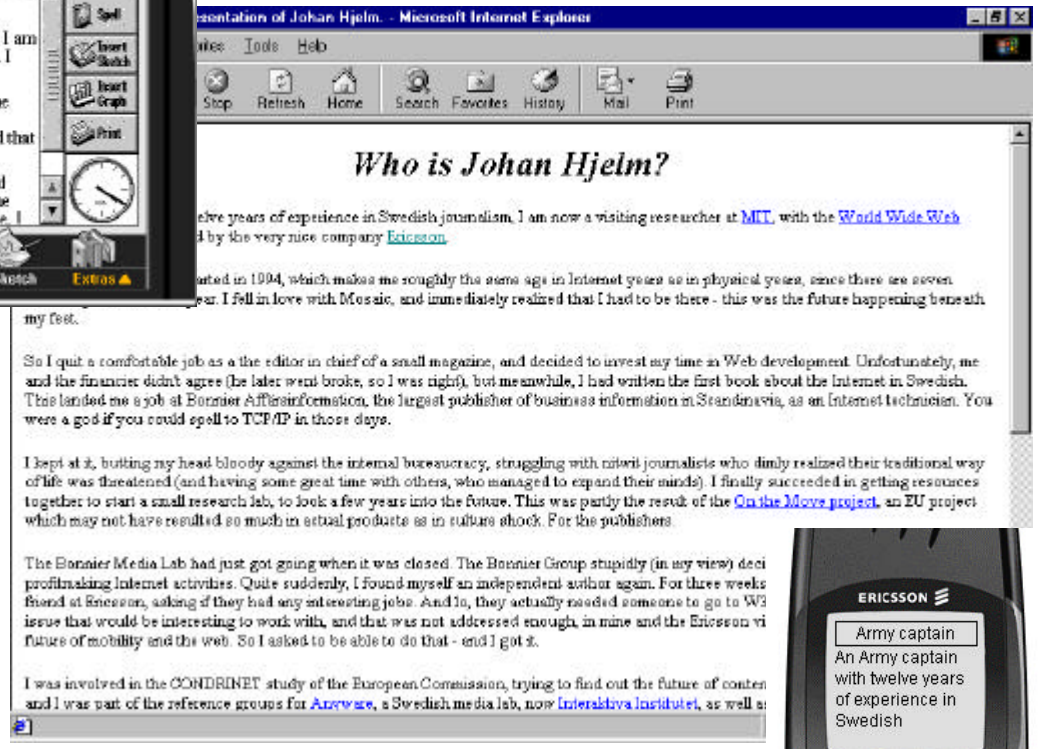
- The keyword is contextualized computing
- Your computer is “aware” of you and your environment, and adapts to it
- This means that it will become extremely personal, an extension of yourself



ERICSSON

Transformations must be more than transformation of presentation





The same content can be adapted to different formats. But you can not just take the content and display it on the different devices. You have to adapt it to the presentation format and the dialog model.

Unique to me

Credit card number

Taste in music

Language

Position

Temperature

Wheel rotation

Screen size

Shared with others

Context-dependent information, in Schilits definition, can be further broken down into two aspects: Environmental adaption, and personalization.

Environmental adaption was not very interesting until the mobile came along - because up till then, your environment did not change.

Where the border between personalization and environmental adaption runs is unclear - but the evidence is that it includes position information.

The Infinite Number of Profiles

Object (Device) Descriptors:

- CC/PP - UAPROF
- SyncML
- DevMan
- SDP
- (…)

Object (User) Descriptors

- CPEX
- MS Passport
- 3GPP User Profiles

Object (Document) Descriptors

- OGIS GML
- Dublin Core

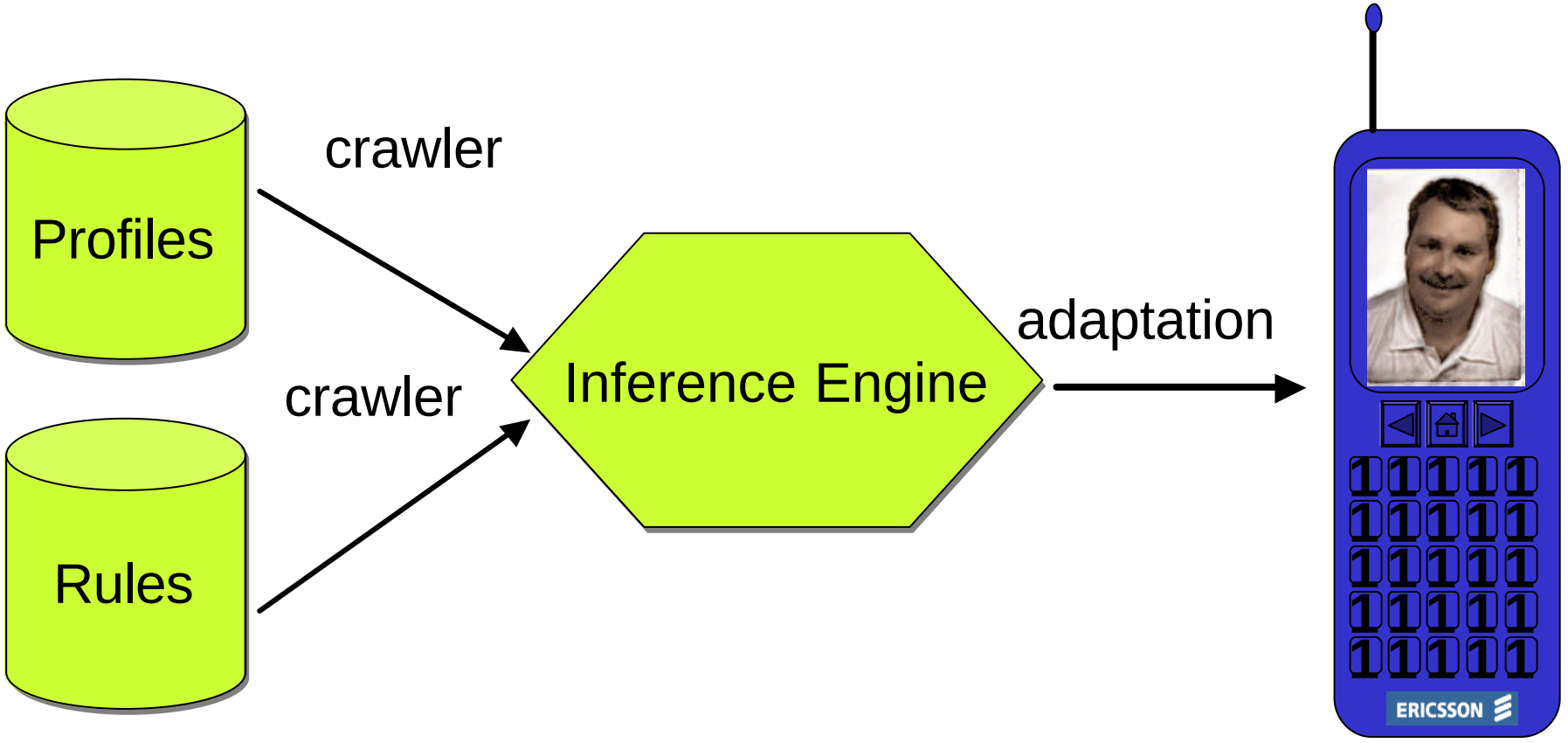
Object (Service) Descriptors

- UDDI
- WSDL
- XML Schema

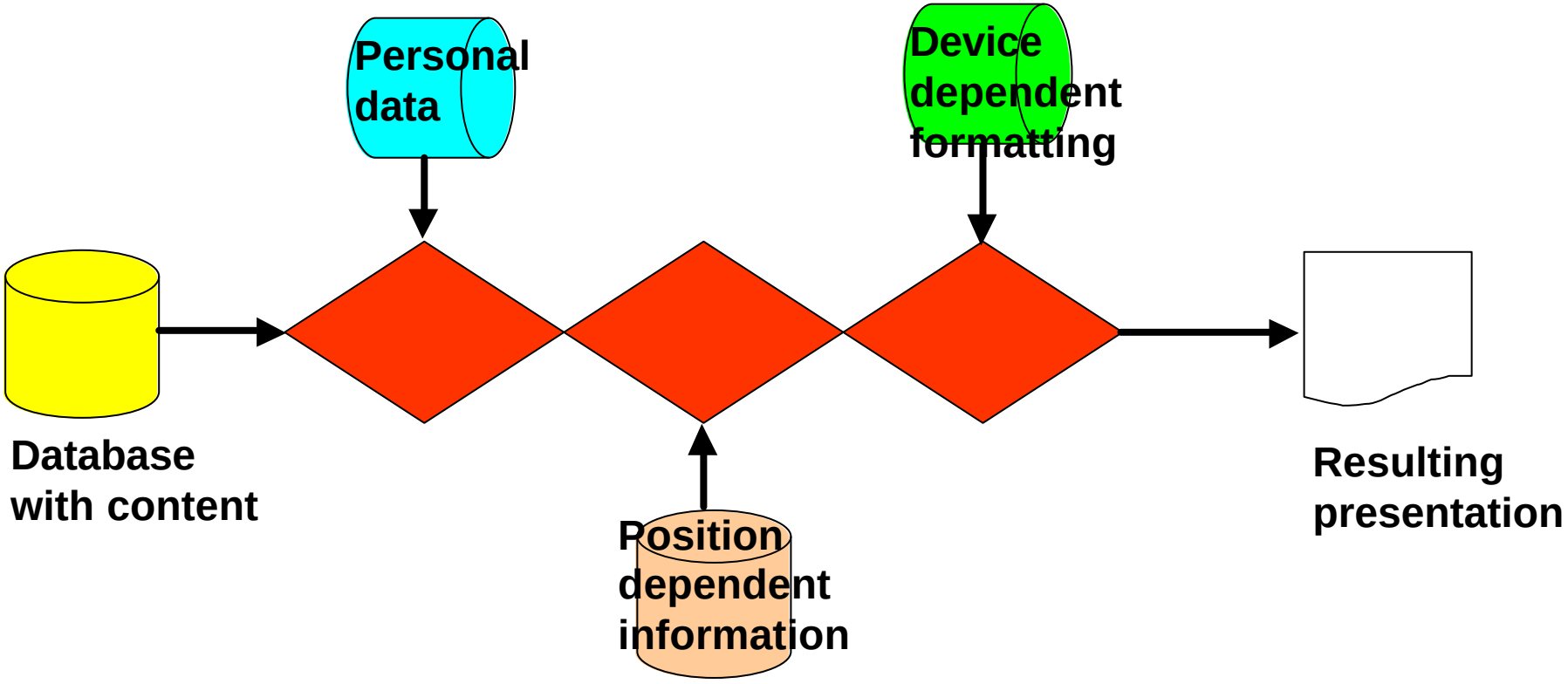
The taxonomy is still a research topic…

All these can be expressed in an ontology language, such as DAML+OIL (where the datatype issue has been addressed)

Facts & Rules


Profiles
crawler
crawler
Rules
Inference Engine
adaptation
ERICSSON

Personalization in multiple dimensions



- Different position, personal data will mean different content
- Different presentation preferences will mean different presentations

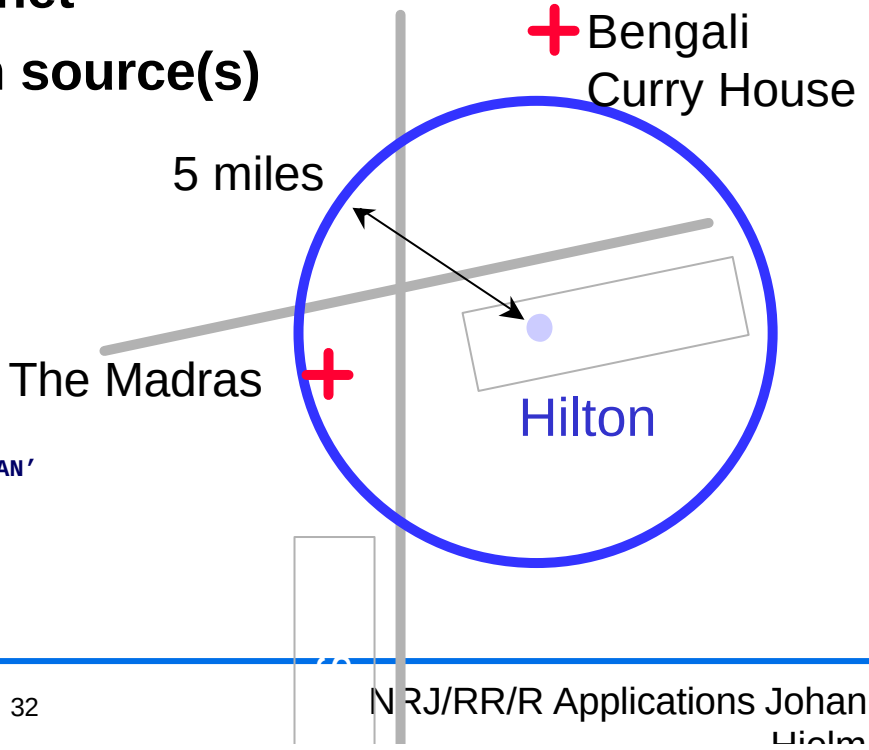
Mobile information is crucial in personalization

- Position-dependent information
  - How to avoid becoming annoying?
  - Creating buy-in among users
- Billing and charging using the telecom bill
  - Unexpected credit risk for operators
  - 3GPP user profiles
- Optimal presentation
  - Depends on preferences, not just capabilities
  - Delivery context

Introducing User Location

“Find all indian restaurants within 5 miles of my hotel”

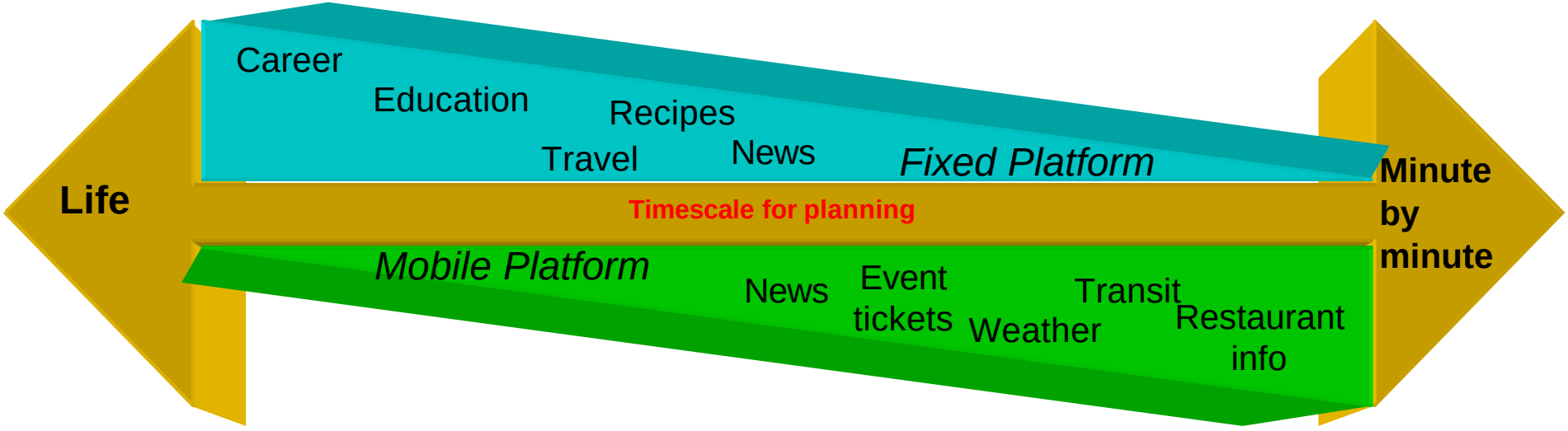
- **Post query**
  - **Include position - or -**
  - **Retrieve position from network or device**
- **Make query in database or Internet**
- **Compile result from information source(s)**
- **Present result to user**




```
SQL> SELECT R.Restaurant_name
  2>   FROM Restaurants R,
  3>        Hotels H
  4>  WHERE R.Restaurant_type = 'INDIAN'
  5>    AND H.Hotel_name = 'HILTON'
  6>    AND SDO_WITHIN_DISTANCE(
  6>         H.Location, R.Location,
  7>         'distance=5') = 'TRUE';
```

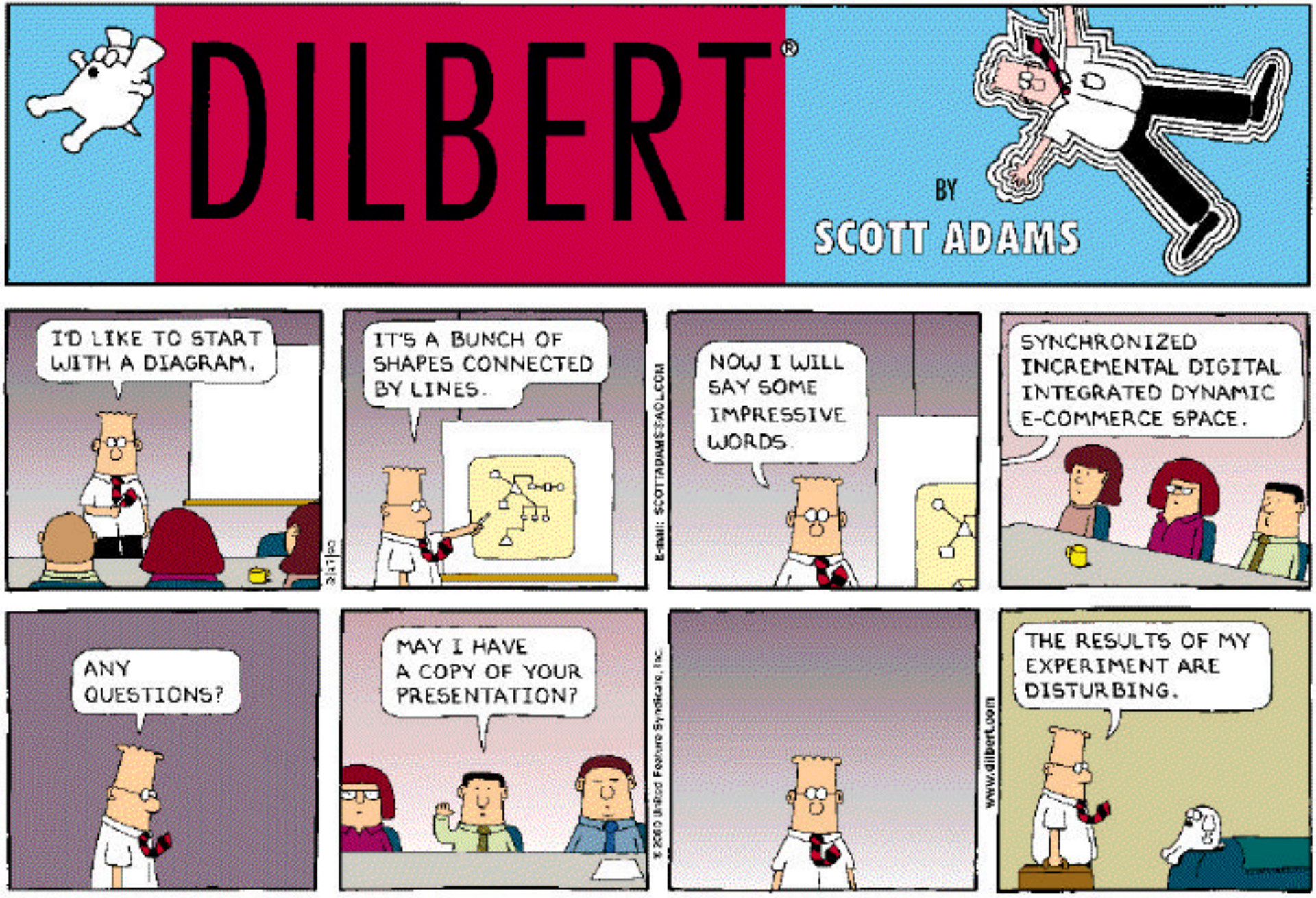
The Platform Depends on The Task

- **Mobile content is used as a base when making decisions about the near future**
- **Fixed content is used when making life decisions**
- **People choose platform depending on the timeframe**



Towards the Semantic Web

- **Transformation is the same thing as filtering**
- **The user does not have to be human**
  - **Web services can be more efficient using contextualization**
  - **For a user, you can make a totally personal presentation**
- **Location dependent services are just one aspect of environmental adaption of content**


DILBERT®
BY
SCOTT ADAMS
I'D LIKE TO START WITH A DIAGRAM.
IT'S A BUNCH OF SHAPES CONNECTED BY LINES.
E-mail: SCOTTADAMS@AOL.COM
NOW I WILL SAY SOME IMPRESSIVE WORDS.
SYNCHRONIZED INCREMENTAL DIGITAL INTEGRATED DYNAMIC E-COMMERCE SPACE.
ANY QUESTIONS?
MAY I HAVE A COPY OF YOUR PRESENTATION?
© 2000 United Feature Syndicate, Inc.
THE RESULTS OF MY EXPERIMENT ARE DISTURBING.
www.dilbert.com