Cinquième école thématique du CNRS sur les EIAH : personnalisation des EIAH
7 - 12 juillet 2007

Hôtel Grenoble-Charmeil, Saint-Quentin sur Isère, France

User Modelling for Multi User Environments

Julita Vassileva
University of Saskatchewan
Canada

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

RhôneAlpes Région

kaleidoscope
shaping the scientific evolution of Technology Enhanced Learning

Contents

- Evolution of Multi-User Environments
- User Modeling Servers
- Collaborative Recommender Systems
- Expert finding
- Decentralized User Modeling
- Purpose-Based User Modeling
- Policies for User Modeling in Online Communities
- New uses for User Modeling: adaptive rewards

UM in MUE

UM Servers

Collaborative Recommenders

Expert Finding

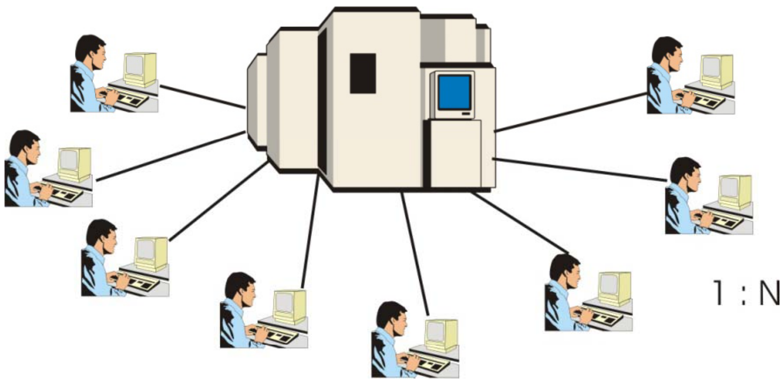
Decentralized UM

Purpose-based UM

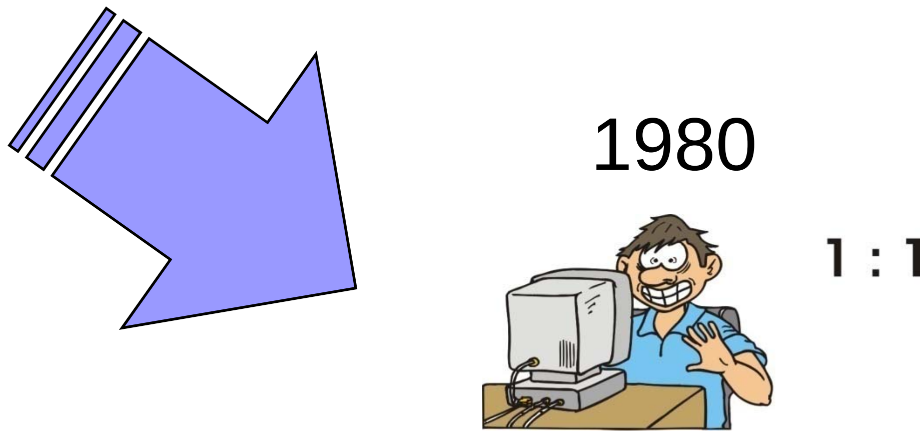
Policies in OC

Stone age and Antiquity

1970



1980



UM in MUE

Middle-ages and Present

UM Servers

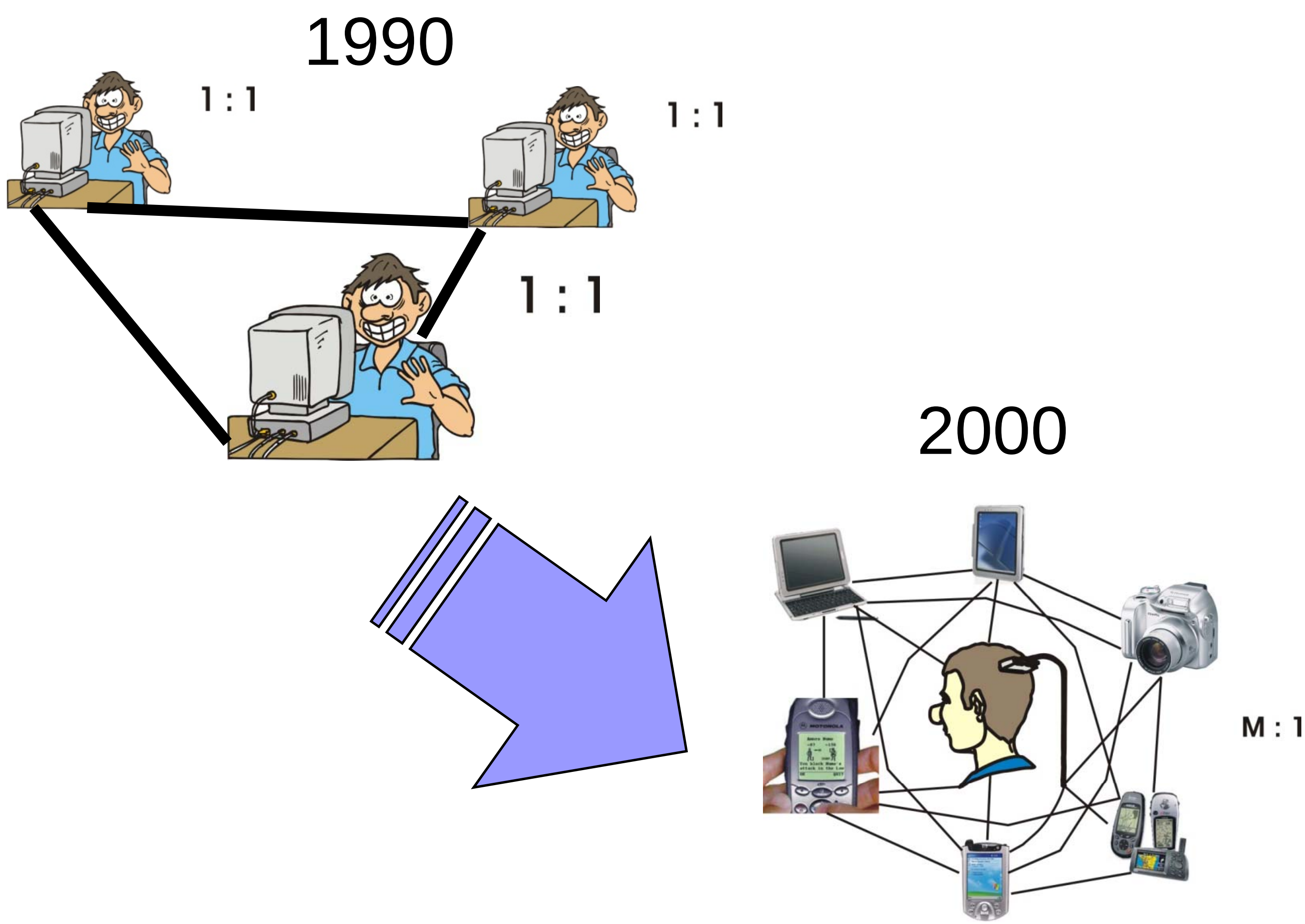
Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC



UM in MUE

Present and Future

UM Servers

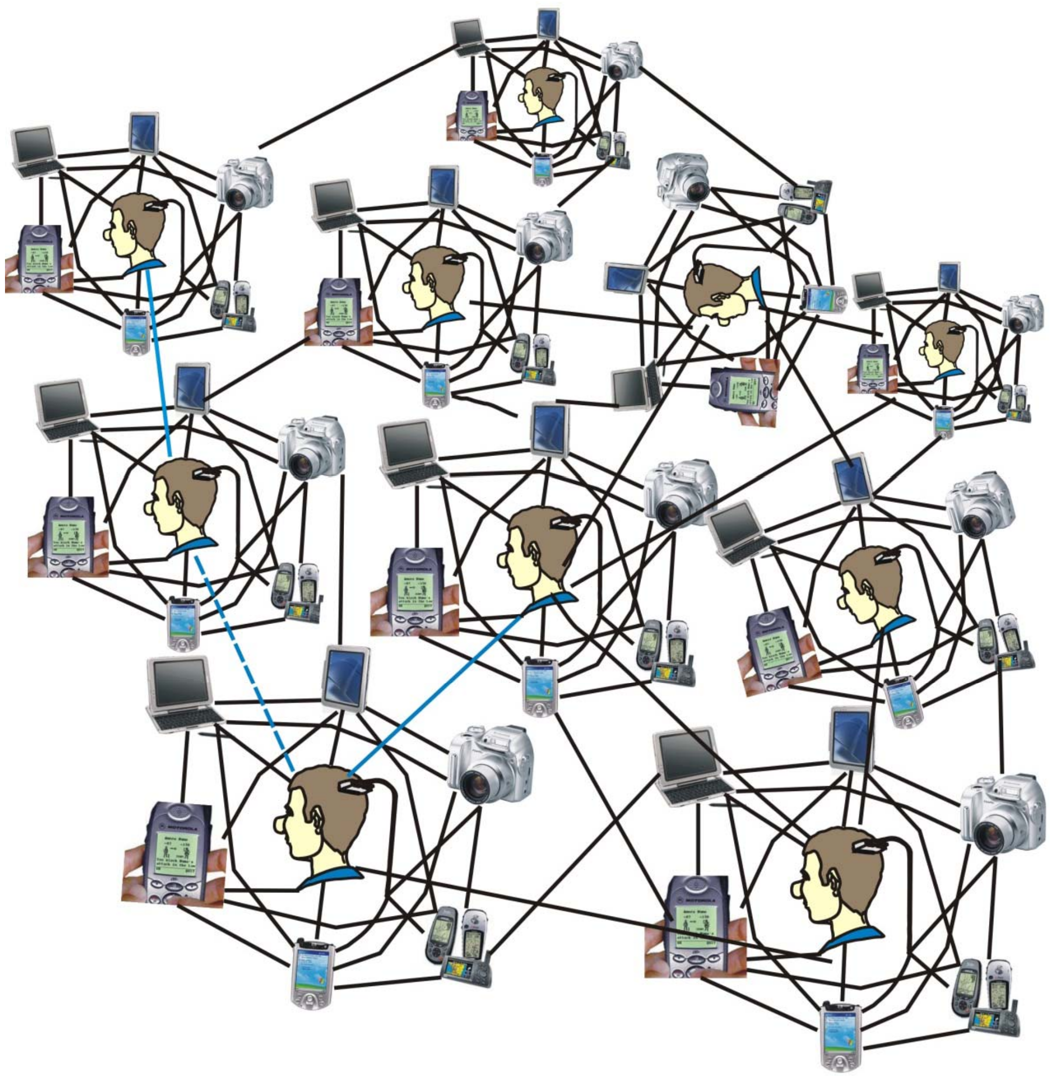
Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC



User Modeling in Multi-User Environments

WHAT ARCHITECTURES?

UM in MUE

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

Types of UM architectures for MUE

Centralized (UM Servers)

- Recommender systems
- Expert Finding systems

Decentralized

- Multi-agent systems
- Ubiquitous Computing
- Online Communities

CENTRALIZED ARCHITECTURES

UM in MUE

Centralized: UM Servers

UM Servers

Collaborative Recommenders

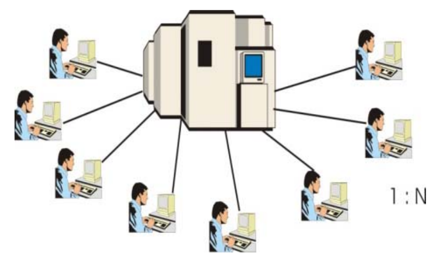
Expert Finding

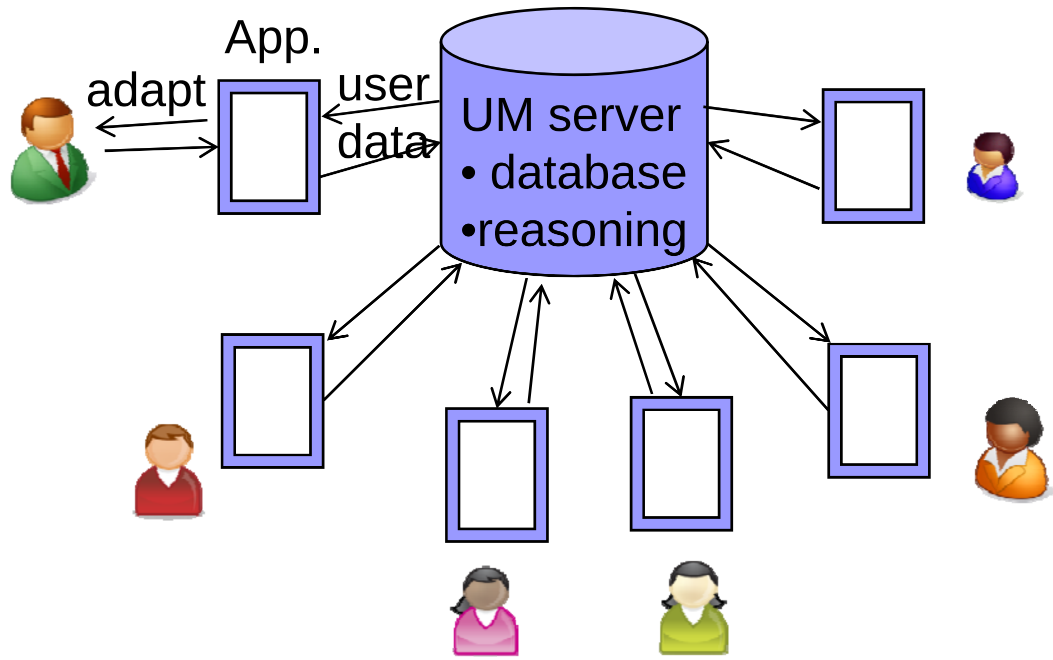
Decentralized UM

Purpose-based UM

Policies in OC

1970





UM Servers: Main Advantages

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

- Information maintained centrally or at a virtually central location can be used by many (distributed) applications; also information inferred by one application can be used by another application at the same time for a different purpose.
- Information is stored in a non-redundant manner in centralized user modelling, which more readily supports consistency and coherence of the user information.
- User modelling servers allow the application of methods and tools for security, identification, authentication, and access control.
- Information about user groups can be obtained more easily, as well as group user information can be added to the user models.

Kobsa, A. (2001): Generic User Modeling Systems. *User Modeling and User-Adapted Interaction* 11(1-2), 49-63.

Kobsa, A. and J. Fink (2006): An LDAP-Based User Modeling Server and its Evaluation. *User Modeling and User-Adapted Interaction: The Journal of Personalization Research* 16(2), 129-169.

UM Servers – More Advantages

- Backed by a database
  - Efficient
  - Secure
  - Scalable
  - Standard technology
- Hosted at one place
  - Legal responsibility - regulated by privacy legislation
  - Allows collecting data about large user population
    - Data-mining
    - Using demographic data for adaptation
    - “Cross-fertilization” – using user data across different subscriber applications
  - Allows more complex inferences, consistency maintenance, more raw data available → better quality of deduced data

Disadvantages

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

- A centralized point of failure
  - May crack under high load of requests
  - Provides a tempting target for hacking and stealing user data
- Requires a standard representation of user data
  - All subscribing applications have to adhere to standard representation / shared ontology
  - Many applications use just a fraction of the data that they can be served
- User data is out of the context in which it was harvested
  - Can be interpreted in a incorrect / harmful way in a different context; generalization not straightforward

UM in MUE

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

Representation and storage

- Stereotypes and personalized stereotypes:
  - BGP-MS (Kobsa,1990), UMT (Brainik&Tasso,1994)
- First-order logic and stereotypes
  - Tagus (Paiva&Self,1994), um (Kay,1995), Personis (Kay,2002)
- Vectors (profiles)
  - GroupLens (Breese et al, 1998, Herlocker, 1999)
- Directories (LDAP servers)
  - Standardized, security and authentication
- Usually, storage of user data was not considered important; typically read from secondary storage – file or db – at launch time or were part of the program code)

Collaborative Recommendation

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

- Example: MovieLens – a system for recommending movies
  - Users first rate a set of 20 movies given by the system
  - System computes similarity between user ratings
  - If two users have rated similarly a number of movies the system predicts that movies that were rated positively by one of the users, but not seen by the other one will be rated positively by the other user too.
  - It recommends these movies to the other user

  Amazon recommendations – people who have bought similar things in the past are more likely to continue buying similar things.
- Advantages:
  - Easier to get user data (clicks, buys, ratings)
  - Solid mathematical basis – Pearson correlation

UM in MUE

CR – basics

UM Servers

Collaborative Recommenders

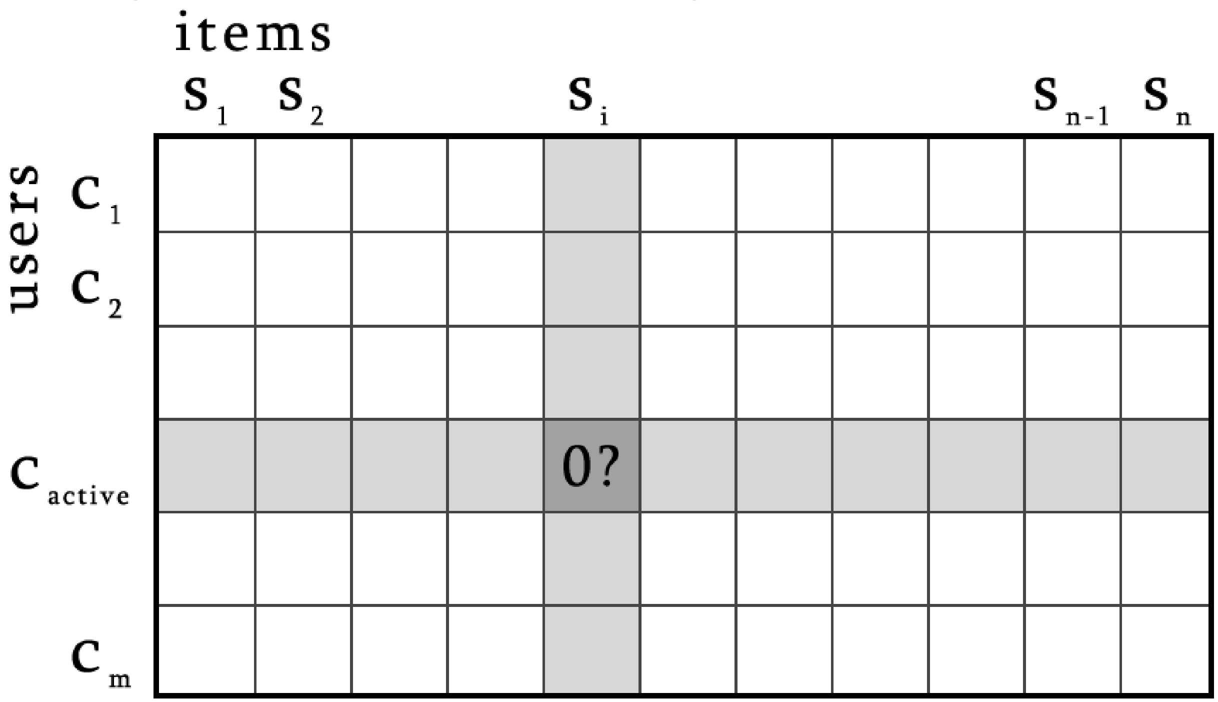
Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

- *user-item* matrix: *R*, where entry $r_{c,s}$ indicates the rating score user *c* has given item *s* .



CR – basics

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

- The goal is to predict the score of empty cells for the *active user*, the user currently requesting recommendations.
- Once scores have been predicted, a straightforward approach is to recommend the *Top-N* items that have the highest scores.

CF - basics

- Variety of recommendation strategies.
- Collaborative filtering algorithms are divided into two categories:
  - *memory-based* - uses statistical methods to derive predictions from like-minded users;
  - *model-based* - builds models of user ratings through *machine learning* algorithms (e.g. Bayesian networks)

Memory-based CF algorithms

- If the active user c has not rated item s, but the recommender system can find similar or correlated users c' (i.e. a set of *neighbours* $\hat{C}$) who have, then a rating score can be predicted using:

$$r_{c,s} = \bar{r}_c + k \sum_{c' \in \hat{C}} sim(c,c') \times (r_{c',s} - \bar{r}_{c'})$$

- Where $sim(c, c')$ is the Pearson correlation coefficient (next slide)
- $k = 1 / \sum_{c' \in \hat{C}} sim(c,c')$ is a normalizing coefficient

Similarity between 2 users

$$sim(c,c') = \frac{\sum_{s \in S_{cc'}} (r_{c,s} - \bar{r}_c)(r_{c',s} - \bar{r}_{c'})}{\sqrt{\sum_{s \in S_{cc'}} (r_{c,s} - \bar{r}_c)^2 \sum_{s \in S_{cc'}} (r_{c',s} - \bar{r}_{c'})^2}}$$

- Users tend to use ratings scales differently.
  - For example, on 5-scale rating, the active user may seldom rate 1 or 5 while a neighbour *only* rates 1 and 5.
- Therefore, the average rating of the active user and current neighbour ($\bar{r}_c$ and $\bar{r}_{c'}$, respectively) are used to smooth out this inconsistency.
- $S_{cc'}$ is the set of items both users have rated in common.

Limitations

- The lack of ratings presents well-known challenges: if rating information does not exist, then predictions will be inaccurate or undefined
  - new users must rate some number of items before recommendations can be made,
  - new items must receive ratings before being recommended to users
- The computational complexity of memory-based algorithms requires offline calculations (e.g. at night when resource demand is low), and they do not scale well under millions of items and users

Adomavicius, G. & Tuzhilin, A. 2005, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions", IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 6, pp. 734-749.

Expert Finding

- Example: the PHelpS system (Prison Help System)
  - The system discovers the area of expertise/knowledge of each user (prison guard/ employee) by monitoring how many times they completed certain tasks
  - User model – a overlay of a domain task model, together with information about location, time-zone, telephone, language (En/Fr)
  - When a user asks for help with a certain task (filling out a form), the system finds a suitable user (location, time zone, language) who has completed this task a number of times before and shows her phone number.
- Many such systems since 1986…

Expert Finding

- Modeling the expertise of experts
  - Latent Semantic Analysis of documents, e.g. e-mails (www.tacit.com), newsgroup postings,
  - Analysis of web-browsing patterns (systems called Memoir, Expert Browser),
  - Social network analysis from emails, paper co-authorships etc.
- Matchmaking
  - Keyword matching
  - Similarity matching (e.g. correlation like in CF, or vector-space based methods)
  - Domain-model enhanced (with concept- or task-relationships that imply knowing sub-concepts or tasks)

UM in MUE

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

DECENTRALIZED ARCHITECTURES

Multi-Agent Expert Finder System

- Each user represented by an agent who keeps the UM of the user
- Resources and applications are also represented by agents with application models (AM)
- User model is private, user can interact with the agent and change the UM as she wants to be perceived by others
- Agent connects with other agents and searches for good matches for the user's need
- Agent society, agent economy…

UM in MUE

UM Servers

Collaborative Recommenders

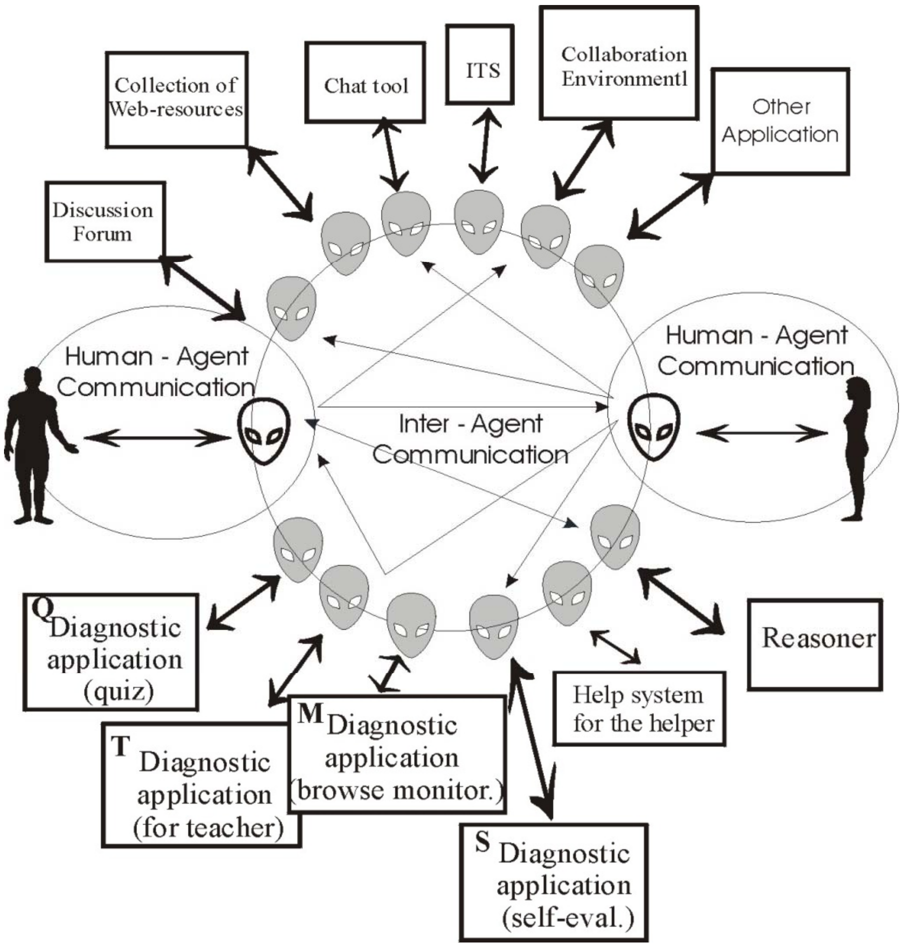
Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

I-Help Multi-Agent Architecture



Personal / Application Agents:

- **goal-driven, social**
- **maintain user/application models**
  - calling upon diagnostic applications
- **negotiate** on behalf of users / applications about price

User Model:

- **resource repository** for the personal agent
  - knowledge
  - cognitive style
  - social model
  - relationships
  - time available
  - currency

Goal-driven social agents

- *Agents are autonomous and rational*
- Types of Agent Goals:
  - intrinsic: *maximize utility, save state, maintain resource repository, evaluate goal importance*
  - extrinsic: *adopted from user, from other users*
- Intra- Agent Process:
  - Monitor the environment
  - Calculate goal importance
  - Calculate and maximize utility (Utility function --> agent “character”)
  - Retrieve plans for goal achievement / plan
  - Execute plans
  - Delegate goals
- Inter Agent Process:
  - Agents can communicate
  - Agents can not achieve their goals alone --> need to collaborate: *delegate goals, provide services, share resources*, and to compete for resources
  - Agents offer excess resources or services in exchange for money
  - Agents negotiate to resolve conflicts
  - Agents make models of other agents

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

UM in MUE

UM Servers

Collaborative Recommenders

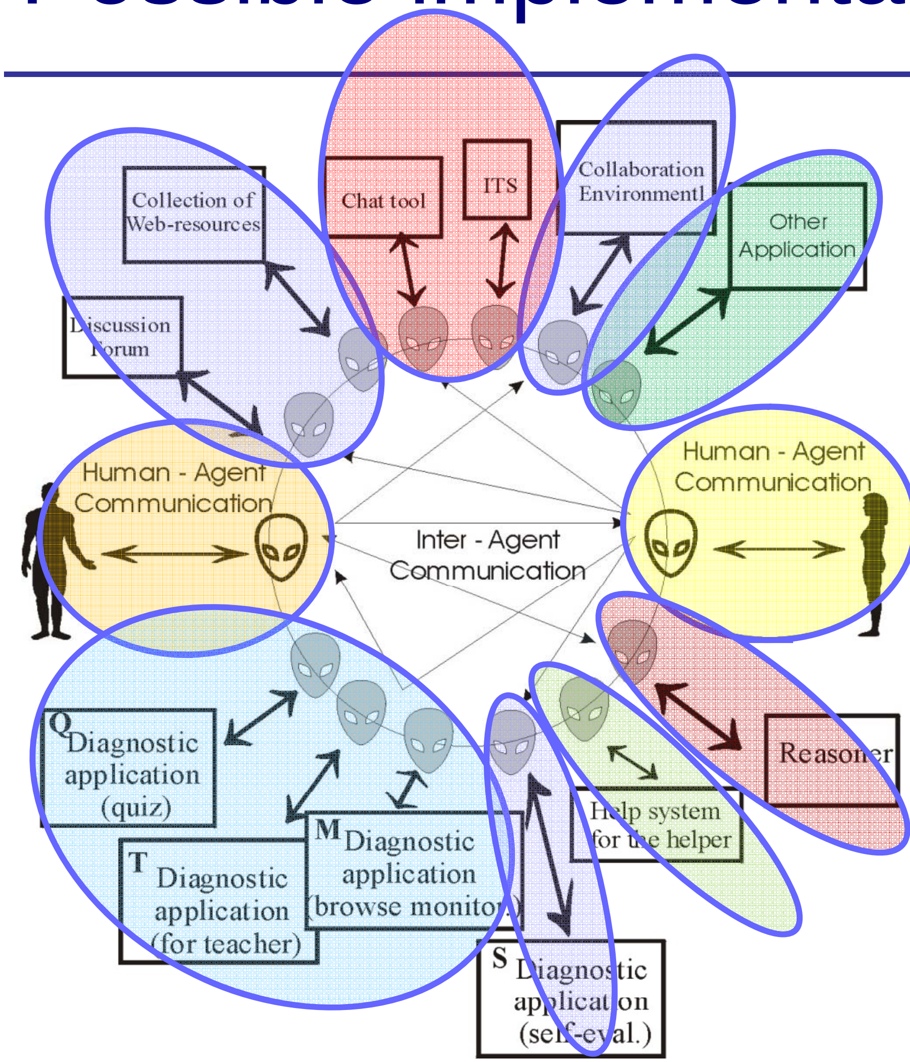
Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

Possible implementations



Physically distributed on Different machines (P2P)

Vassileva J. , J. Greer, G. McCalla, R. Deters, D. Zapata, C. Mudgal, S. Grant (1999) A Multi-Agent Approach to the Design of Peer-Help Environments, in *Proceedings of AIED'99*, Le Mans, France, July, 1999, 38-45

UM in MUE

Possible implementations

UM Servers

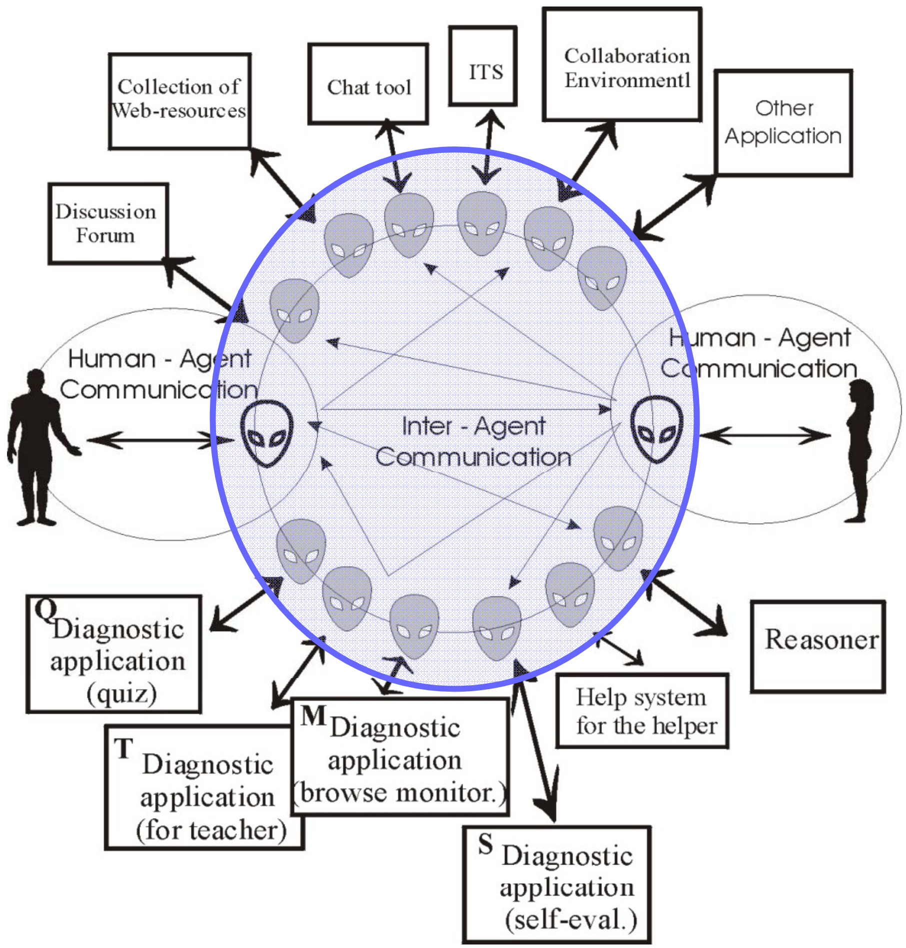
Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC



Physically centralized
- All agents reside on the same server

Why??
- to minimize communication costs

Yet, logically decentralized!
- preserved autonomy of each agent - can have its own UM representation

Agent Negotiation

- Agents negotiate on behalf of their users depending on
  - the importance of a request
  - the importance of own goals
  - the importance of saving or earning money
  - the importance and sign of the relationship
- Each agent computes a utility function
  - one for the utility of getting help
  - one for the utility of providing help
  - **price** affects utility; quality affects **price**
  - agents have limited knowledge about each other
- Decision theoretic approach
- Negotiation protocol

UM in MUE

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

UM challenge

What User Models in such a decentralized environment?

- User models AND Application models
- Agents modelling other agents in negotiation
- User model as a repository of user resources:
  - knowledge
  - social competence
  - time, money and relationships with other users
- Distributed, domain/task specific diagnostic applications
- **User information that is context dependent?**

Ubiquitous Environments and UM

- Smart appliances have appeared on the market that maintain use characteristics and adapt their functionality
- Examples:
  - Refrigerators keep track of user preferences and stored food and order out-of-stock items via the internet
  - Mobile phones preload pages that are assumed relevant
  - DVD recorders record programs deemed interesting for the viewer based oh her viewer patterns
  - Museum guides that adapt presentation and tours
  - Car-radios and navigation systems …
- User data fragments everywhere:
  - Variety of user features modeled, representations, adaptation techniques (what is adapted, how).

UM in MUE

Context is important in DUM

UM Servers

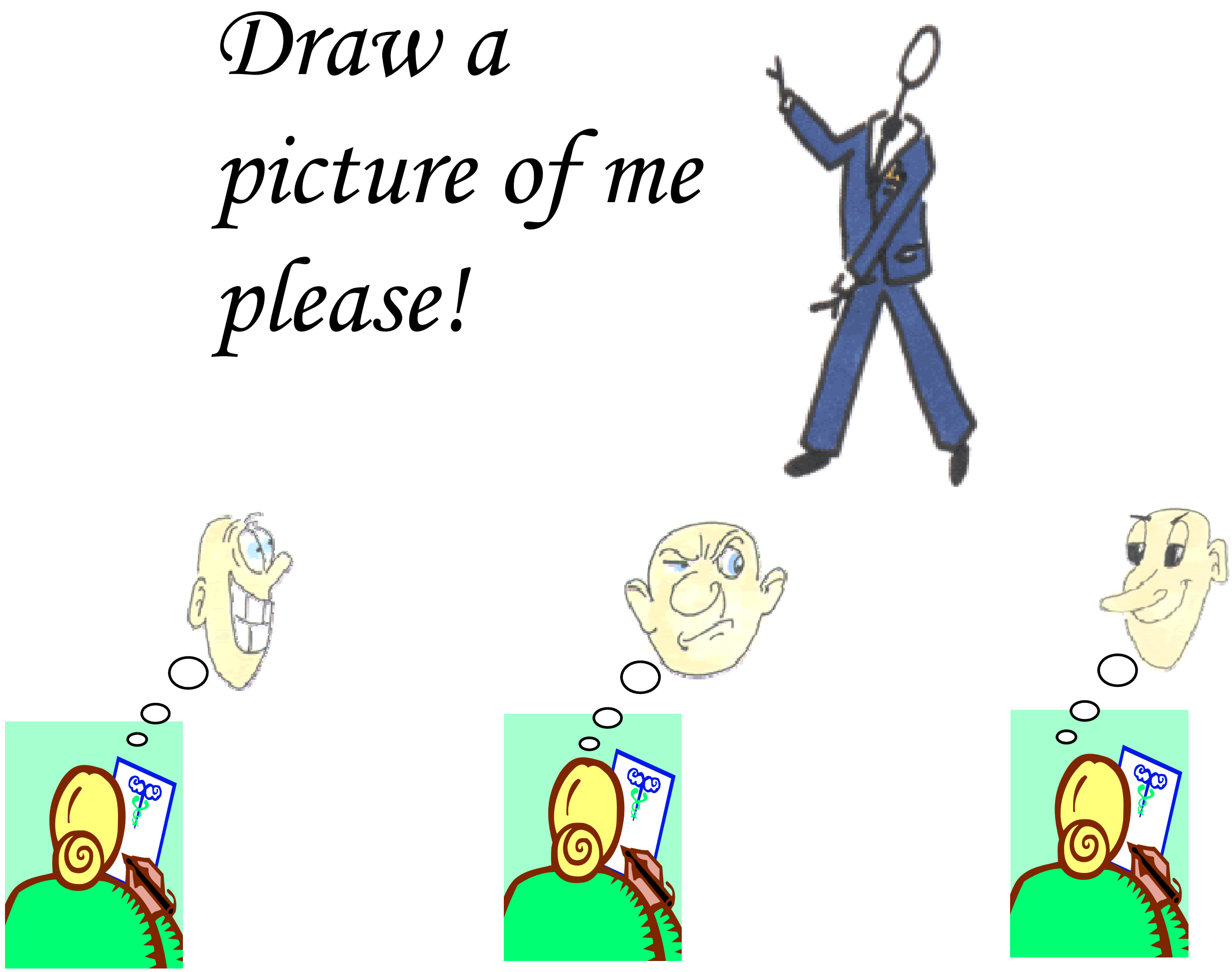
Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC



UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

Decentralized UM

- Every application / agent stores user data locally in its own representation format
  - data is lose to the context of its harvesting and use
- Applications communicate and share user data
  - only on a “need to know” basis
  - for particular purpose
- User modeling
  - *Searching, retrieving and integrating fragmented user information from diverse sources at the time when it is needed for a particular purpose.*
  - “to model” *(verb)* is a process, not a data-structure!!

Decentralized vs. Centralized UM

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

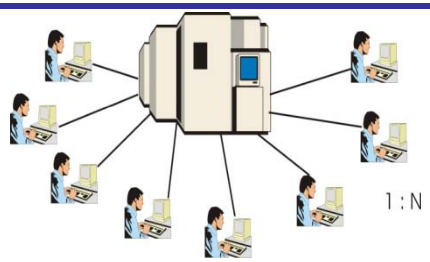
Policies in OC

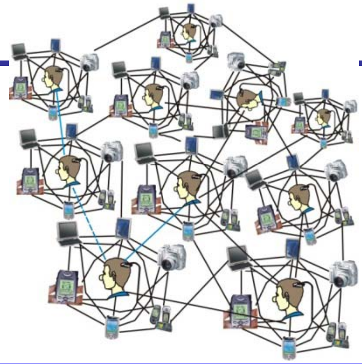
- Centralized
  - collecting at one place as much information as possible
  - about many users,
  - to make sure it is correct and consistent,
  - so that it can be used for many purposes.
- Decentralized
  - user information fragmented among many applications
  - each application models one or more users
  - inherently inconsistent (in different contexts, applications autonomous and by different designers)
  - fragments are retrieved and combined just in time for one specific purpose only

UM in MUE

Centralized vs. Decentralized UM

- UM Servers
- Collaborative Recommenders
- Expert Finding
- Decentralized UM
- Purpose-based UM
- Policies in OC





Centralized		Decentralized
Knowledge Representation	→	Modelling Process
Maintain Consistency	→	Determine Relevance
Long Term Modelling	→	Just-in-time Computing

Purposes for modeling

- A “Purpose” is like a recipe – a procedural knowledge representation construct
  - Retrieval – which are the relevant sources to get user data from
  - Interpretation – mapping information to own representation / context
  - Integration – reasoning based on the user data and possibly generating new user data
  - Adaptation – using the user data to make a decision about adaptation of interface or functionality.
- Example: Selecting new graduate students
  - Retrieve data from transcripts, letters of reference (not from his mom)
  - Interpret the marks: 6 in Bulgaria corresponds to 1 in Germany, to 92-95% in Canada
  - Integrate the interpreted data from all sources, for all considered students
  - Adaptation – generate a ranked list

Purposes

- Procedural, decentralized
- Computed just-in-time on demand
- A purpose defines the context and the processing (both user modeling and adaptation) to be done in that context
- Purposes can be designed as clichés and re-used in different situations
- Purpose libraries for designers of adaptive autonomous interconnected applications

X. Niu, G. I. McCalla, and J. Vassileva, (2004) Purpose-based Expert Finding in a Portfolio Management System. *Computational Intelligence Journal*, Vol. 20, No. 4, 548-561.

UM in MUE

UM Servers

Collaborative Recommenders

Expert Finding

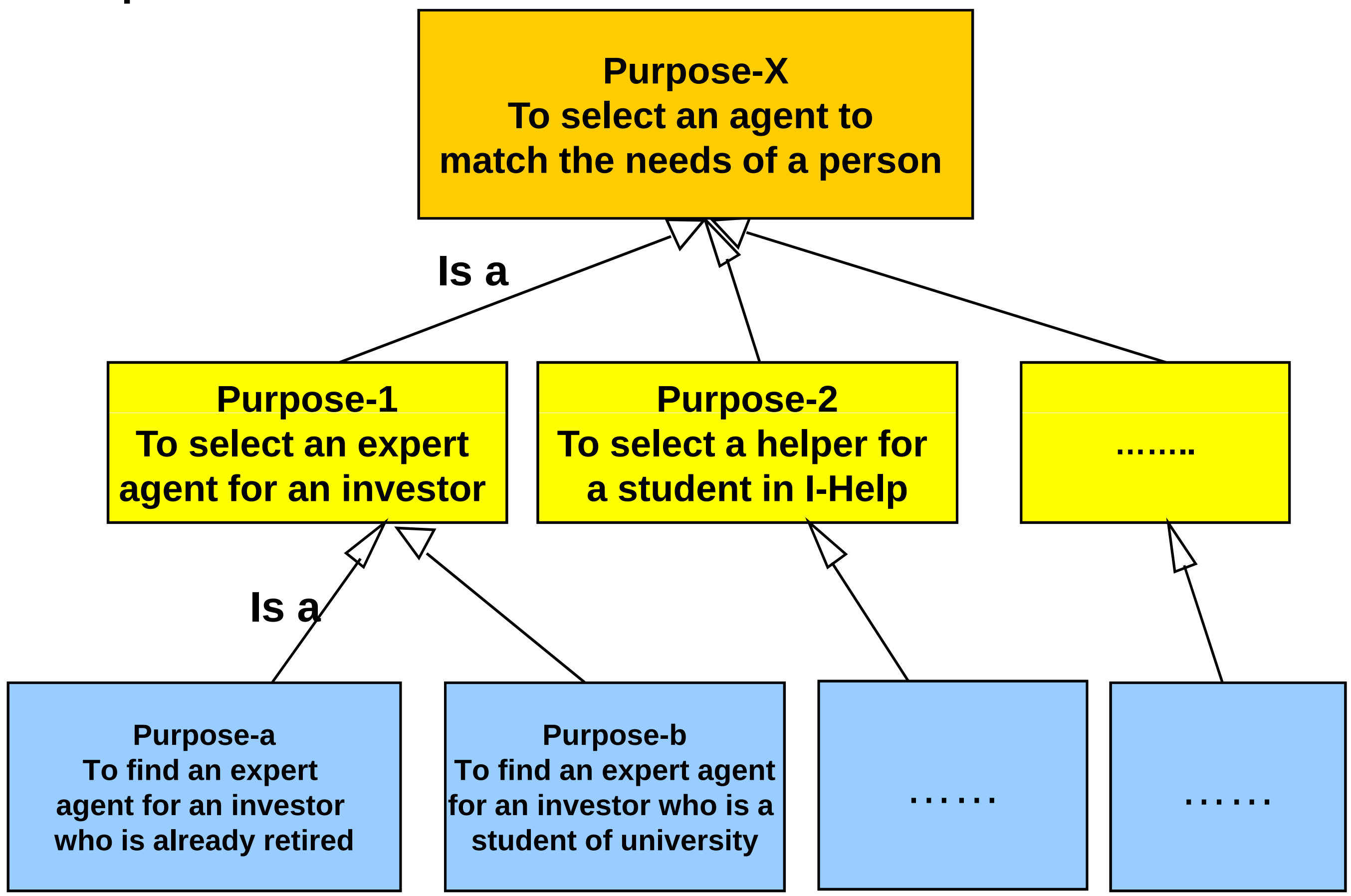
Decentralized UM

Purpose-based UM

Policies in OC

Purpose hierarchies

- Specialization



UM in MUE

UM Servers

Collaborative Recommenders

Expert Finding

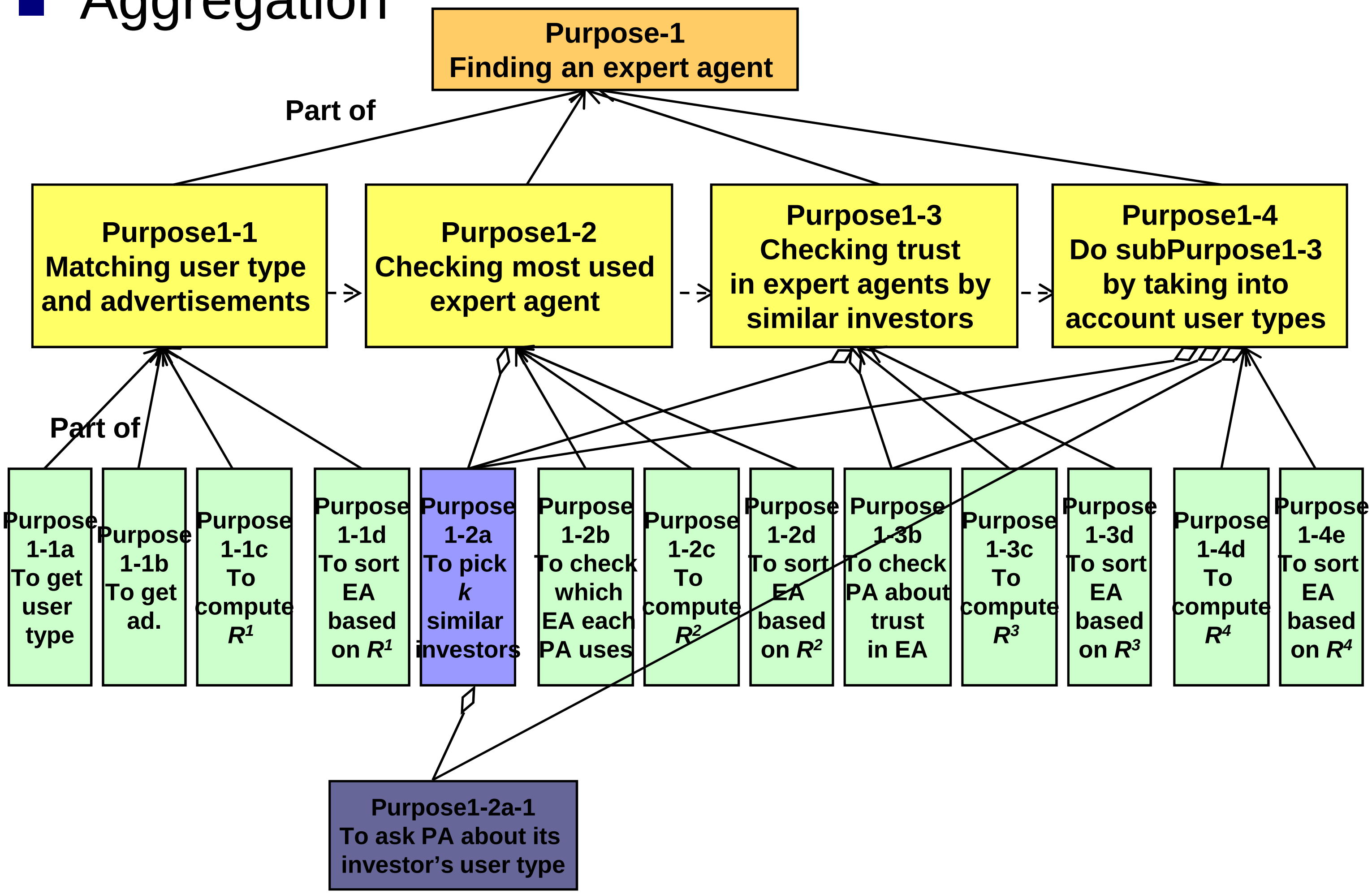
Decentralized UM

Purpose-based UM

Policies in OC

Purpose Hierarchies

- Aggregation



UM in MUE

UM Servers

Collaborative Recommenders

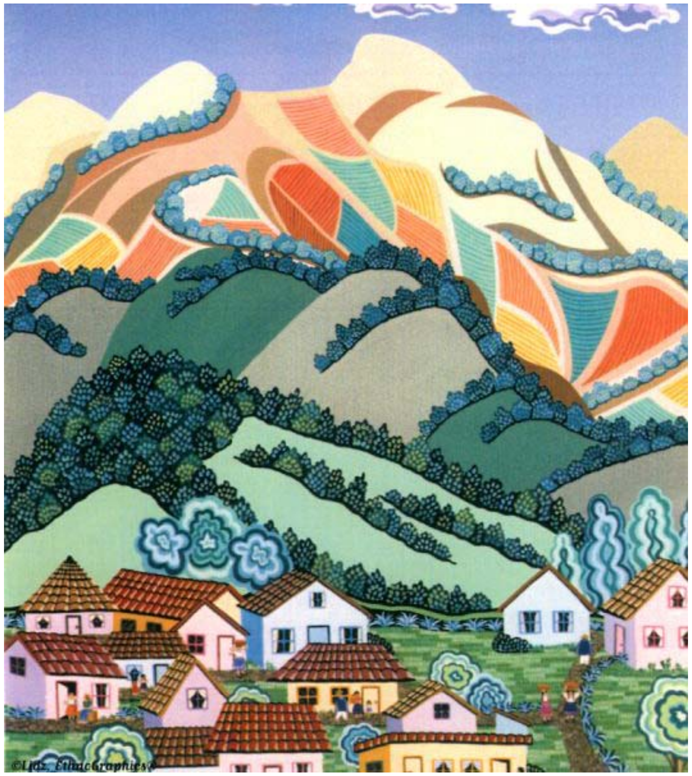
Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

Online Communities



Many communities exist

Few collaborate and share users yet, but in the future they will.

Users will be travelling seamlessly across online communities, as they travel from city to city in the real worlds.

Do they have to be strangers when they go to a new place?

How to share user data across?

- Authentication (login data)
- Profile (interests, status, friends, resources?)

UM in MUE

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

User Models in OC

- Fragments stored by logically and physically distributed nodes
- How to collaborate and share data, without giving up the autonomy
- How to update and synchronize models of users who are members of many communities

- Example:
  - European Countries have agreed to a standard format and security features of their national passports and have implemented standard policies for sharing data to facilitate travel within the EU (Schengen)

Policies in OC

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

- UM in OC are based on policies describing the role, status, and rights of each user
- Roles, status, imply rights and adaptation of the functionality and interface of the OC to the user.
- Examples:
  - "New users can not delete links" =
    If user_participation_C1< threshold
    disable "delete link" functionality.
  - "Users from community C2 are not treated as new users".
    If user_participation_C2 <> 0,
    user_participation_C1 = user_participation_c2
- The purpose-based approach can be implemented through policies
  - Transparent
  - Editable by users in certain roles (moderators)

Comtella Framework for OCs

- Every user can create a community → "owner"
- Communities can be hosted at different websites (Comtella nodes)
- Every owner defines the policies for rewarding participation (e.g. bronze, silver, gold status), the privileges with each status level, the roles that users can take (e.g. guest, member, moderator) and the rights associated with the role.
- Policies are like decision making procedures that use UM data to generate new UM data or to make an adaptation decision – enabling or disabling a particular interface feature.
- UM data can be from any community in the NW

UM in MUE

UM Servers

Collaborative Recommenders

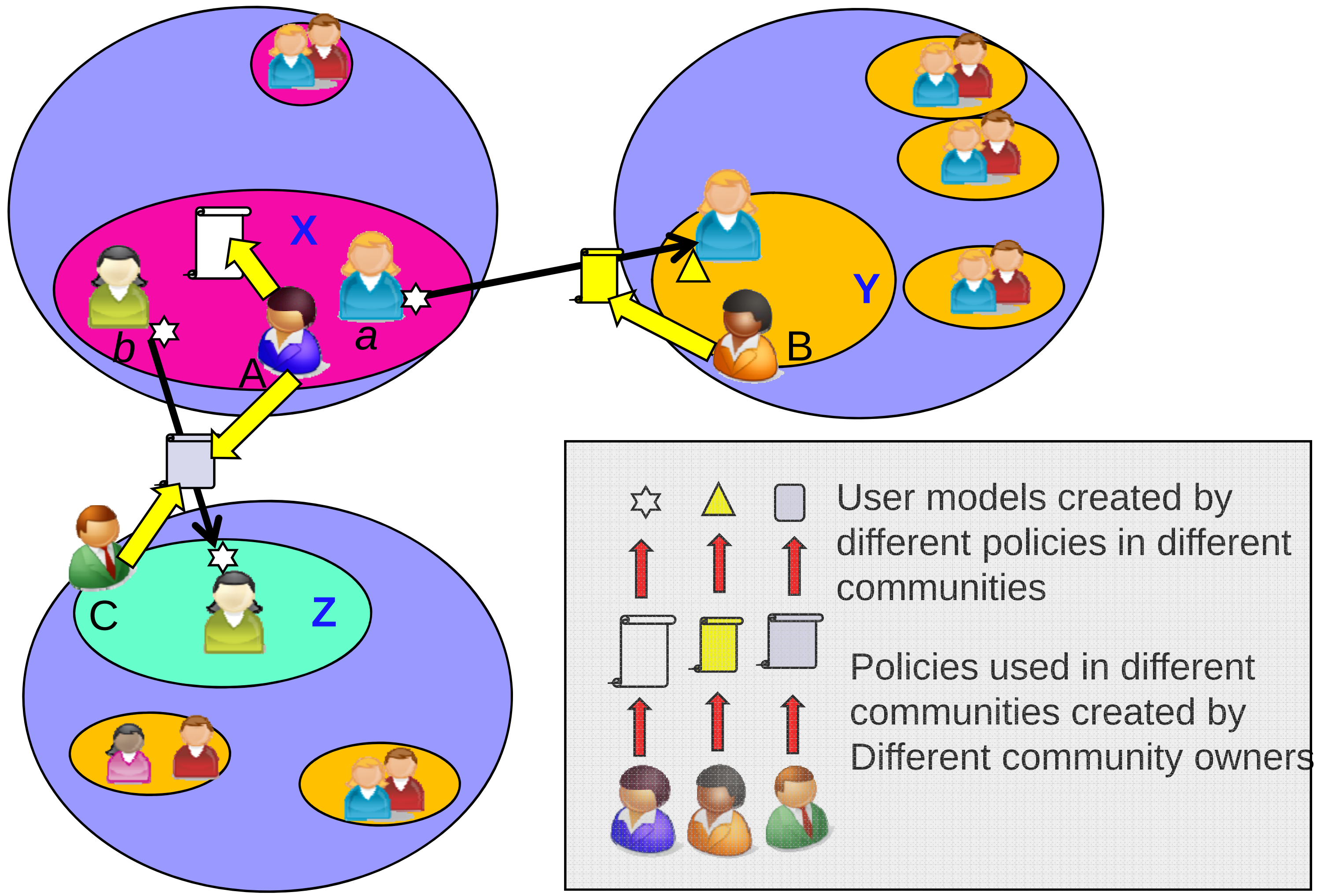
Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC

Policies in Comtella: user editable UM processes



Muhammad T., Vassileva, J. (2007) Policies for Distributed User Modeling in Online Communities, Proc. Workshop UbiDeUM'2007, at the 11th International Conference UM'2007, Corfu.

Comparison

Centralized UM

- Fixed schema (user features)
- De-contextualized
- Same update mechanisms
- Maintaining consistency is a goal
- Designed

Decentralized UM

- No fixed representation schema - diverse user features
- In context
- Purpose-defined update mechanisms
- No attempt to maintain consistency
- **Participative**

UM in MUE

UM Servers

Collaborative Recommenders

Expert Finding

Decentralized UM

Purpose-based UM

Policies in OC