ICALT

Toward a fully automatic learner model based on web usage mining with respect to educational preferences and learning styles

Mohamed Koutheaïr KHRIBI
University of Tunis
mk.khribi@uvt.rnu.tn

Mohamed JEMNI
University of Tunis
mohamed.jemni@fst.rnu.tn

Olfa NASRAOUI
University of Louisville
olfa.nasraoui@louisville.edu

Sabine GRAF
Athabasca University
sabineg@athabascau.ca

Kinshuk
Athabasca University
kinshuk@athabascau.ca

Learner Modelling Approaches

Collaborative   Automatic

- Advantages of automatic learner modeling :

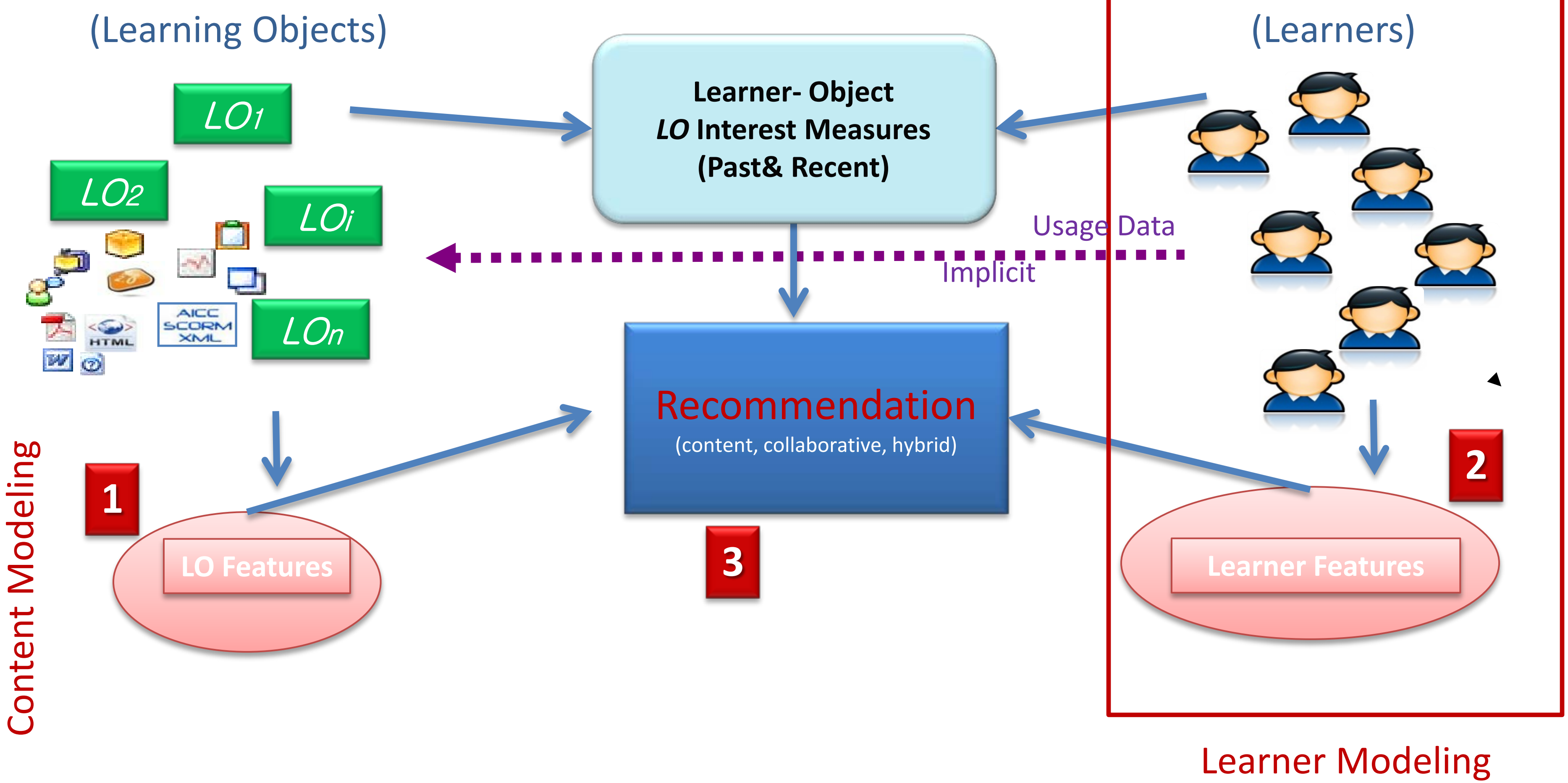
→ No additional work for learners ;

→ Uses information from a time span; higher tolerance

→ Allows dynamic updating of information

General Aim

Building an automatic recommendation system for Learning Management Systems.



Research Question

How to automatically model learners and groups of learners based on implicit data from their interactions and online activities in the learning management system, taking into account the learners'

* educational preferences
* learning styles

General Features of the Proposed Approach

→ Automatic, dynamic and based solely on learner usage sessions ;

→ Input data is composed from collected implicit data tracked and saved in LMS database and/or web server log files ;

→ Based on web usage mining techniques ;

→ Educational preferences and learning styles are considered and identified automatically.

Proposed Learner Model

→ Proposed Learner model components :

$LM_i = \{PR_i, LK_i, LEP_i\}$	
PR_i : i^{th} Learner Profile	General student information such as Identification data, Demographic information, etc.
LK_i : i^{th} Learner's Knowledge Component	Sequences of weighted visited learning objects i.e. vectors of visited learning objects or curriculum elements in which the student was interested (learner's knowledge)
LEP_i : i^{th} Learner's Educational Component	Learner's educational preferences and Learning style.

Learner's Knowledge Component

Let *LO* be a set of *n* unique visited learning objects: $LO = \{LO_1, LO_2, LO_3, \ldots, LO_n\}$.

The learner knowledge model LK_i corresponding to the learner *i* is represented by a set of *p* sessions S_j^i extracted from tracked data:

$LK_i = \{S_1^i, S_2^i, S_3^i, \ldots, S_p^i\}$, where each S_j^i is a subset of *k* weighted visited learning objects LO_l, $S_j^i : < w(LO_1^{S_j^i}), w(LO_2^{S_j^i}), w(LO_3^{S_j^i}), \ldots, w(LO_k^{S_j^i})>$, where each $LO_k^{S_j^i} = LO_l$ for some $l \in \{1, \ldots, n\}$, and $w(LO_k^{S_j^i})$ is the weight associated with the visited learning object $LO_k^{S_j^i}$ in the session S_j^i corresponding to the i^{th} learner.

The learning object weight $w(LO_k^{S_j^i})$, also called, learning object interest measure *LOIM*, is computed as a function of a number of features (based on the frequency of occurrence of the learning object in a session and/or the time a learner spends when visiting a learning object).

Learner’s Knowledge Component

To compute the weights of visited learning objects within the learner’s sessions, we apply two heuristics: *duration and frequency of visit*. Hence, the learning object weight $w(LO_k^{S_j^i})$, is obtained as follows:

$$w(LO_k^{S_j^i}) = (F(LO_k^{S_j^i}) / \sum_k F(LO_k^{S_j^i}) + T(LO_k^{S_j^i}) / \sum_k T(LO_k^{S_j^i})) / 2$$

$F(LO_k^{S_j^i})$ represents the frequency of visit of the learning object $LO_k^{S_j^i}$ within the session S_j^i, $\sum_k F(LO_k^{S_j^i})$ represents the sum of all visit frequency values related to the entire set of learning objects $LO_k^{S_j^i}$ within the session S_j^i, $T(LO_k^{S_j^i})$ represents the duration of a single visit to the learning object $LO_k^{S_j^i}$ in the session S_j^i, and $\sum_k T(LO_k^{S_j^i})$ represents the sum of all duration visits to the entire set of learning objects $LO_k^{S_j^i}$ in the session S_j^i.

Learner's Knowledge Component

→The learner's knowledge component LK_i can be represented as a matrix *M(p, n),* where *p* is the total number of learner's sessions and *n* the cardinality of unique visited learning objects.

$$
LK_i = \begin{array}{c|c} & LO_1 \quad LO_2 \quad LO_3 \quad LO_4 \quad \ldots \quad LO_n \\ \begin{array}{c} S_1^i \\ S_2^i \\ S_3^i \\ .. \\ .. \\ S_p^i \end{array} & \left(m_{xy} = w(LO_y^{S_x^i}), x \in \{1, \ldots, p\}, y \in \{1, \ldots, n\} \right) \end{array} : \begin{pmatrix} w(LO_1^{S_1^i}) & \cdots & w(LO_n^{S_1^i}) \\ \vdots & \ddots & \vdots \\ w(LO_1^{S_p^i}) & \cdots & w(LO_n^{S_p^i}) \end{pmatrix}
$$

Learner's Educational Component

→Composed of the **learner's preferences** among visited learning objects and his/her **learning style**.

→**Detection of the learner's preferences**:

What kind of learning object does a learner prefer?

Learning objects available in LMS are characterized by many attributes (e.g. format), each of which may have several values (e.g. for format: text, image, video, etc.) that could be preferred or not by the learner. The preferences of a learner *i* upon these values can be represented, as a vector of interest measures : $V_{EP}^{i}:\{LOIM\_Att_k\_LO_t^i\}_{k,t}$

Attributes	Related values
Type_LO(Learning object type)	*{Resource, Activity}*
Shape_LO(Learning object shape, if Type_LO = Activity)	*{Exercise, Simulation, Questionnaire, Assessment, Forum, Chat, Wiki, Assignment}*
Format_LO(Learning object format, if Format_LO = Resource)	*{Text, HTML, Image, Sound, Video}*

Learner's Educational Component

Let $LOIM\_Type\_LO_t^i$, $LOIM\_Format\_LO_t^i$, and $LOIM\_Shape\_LO_t^i$ be the interest measures composing the vector V_{EP}^i, which are computed over all learner's sessions S^i as follows:

$$LOIM\_Type\_LO_t^i = \frac{\sum_j LOIM\_Type\_LO_t^{S_j^i}}{Card(S^i)}, \; t \in \{1,2\}$$

$$LOIM\_Shape\_LO_t^i = \frac{\sum_j LOIM\_Shape\_LO_t^{S_j^i}}{Card(S^i)}, \; t \in \{1,8\}$$

$$LOIM\_Format\_LO_t^i = \frac{\sum_j LOIM\_Shape\_LO_t^{S_j^i}}{Card(S^i)}, \; t \in \{1,5\}$$

Where $Card(S^i)$ represents the number of sessions S_j^i in S^i belonging to the learner i, $LOIM\_Type\_LO_t^{S_j^i}$, $LOIM\_Format\_LO_t^{S_j^i}$ and $LOIM\_Shape\_LO_t^{S_j^i}$ represent the interest measures computed over each learner's session S_j^i .

Learner’s Educational Component

Attributes and corresponding values		Interest measures
Type_LO	Resource	$LOIM\_Type\_LO^{i}_{Resource}$
	Activity	$LOIM\_Type\_LO^{i}_{Activity}$
Shape_LO	Exercice	$LOIM\_Shape\_LO^{i}_{Exercice}$
	Questionnaire	$LOIM\_Shape\_LO^{i}_{Questionnaire}$
	test	$LOIM\_Shape\_LO^{i}_{Test}$
	Forum	$LOIM\_Shape\_LO^{i}_{Forum}$
	Chat	$LOIM\_Shape\_LO^{i}_{Chat}$
	Wiki	$LOIM\_Shape\_LO^{i}_{Wiki}$
	Navigation	$LOIM\_Shape\_LO^{i}_{Navigation}$
	example	$LOIM\_Shape\_LO^{i}_{Exeamle}$
Format_LO	Text	$LOIM\_Format\_LO^{i}_{Text}$
	Html	$LOIM\_Format\_LO^{i}_{Html}$
	Image	$LOIM\_Format\_LO^{i}_{Image}$
	Audio	$LOIM\_Format\_LO^{i}_{Audio}$
	Vidéo	$LOIM\_Format\_LO^{i}_{Video}$

$$LOIM\_Type\_LO_t^{S_j^i} = \sum_k w(LO_{k,t}^{S_j^i}) \Big/ Nbre_{LO_t^{S_j^i}}, t \in \{1,2\}$$

$$LOIM\_Type\_LO_t^i = \sum_j LOIM\_Type\_LO_t^{S_j^i} \Big/ Card(S^i), t \in \{1,2\}$$

$$LOIM\_Shape\_LO_t^{S_j^i} = \sum_k w(LO_{k,t}^{S_j^i}) \Big/ Nbre_{LO_t^{S_j^i}}, t \in \{1,8\}$$

$$LOIM\_Shape\_LO_t^i = \sum_j LOIM\_Shape\_LO_t^{S_j^i} \Big/ Card(S^i), t \in \{1,8\}$$

$$LOIM\_Format\_LO_t^{S_j^i} = \sum_k w(LO_{k,t}^{S_j^i}) \Big/ Nbre_{LO_t^{S_j^i}}, t \in \{1,5\}$$

$$LOIM\_Format\_LO_t^i = \sum_j LOIM\_Format\_LO_t^{S_j^i} \Big/ Card(S^i), t \in \{1,5\}$$

Learner's Educational Component

→ **Detection of the learning style** (Graf et al., 2008)

Sensing/Intuitive

Actifve/Reflective

F S L S M

Visual/Verbal

Sequential/Global

Commonly incorporated features in LMS

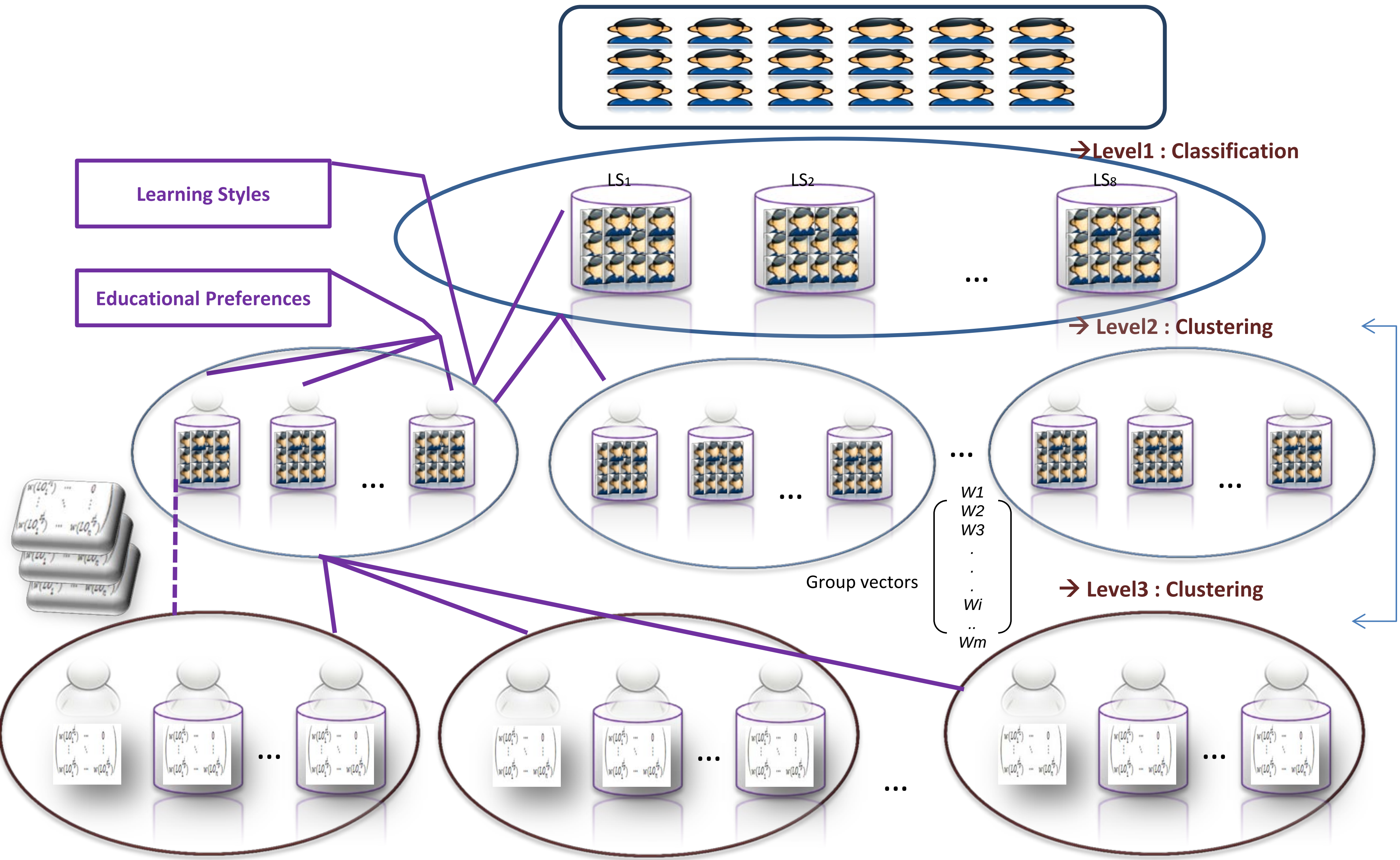
- Content
- Outline
- Example
- Self-Assessment
- Exercice
- Forum
- Navigation

Active/Reflective	**Sensing/Intuitive**	**Visual/Verbal**	**Sequential/Global**
Content_visit(-)	Example_visit(+)	Ques_text(-)	Outline_stay(-)
Content_stay(-)	Example_stay(+)	Forum_visit(-)	Ques_detail(+)
Outline_stay(-)	Selfass_visit(+)	Forum_stay(-)	Ques_overview(-)
Selfass_stay(-)	Selfass_stay(+)	Forum_post(-)	Ques_interpret(-)
....	....	....	....
....	....	....	....

Group Modeling

$$\left(LP_i \quad , \quad \begin{pmatrix} w(LO_1^{s_1^i}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ w(LO_1^{s_p^i}) & \cdots & w(LO_n^{s_p^i}) \end{pmatrix} , \quad \{(LEP_1, .. , LEP_p), LS_i\} \right)$$

→Once learner models are built, we apply a **hierarchical multi-level model based collaborative filtering approach** on these models, in order to assign learners with common preferences and interests to the **same groups**, so that **feedback from one learner can serve as a guideline** for information delivery to the other learners within the same group.


Learning Styles
Educational Preferences
→Level1 : Classification
LS1
LS2
LS8
…
→ Level2 : Clustering
W1
W2
W3
Wi
Wm
Group vectors
→ Level3 : Clustering

Implementation and Proof of Concept Evaluation

→ The proposed approach has been implemented and an experiment was performed **as part of a recommendation approach**

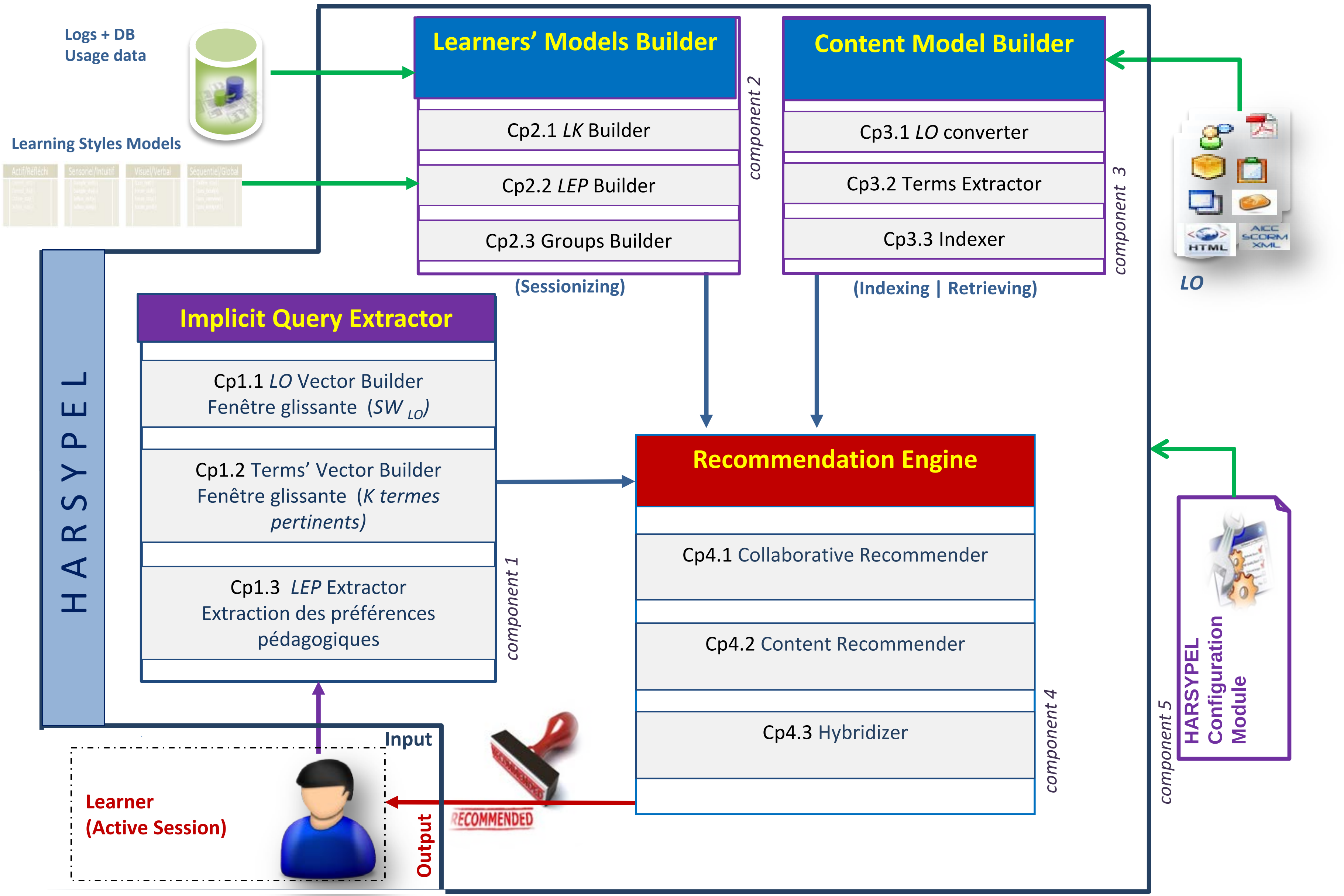
→ **Recommendations are computed** with respect to the learner's clickstreams, his/her learning style and educational preferences, as well as exploiting similarities and dissimilarities among the learner models and educational content.

→ **Moodle** LMS is used to implement the proposed approach.

→We used an online hybrid course (C2i) with **704 learners**.

→**Tracked data was successfully extracted** from various Moodle tables (primarily from mdl_log table).

HARSYPEL Architecture



Conclusions and Future Work

→ Developed an approach for automatically modeling learners (groups) within LMS taking into account their educational preferences and learning styles

→ Proposed approach falls within the scope of building an **educational automatic hybrid recommender system** providing suitable recommendations to learners for personalized technology enhanced learning ;

Future work:

→ Evaluation of the recommender system