



DUINE RECOMMENDER

SUPPORTING PEOPLE IN FINDING INFORMATION

**HYBRID RECOMMENDER SYSTEMS AND GOAL-BASED
STRUCTURING**

1

Project de cours INF6304

Automne 2009

Rashin ESMAEILI

Matricule:1473412

PLAN DE PRÉSENTATION

- 1) **Introduction**
- 2) **Définitions et concepts de base**
- 3) **Duine Conseilleur**
- 4) **Conclusion**

1) INTRODUCTION

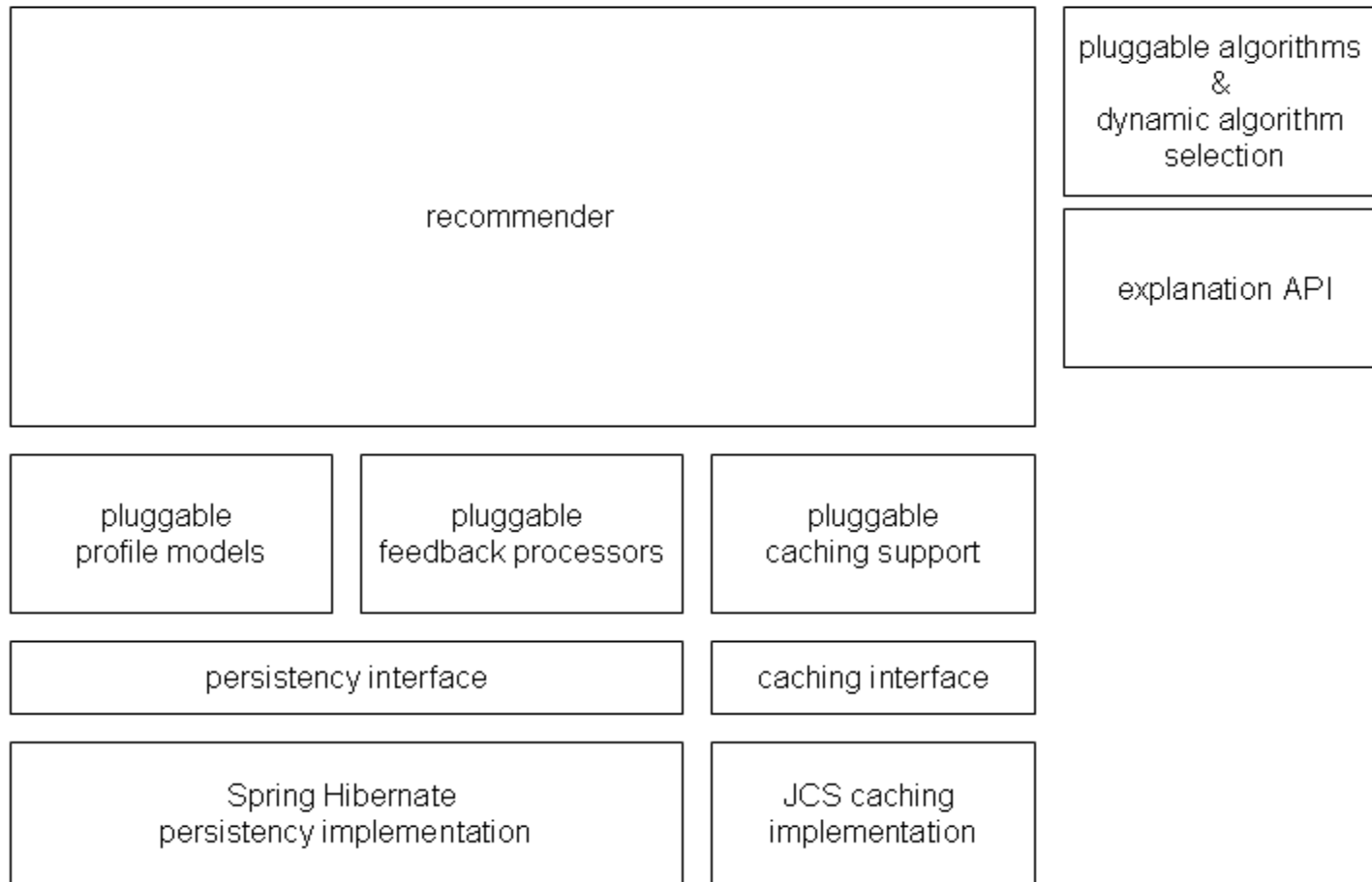
- **Qu'est- ce que c'est un conseiller Duine?**
- **Comment ca fonctionne le conseiller Duine?**
- **A quel niveau le conseiller Duine est-il utile?**
- **Quels d'autres services sont offert par le conseiller Duine?**



QU'EST- CE QUE C'EST UN CONSEILLEUR DUINE?

- **Duine Framework:**le cadre du Duine (Duine Framework) est une (collection des) bibliothèques logicielle qui permet aux développeurs de créer les moteurs de prédiction pour ses propres applications
- Duine est un mot gaélique irlandais qui signifie une personne et est prononce comme “dinne” (diner sans un ‘r’).
- Duine a été développe par [Telematica Instituut/Novay](#) , basée sur une recherche scientifique sur la personnalisation et plus précisément les systèmes conseillers
- C'est un module logiciel qui calcule a quelle niveau les informations fournis sont intéressants pour un utilisateur donne.
- Le résultat de ce calcule est quantifie par un chiffre, dit la prédiction, avec une valeur entre -1 (i.e. pas de tout intéressant) et +1 (i.e. tout a fait intéressant)
- Le conseiller Duine élabore et sauvegarde les évaluations dont les utilisateurs font pour les items d'information. Tous les données associées à un utilisateur est ensuite sauvegarde dans son profile d'utilisateur.
- Duine est capable d'apprendre .Quand un utilisateur évalue un item d'information, le conseiller extrait les données liées à cet item afin déterminer les intérêts de l'utilisateur. En utilisant les algorithmes intelligents d'apprentissage, il va modifier le profile de l'utilisateur après chaque évaluations, basée sur ces infos.

Une vue d'ensemble des caractéristiques de Duine





COMMENT FONCTIONNE LE CONSEILLEUR DUINE?

- Duine calcule les prédictions en exécutant des algorithmes qui raisonnent à propos d'un utilisateur et les items d'information, et aboutit à une valeur de prédiction .Ces algorithmes sont appelés les techniques de prédictions. Les techniques de prédictions utilisent les profils d'utilisateurs et les items d'information, en tant que des données d'entrée pour ses calculs.
- Exemple des techniques de prédictions utilisées:
 - **Les Filtres Collaboratifs**: Les gens qui ont évalués les mêmes items de la même façon dans le passé, probablement ont le même goût (*l'approche utilisateur-utilisateur* ☺)
 - **Les Filtres d'Information**: Extrait les information d'un item (e.g. la description d'un programme TV), fait une analyse textuel et compare les résultats avec les profils d'utilisateurs afin de déterminer à quel niveau l'item sera intéressant pour un utilisateur (*l'approche item-item* ☺)
 - **Case-based Reasoning** (basée sur les similarités entre les items qui sont déjà été évaluées par un utilisateur dans le passe)
 - **GenreLMS** (raisonnement sur les genres),
 - **TopNDeviation** (popularité),
 - **AlreadyKnown**
 - **UserAverage**

A QUEL NIVEAU LE CONSEILLER DUINE EST-IL UTILE?

- Duine est un conseiller prétendu **hybride** ! Ca veut dire qu'il combine plusieurs techniques de prédictions au lieu de compter sur une seule technique spécifique. Un conseiller hybride est mieux qu'un conseiller algorithme-unique, car:
 - Il donne des résultats plus corrects
 - Il diminue le problème de démarrage à froid (quand il n'y pas assez de données concernant un utilisateur ou un item pour faire la recommandation)
- Chaque différent technique a ses propre points forts et points faibles. L'exactitude et utilité d'un technique dépende largement aux types de données qui est disponible dans les profiles d'utilisateurs, le nombre des utilisateurs et les métadonnées disponible pour les items.
- Les règles de décision qui gère le comportement changeant sont appelés les stratégies de prédictions.
- Les stratégies et les techniques de prédictions peuvent être combine par une configuration. Cette configuration ne nécessite aucun codage du logiciel. Pour chaque domaine d'application, la configuration doit être adapte pour les caractéristiques spécifiques de ce domaine.
- L'exactitude d'un conseiller est déterminée par l'optimisation de la configuration pour le domaine d'application.



QUELS D'AUTRES SERVICES SONT OFFERT PAR LE CONSEILLER DUINE?

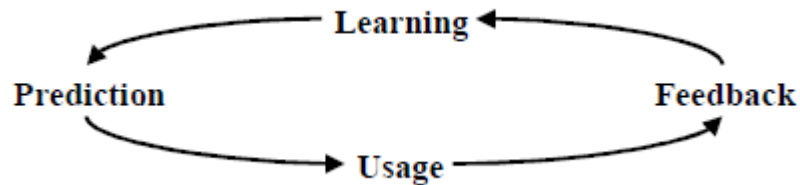
- **Extensibilité** : Duine peut s'étendre avec les nouvelles techniques et stratégies de prédictions, aussi avec des nouveaux modèles de profils et un processeur de rétroaction (feedback processor). De plus, l'implémentation fournie avec la persistance de spring hibernate peut être remplacée par une autre implémentation, et on peut remplacer la cache JCS (Java *Caching* System) par une autre.
- **Validation** : le cadre de Duine vient avec des outils de validation. En donnant un set de données (dataset) avec les évaluations faites par les utilisateurs, le cadre de validation peut mesurer l'exactitude des prédictions. Le cadre de validation peut être utilisé pour optimiser la configuration pour un domaine d'application spécifique.
- **Explication API** : Duine peut expliquer comment les résultats de prédictions sont déterminés. Cette explication API peut être utilisée pour créer les explications "end-user friendly" pour les applications.

DÉFINITIONS ET CONCEPTS DE BASE

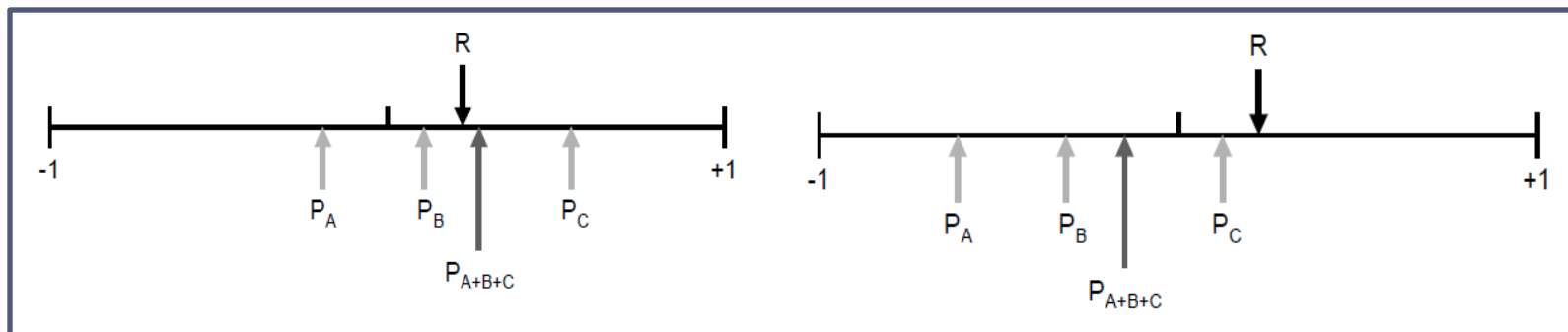
- **Un Système conseiller**
- **Recommandations, prédictions et les votes**
- **Les techniques de prédiction**
- **Les techniques Hybride de prédiction**
- **Les méthodes d'hybridation**
- **Couplage**
- **Certitude et l'apprentissage**
- **Les techniques de l'apprentissage**

UN SYSTÈME CONSEILLEUR

Le Processus de Prédiction

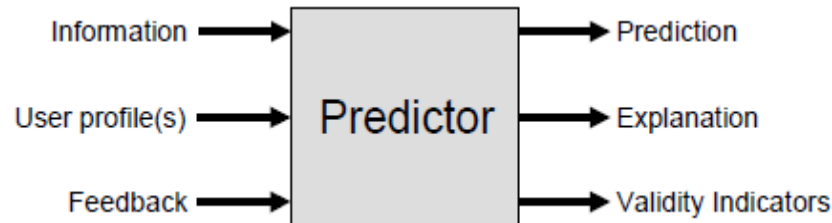


1. *La Phase de prédiction* : une prédiction est générée pour un item pour un utilisateur (la sélection de prédicteurs, La combinaison de prédicteurs).
2. *La Phase de l'usage* : l'utilisateur et/ou d'autres parties de systèmes de conseiller utilise prédictions générés (filtrage, classification (structuring) & présentation).
3. *La Phase de feedback* : Après qu'un utilisateur a utilisé un ou plusieurs items, il peut fournir feedback pour le système à propos de ses intérêts actuels aux items. (Implicite et explicite).
4. *La Phase de l'apprentissage*: Le système apprend de feedback fourni par l'utilisateur, qui est utilisé pour augmenter l'exactitude des prédictions de future (maximum path, breadth first technique)



UN SYSTÈME CONSEILLEUR

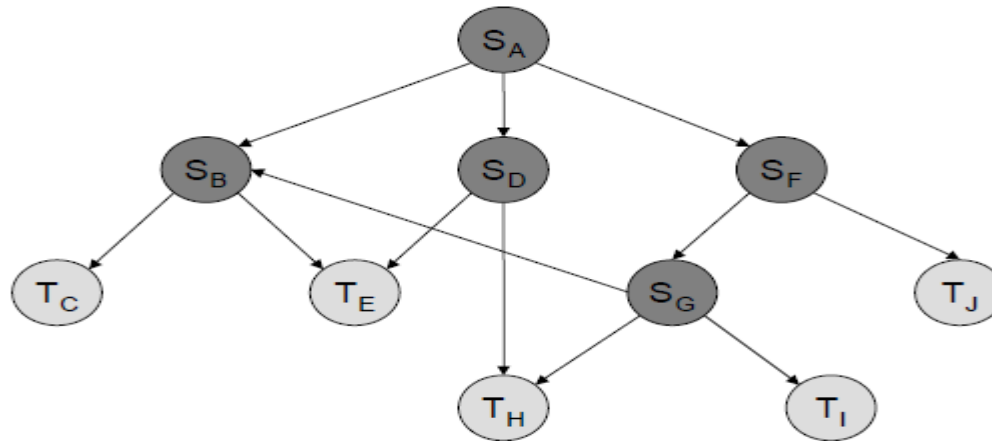
Le modèle générique de prédicateur



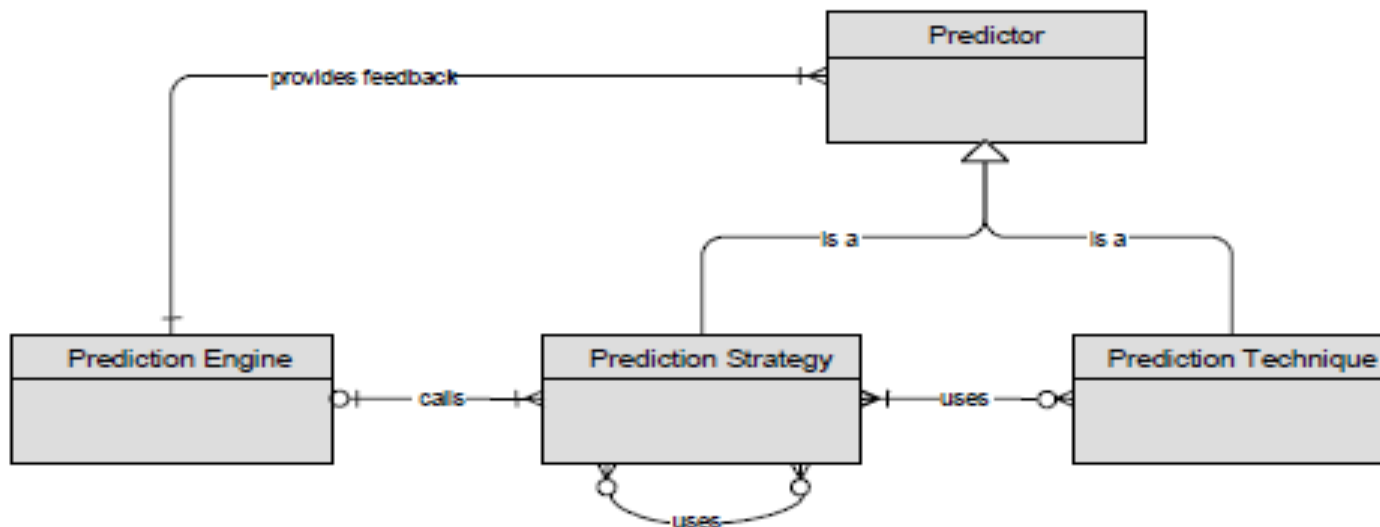
Le moteur de prédiction: un composant de logiciel qui élabore chaque demande entrant de prédictions et est le point d'entrée principal pour le traitement des feedback.

- 1. Séparer une demande de prédictions des items multiples en demandes de prédiction pour l'item unique**
- 2. Séparer une demande de prédictions des utilisateurs multiples en demandes de prédiction pour chaque utilisateur individuel**
- 3. demander directement pour divers classes des items aux stratégies de prédiction exacte.**
- 4. Elaborer les feedback ramasse de l'utilisateur (explicitement ou implicitement) fournit pour apprendre des feedback pour améliorer les prédictions de future)**

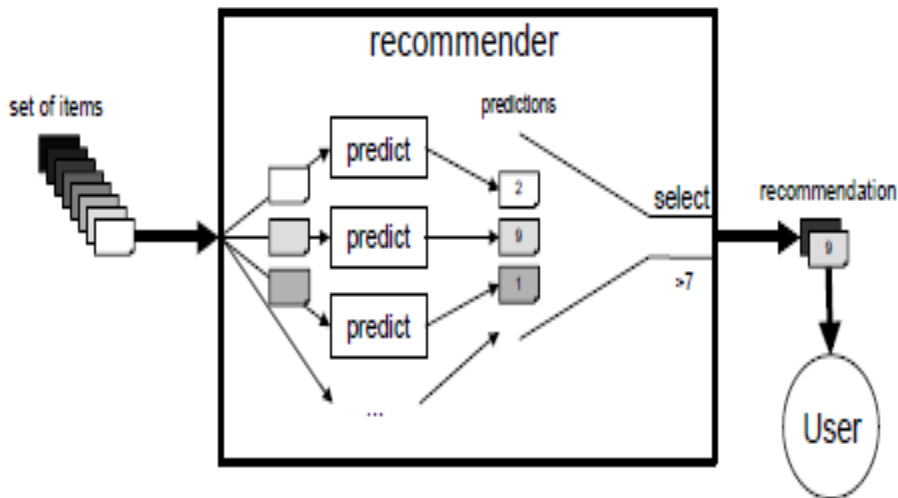
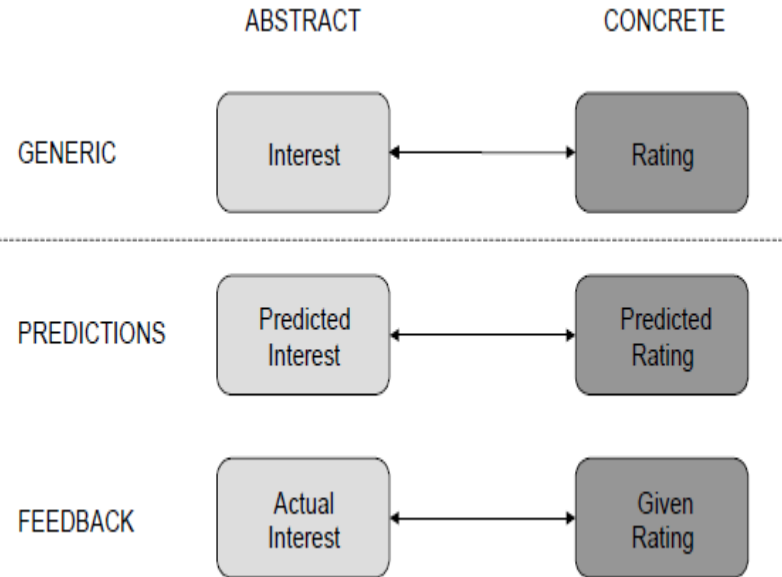
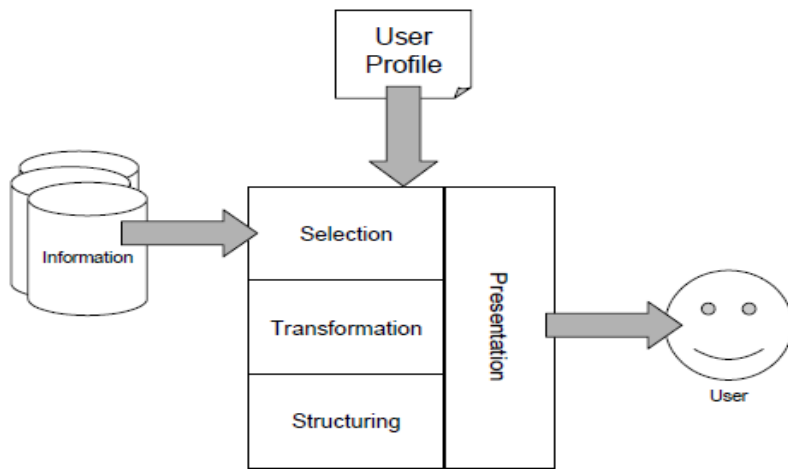
UN SYSTÈME CONSEILLEUR



Les stratégies de prédictions hiérarchiques sont les graphes orientés acycliques . Tous les nœuds sortants (local sinks) sont les techniques de prédiction (T) et tous les nœuds réguliers sont les stratégies de prédiction (S). Une stratégie de prédiction (SA) est la source globale de graph et c'est le point de démarrage pour la demande de prédiction .



RECOMMANDATIONS, PRÉDICTIONS ET LES VOTES



LES TECHNIQUES DE PRÉDICTION :

- **Social-based prediction techniques:**
 - *Collaborative filtering*
 - *Item-item filtering*
 - *Stereotypes and demographics*
 - *Popularity*
 - *Average*
- **Information-based prediction techniques:**
 - *Information filtering*
 - *Case-based reasoning*
 - *Attribute-based prediction techniques*

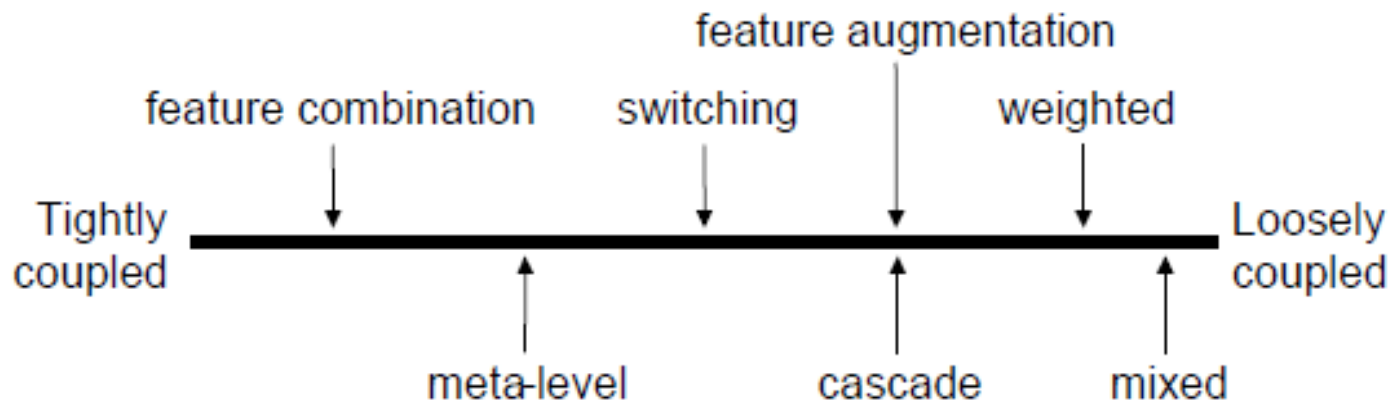
LES TECHNIQUES HYBRIDE DE PRÉDICTION

LES MÉTHODES D'HYBRIDATION

- Sachant que chaque technique de prédiction a ses vigueurs et faiblesses, il y a un besoin de combinaison des différentes techniques de prédiction afin d'augmenter l'exactitude des systèmes de conseiller. L'idée derrière des techniques de prédiction hybride c'est que la combinaison des algorithmes peut fournir les recommandations plus correctes, qu'avec un algorithme unique, car les désavantages d'un algorithme peuvent surmonter des autres
- **Les méthodes d'hybridation :** pondère (weighted), changement (switching), mélangé (mixed), combinaison des caractéristiques (feature combination), cascade, augmentation des caractéristiques (feature augmentation) et méta-niveau (meta-level)

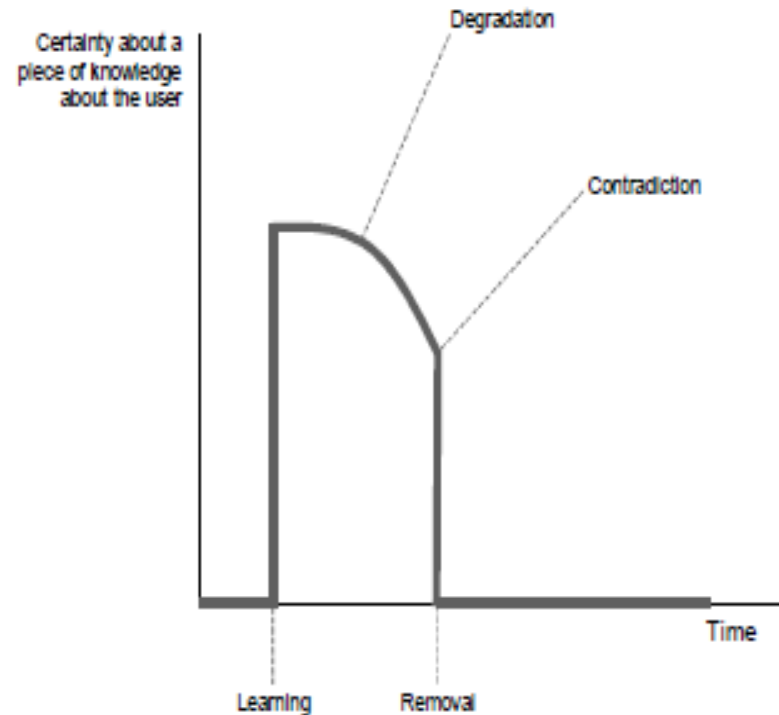
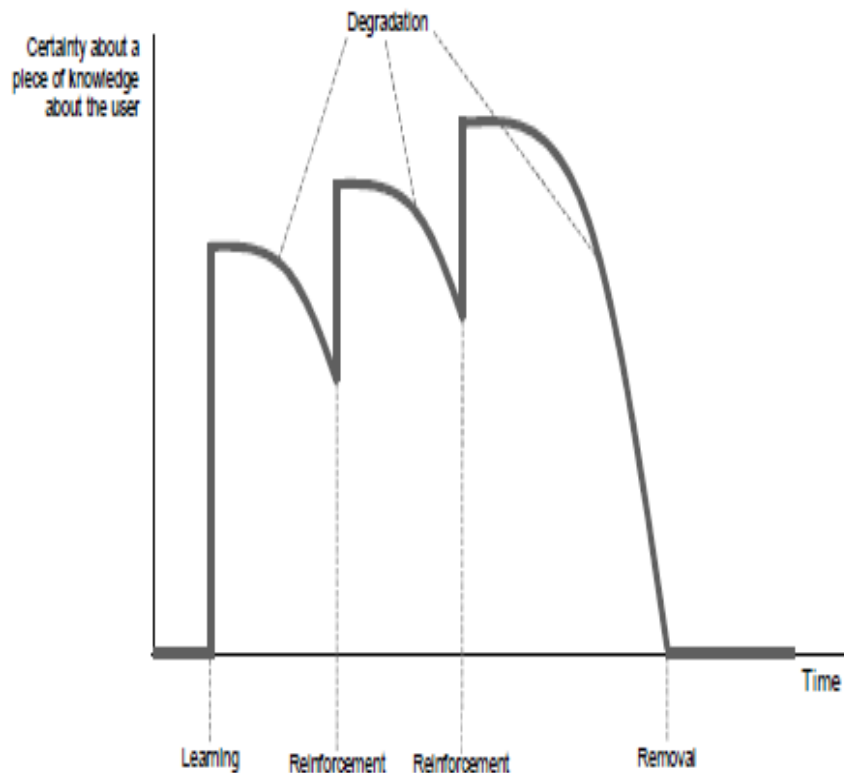
COUPLAGE :

- Le Couplage tendue (*Tight coupling*) : deux ou plus de techniques de prédictions (A, B) sont combinées dans une techniques de prédictions (C), ou la nouvelle techniques de prédictions possède et utilise la connaissance de la mécanismes interne des techniques combinées = $f(A, B)$
- Le Couplage relâche (*Loose coupling*) : deux ou plus de techniques de prédictions (A, B) sont relâchement couplées dans une nouvelle technique de prédiction (C), ou C combine des résultats des techniques couplées et elle n'a aucune connaissance de la mécanismes interne des techniques combinées = $g(f(A), f(B))$



CERTITUDE ET L'APPRENTISSAGE

Il y a 5 événements qui impactent la certitude de la connaissance d'un utilisateur: apprentissage, renforcement, dégradation, enlèvement et contradiction



LES TECHNIQUES DE L'APPRENTISSAGE

Decision Technique	Learning	Type	Updating
Decision Trees and Decision Rules	Manual	Model	Static or Staged
	Automatic		
Case-Based Reasoning	Automatic	Instance	Static, Staged or Dynamic
Backpropagation Artificial Neural Network	Automatic	Model	Static or Staged
Bayesian Probabilities	Automatic	Instance	Dynamic
		Model	Static or Staged

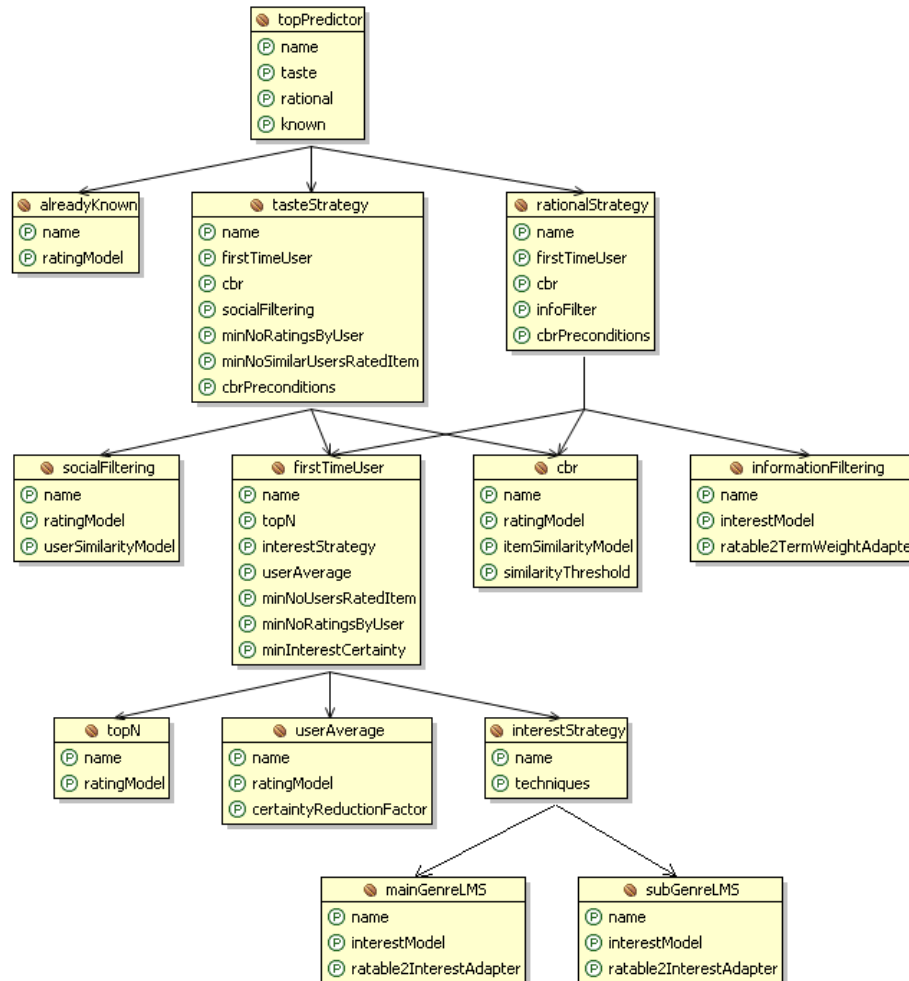
3) DUINE CONSEILLEUR

- Les Concepts principaux et les caractéristiques**
- Modèles de Profile**
- vue d'ensemble des techniques de prédiction**
- Configuration**
- Validation**
- extension de Cadre(Framework)**

LES CONCEPTS PRINCIPAUX ET LES CARACTÉRISTIQUES

- **Le cadre (framework) de prédiction** : Independent aux domaines, utilise les systèmes de conseillers hybrides qui peuvent être adaptées aux domaines divers afin de fournir les recommandations plus précises.
- **Switching hybridization** : fournit le meilleur équilibre entre le couplage relâche et le couplage tendu
- **Goal-based structuring**: les techniques de prédictions courant adresse les intérêts à long terme ou à court term, mais pas les intérêts instantané; i.e. Les buts que les gens ont pour les items qu'ils les cherchent .Duine a aussi examine comment les buts des utilisateurs peuvent être mieux utilise dans les systèmes de recommandations
- **La Présentation des prédictions**: c'est le mieux d'utiliser un format traditionnel de présentation. Les gens croient que les systèmes de recommandation doit être capable d'expliquer les prédictions clairement, mais seulement sur la demande! (Il faut non plus donner trop de détails), Il n'y a pas de préférence pour le type des explications.
- la majorité des gens préfère l'usage des feedback implicites (surveiller le comportement de l'utilisateur) car il a besoin de moindre effort de la part de l'utilisateur

LES CONCEPTS PRINCIPAUX ET LES CARACTÉRISTIQUES PREDICTOR TREE



MODÈLES DE PROFILE

- Les modèles de profils sont utilisés par les prédicteurs pour aller chercher les données dont ils ont besoin comme entrée pour calculer les prédictions. Les modèles de profils peuvent être mis à jour par les processeurs de feedback, lorsqu'un utilisateur fournit le feedback.
- Dans les 4 modèles de profils utilisés, chacun est complètement réservé, et il n'a aucune dépendance avec les autres. Chacun gère son propre persistance. Ils ont l'accès à son couche de persistance via un interface DAO(data access object).
- **Les Modèles d'évaluations(RatingModel):**accès aux évaluations des utilisateurs et les sauvegarder
- **Les Modèles d'intérêts (InterestModel):**Pour éviter de créer un nouvel modèle de profile pour chaque modèle d'intérêt, Duine fournit un composant générique (InterestModel) pour réalisation des modèles d'intérêts
- **Les Modèles de similarité entre utilisateurs(UserSimilarityModel):** les similarités entre les utilisateurs peuvent être accéder via une interface. Il y a en 2 implémentations: *UserSimilarityModel* et *UserSimilarityModelOnTheFly*
- **Les Modèles de similarité entre les items(ItemSimilarityModel):** les similarités entre les items peuvent être accéder via une interface. Il y a en 2 implémentations: *ItemSimilarityModel* et *ItemSimilarityModelOnTheFly*

VUE D'ENSEMBLE DES TECHNIQUES DE PRÉDICTION

Prediction Technique	Profile Model	Ratable Item Adapter	Description
UserAverage	Rating Model		Returns the average of the ratings that the user has already given to other items
AlreadyKnown	Rating Model		Returns the existing rating of the item for this user (if available)
TopNDeviation	Rating Model		Returns a prediction based on all n ratings from other users that already rated the item
Collaborative Filtering	Rating Model, User Similarity Model		People who have rated the same items the same way in the past probably have the same taste. Based on this knowledge one can predict how much a person likes an unseen item when similar users have already rated that item.
Information Filtering	InterestModel	IRatable2TermWeightAdapter	Extracts information from an item (e.g. the description of a TV program), does a textual analysis and matches the result with the user profile to determine how interesting the item is.
Case-based Reasoning	Rating Model, Item Similarity Model	IRatable2SimilarityAdapter	Based on the idea that if two items are similar and if a rating is known for one of them, the rating for the other one will probably be the same.
InterestLMS	Interest Model	IRatable2InterestAdapter	Predicts how interesting an item is bases on specific aspects of this item (e.g. the genre)

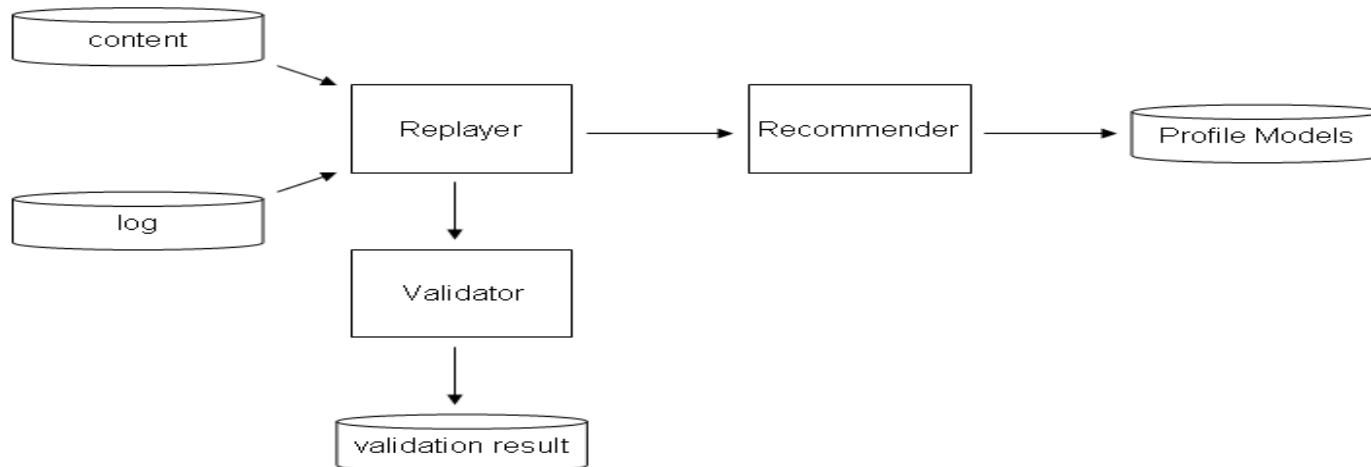
CONFIGURATION

- L'arbre de prédicateur est totalement configurable
- C'est possible d'ajuster les paramètres des prédicateurs existants, et modifier les paramètres de l'apprentissage du processeur rétroaction (feedback)
- La configuration du système peut sembler plutôt compliquée à un premier regard
- La configuration par défaut de Duine peut aider !
- Pour télécharger et puis transférer les configurations, on utilise 'Spring Framework'
- Les fichiers de configuration pertinents sont les suivants:
- Ils sont disponibles dans le répertoire *:src/main/resources*, dans la source d'application demo

spring-recommender.xml	High level configuration of the recommender
spring-predictors.xml	Configuration of the predictor tree
spring-profiles.xml	Configuration of the profile models, the adapters and the feedback processors
spring-datasources.properties	Configuration of the datasources

VALIDATION

Le cadre de validation Demo Pour MovieLens

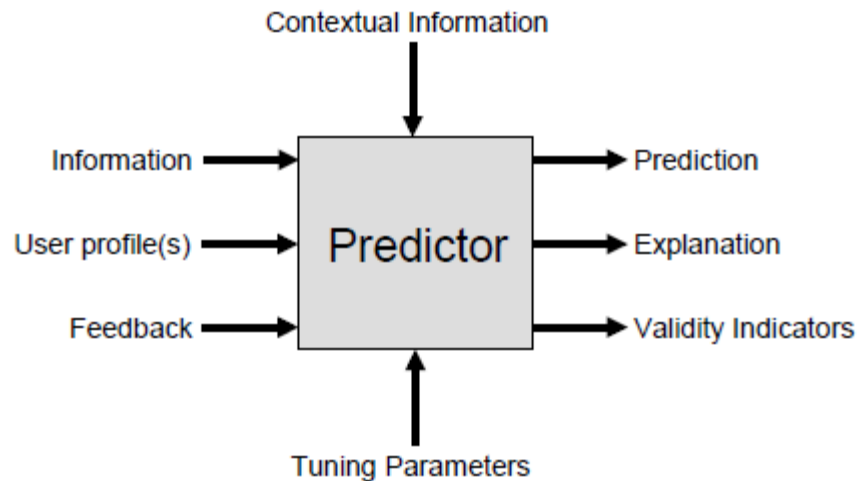


Afin de mesurer et vérifier l'exactitude de conseiller, le Cadre de Duine offre un outil de validation. Les données adoptées pour l'outil de moteur de validation sont les « log usage des utilisateurs ; quels évaluations ils ont déjà fournit pour les items d'information, quels intérêts ont déjà mise à jour explicitement, etc.

Ca sera judicieux de faire les testes de validation de temps en temps, afin d'optimiser les attribues réglable pour chaque modification dans l'usage d'un système. Aussi, si on ajout des nouvelles techniques de prédictions dans le système de validation, ça sera très utile de vérifier l'amélioration des recommandations.

EXTENSION DE CADRE(FRAMEWORK)

- Les paramétrés de réglage(Tuning) : faire le cadre plus the flexible et l'implémentation dans les divers domaines plus facile.
- La conscience de contexte: introduit une nouvelle source de l'information sur lesquels on peut faire les prédictions plus précises en considérant les contextes de l'utilisateur.



4) CONCLUSION

- **Limite de cette recherche**
- **Le travaux Future(Roadmap)**
- **Difficultés rencontrées pour faire une démonstration**
- **Les capture d'écran d'un prototype**
- **Question?**

LES LIMITES DE CETTE RECHERCHE

- **Généralisabilité:** la validation est fait sur deux domaines seulement, mais on peut espérer que ca fonctionne bien sur autres domaines aussi!
- ***Goal-based structuring:*** on peut espérer aussi que le structuring des items base sur les buts attendus sous les formes des utilités et gratification peut aider les gens à trouver les items intéressants, mais comme les utilités et gratifications actuels peut agir de façons différentes dans d'autres domaines, une recherche est nécessaire !
- ***La Présentation des prédictions:*** On peut espérer que la règle de base de « keep it simple » pour les interfaces utilisateur corresponde aussi dans d'autre domaines (Pas besoin de donner des explications compliqués!)

LE TRAVAUX FUTURE(ROADMAP)

- **Les techniques de Prédiction:** ajouter plus de techniques, par exemple la technique des prédictions de factorisation matrice.
- **Enrichissement les Métadonnées:** Ce composant est. Pour le but de supporter la création automatique de métadonnées (e.g. avec Apache UIMA framework) et puis pour un autre composant pour sauvegarder et gérer la cycle de vie de métadonnées
- **Etiquetage de l'utilisateur:** Ce composant doit supporter l'annotation de la contenu avec des étiquettes, particulièrement la suggestion de l'étiquette (c'est très à la mode!)
- **Le Conseiller Duine entant qu'un service:** Avec l'ajout d'un REST API à Duine, c'est possible d'avoir des hôtes de Duine sur une machine différente (sans une application web)
- **Les réseaux Sociaux:** profiter les informations des profile d'un utilisateur a partir de son réseau social préféré
- **La confiance :** un nouveau domaine de recherche en relation avec des réseaux sociaux
- **Contexte:** implémenter un modèle de contexte qui reconnaît les contextes d'un utilisateur et qui réduit ou augmente l'impact de certains attribues d'un profile d'un utilisateur, basée sur son contexte à un moment donné.
- **GUI et les outils de surveillance:** Avoir un éditeur visuel sera préférable qu'un éditeur XML, au moins pour les développeurs non-expert
- **Les modules additionnels:** l'ajout des caractéristiques additionnelles qui sont utiles dans la contexte de recommandation (e.g pour éviter la monotonie)

DIFFICULTÉS RENCONTRÉES POUR FAIRE UNE DÉMONSTRATION :

- Au cours de l'Installation:
 - On peut télécharger le fichier d'installation de: <http://sourceforge.net/projects/duine/files/>
 - *Pour movielens distribution:* sur le repertoire bin, executer *run.bat*
 - Ca exécute l'application java MovieLensRecommenderClient. Le program va automatiquement (ré) créer les tableaux des base de données et exécute 2 petits scenarios de prédictions/feedback s.
 - Apres il faut créer le conseiller avec différents scenarios
 - Les scenarios sont les programmes JAVA! (voir exemple) ☹
- Pour faire une utilisation de ce logiciel:
 - Apparemment c'est possible d'utiliser l'outil de validation de Duine pour faire une démonstration de Duine avec les donnes de teste (par exemple MovieLens data set).
 - Pour ce faire, Il faut:
 - Télécharger la base de donnes de movielens
 - Créer et configure la base de donnes
 - Import la base de donnes de movielens
 - Générer une base de données de log
 - Exécuter le Replay sur la base de données de log
 - Interpréter les résultats de validation
 - En résumé, il faut avoir les connaissances de plusieurs choses (MySQL, JAVA, configuration et transformation des bases de donnes,..) ☹

UTILISATION DE DUINE

Create a recommender

The code snippet below, copied from `MovieLensRecommenderClient.java`, shows how to create a recommender. The code instantiates a new recommender that is configured with the xml files that are provided as arguments. We use the [Spring Framework](#) to load and wire the configuration. The `SpringRecommenderFactory` is the only class of the Duine Framework core that has a dependency with Spring.

```
/**
 * This is where the configuration files are loaded. Together the loaded files represent one global application context.
 */
private void init() throws DuineException {
    IRecommenderFactory factory = new SpringRecommenderFactory();
    recommender = factory.createRecommender("recommender", new String[] { "spring-recommender.xml", "spring-predictors.xml",
        "spring-profiles.xml", "spring-services.xml", "spring-datasources.xml" });
}
```

Scenario 1, part 1: Add a new user and ask for a prediction about a movie

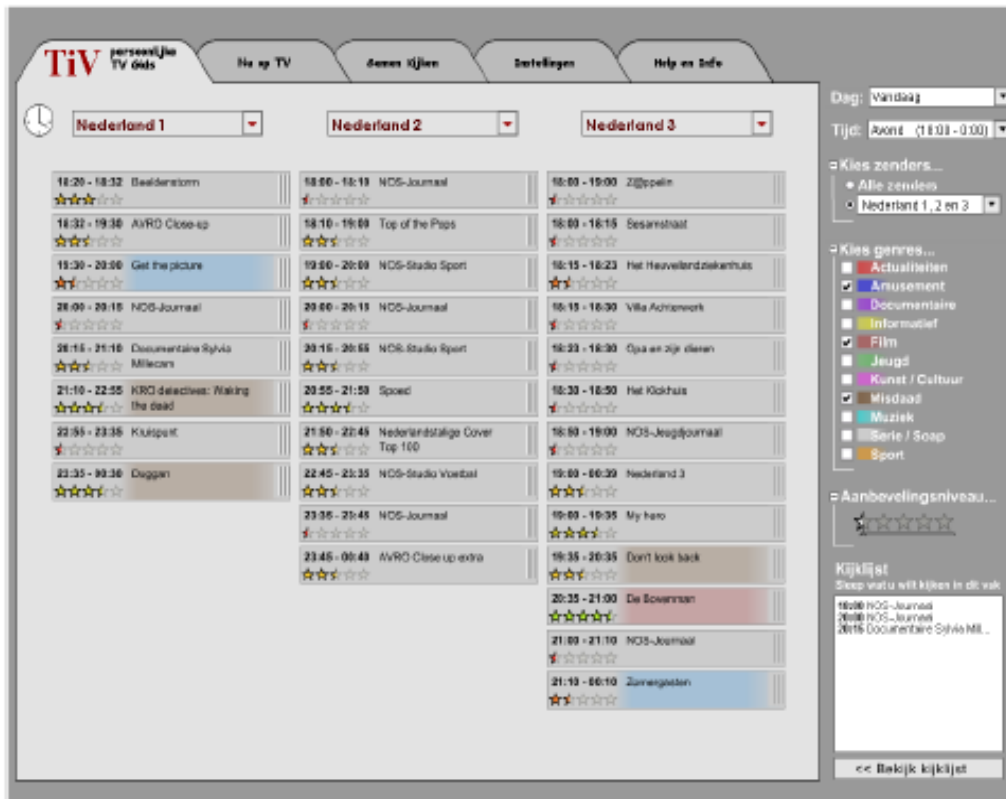
In this scenario we create a new user (`user1`) and a new movie (`movie1`), and ask the recommender to give a prediction for the interest of the user in this movie. Initially the recommender has no knowledge about the user.

```
public void runScenario1() {
    log.info("***** SCENARIO 1 *****");
    log.info("Create a new movie with genre 'sports'");
    RatableItemId item1 = new RatableItemId("1");
    Movie movie1 = new Movie(item1);
    ArrayList<String> genres = new ArrayList<String>();
    genres.add("sports");
    movie1.setGenres(genres);
    movie1.setTitle("movie1");
    log.info("Create a user id 'user1'");
    UserId user1 = new UserId("user1");
    log.info("Calculate a prediction for the interest of this user in this movie");
    Prediction prediction = recommender.predict(user1, movie1);
    log.info("Prediction result: " + prediction);
    .....
}
```

Output:

```
INFO MovieLensRecommenderClient:88 - Give a rating to the movie (value=0.9, certainty=0.8)
INFO MovieLensRecommenderClient:91 - Calculate a prediction for the interest of this user in this movie
INFO MovieLensRecommenderClient:93 - Prediction result: Prediction(item=1, pred=0.9, cert=0.8)
```

PROTOTYPE SCREENSHOTS



First prototype: main screen

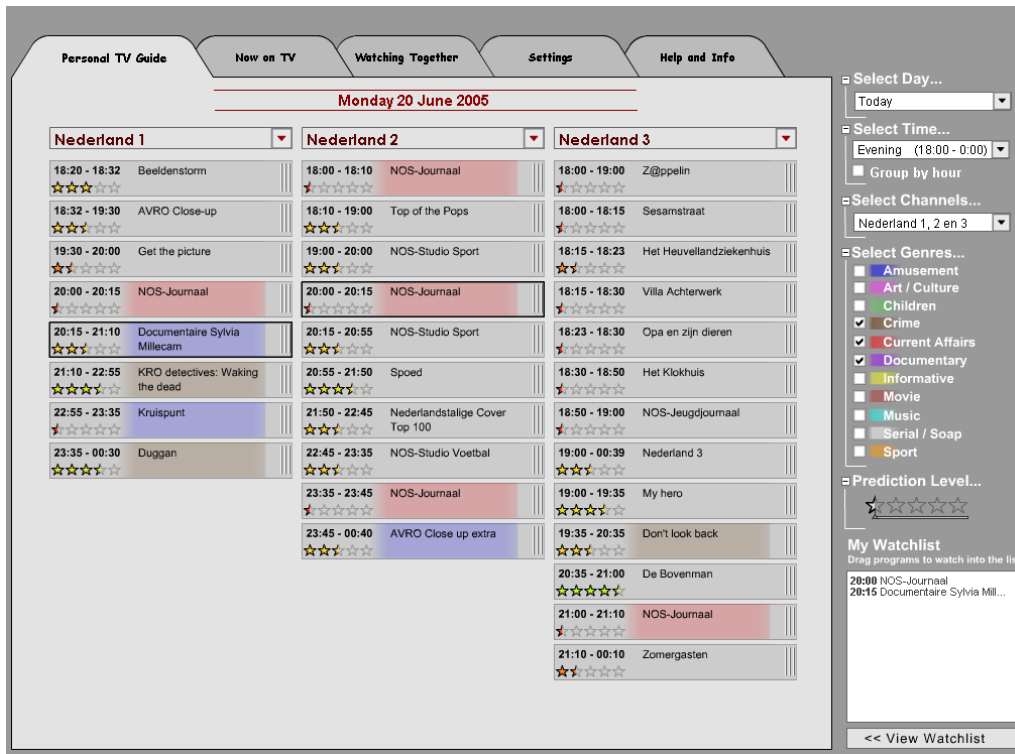


: pop-up for feedback and explanations



pop-up of detailed TV program description.

PROTOTYPE SCREENSHOTS



Improved prototype before usability tests: main screen

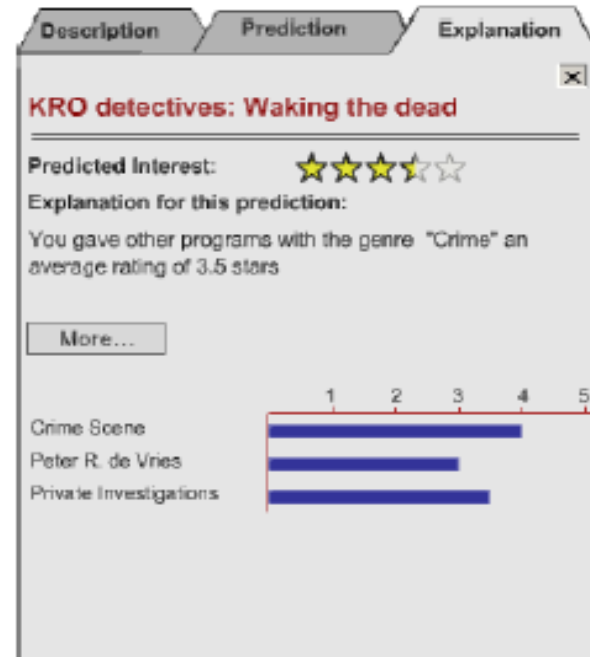


Improved prototype before usability tests: pop-up with detailed TV program description tab active

PROTOTYPE SCREENSHOTS



Improved prototype before usability tests:
pop-up with prediction tab active



Improved prototype before usability tests:
pop-up with explanation tab active

Merci de votre attention!

Question?