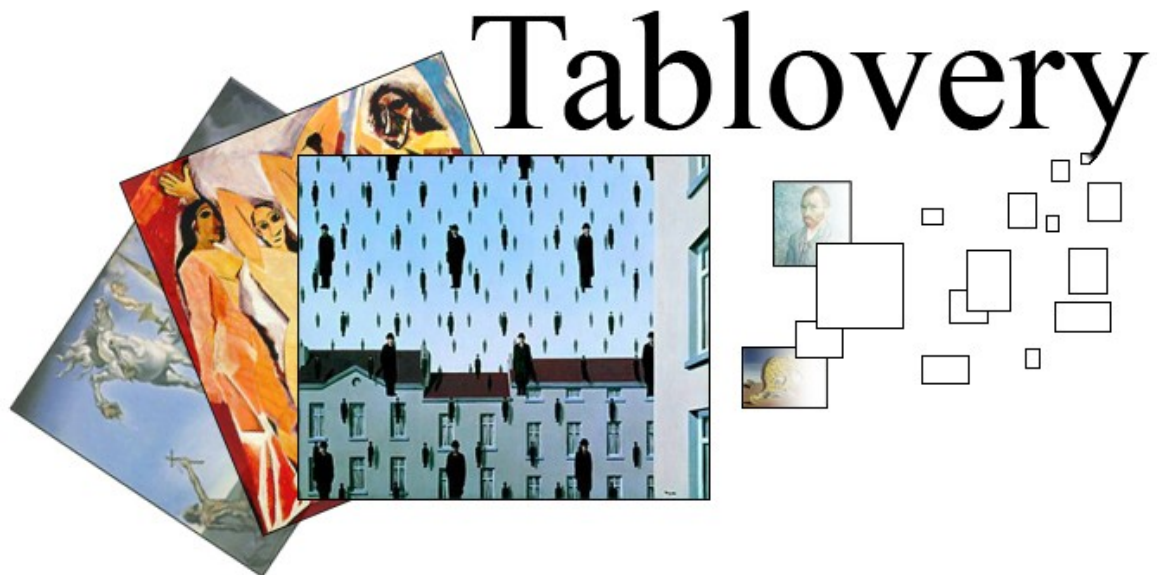
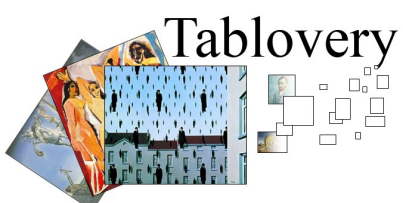


Charles Emmanuel Farkas
Christophe Guyard
Alan Marie
Romain Sechan

MIAGE Paris 5



2006-2007



<u>1. <i>Tablovery</i>.....</u>	<u>3</u>
<u>2. <i>Un site internet</i>.....</u>	<u>4</u>
Pourquoi un site Internet	4
Les fonctionnalités du site	4
<u>3. <i>Le clustering</i>.....</u>	<u>5</u>
Le cluster	5
Le cheminement	6
<u>4. <i>Les évolutions envisagées</i>.....</u>	<u>8</u>
Une inscription	8
Ajout de critères variables	8
Ajout de tableaux	8
Extension à d'autres formes d'art	8
<u>5. <i>Le bilan</i>.....</u>	<u>9</u>
<u>6. <i>Annexe</i>.....</u>	<u>10</u>

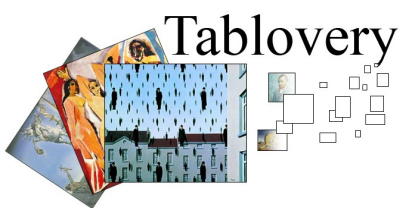
1.Tablovery

Après avoir étudié de manière théorique les concepts du Data Mining, il est intéressant d'en observer les utilisations et de mettre en application les connaissances acquises. Le Data Mining est une science fréquemment utilisée dans la médecine ainsi que dans l'analyse de données industrielles et économiques. Mais le Data Mining permet aussi la mise en place de projets culturels et artistiques très intéressants puisqu'il rend possible la découverte de similitudes non évidentes entre différents objets. Plusieurs projets actuellement en développement tels que Musicoverly, ou Steve (<http://www.steve.museum/>) en sont une très bonne illustration.

Musicoverly propose aux internautes une banque de musique très étendue que l'on parcourt en fonction de choix d'ambiances, il est alors possible d'écouter un ensemble de titres non pas par choix précis de tel ou tel artiste, mais en fonction de notre humeur, de nos envies.

Steve propose quant à lui de créer des salles virtuelles d'exposition en fonction des goûts et des appréciations des internautes. Ce type de projets permet de briser les classifications et postulats qui classifient souvent de manière rigide le milieu artistique.

C'est dans cette optique qu'est fondé le projet Tablovery. L'histoire de l'art a déjà classifié les œuvres de manière académique. Par courants, par techniques, par époques. Un site Internet tel que Tablovery a pour ambition de reclasser les œuvres en fonction de la réaction des internautes plutôt que sur ces classifications académiques. Avec la possibilité de donner leur avis sur les tableaux, et d'influencer de cette manière les classifications, les internautes vont contribuer à la création d'une réelle encyclopédie artistique originale, une sorte de Wikipedia de l'art permettant à tous de se cultiver et d'apprendre, et ainsi de répondre aux objectifs d'une science telle que le Data Mining.



2.Un site internet

Pourquoi un site Internet

Tout dans ce projet tourne autour de la notion de partage et si avons choisi de classifier des tableaux c'est parce qu'une analyse sur des œuvres d'art permet de répondre totalement aux objectifs du Data Mining. En effet, le Data Mining a une vocation éducative et vise à éclairer des similitudes cachées, à classifier selon des critères non évidents des éléments semblables.

L'art fait appel aux sentiments, aux émotions, au vécu, et, est ainsi très difficile à classifier. Certes, il existe déjà une classification académique des tableaux, classification arbitraire que l'on retrouve généralement dans les livres, dans les musées, dans les galeries etc. Notre but est donc d'outrepasser ces classements conventionnels pour faire intervenir dans la classification le facteur émotion des internautes.

Le choix d'une application Internet va donc de paire avec le projet Tablovery dans le sens où Internet est aujourd'hui le moyen de communication qui permet de toucher le plus grand nombre de personnes.

Les fonctionnalités du site

Consulter des tableaux en fonction de critères définis par tous

Le but de Tablovery est de permettre aux internautes de consulter un panel important de tableaux de maîtres, en les regroupant en fonction de critères d'ambiance sélectionnés lors de la recherche. Ces critères illustrent les sensations et sentiments que peuvent susciter les tableaux, tels que la tristesse, la joie, le calme, ou la peur. Il suffit de sélectionner et de situer sur une échelle de 1 à 100 un ou plusieurs de ces critères de sélection, et de lancer la recherche pour obtenir en réponse un ensemble de tableaux qui se rapprochent le plus possible des choix énoncés.

Possibilité de recherche concernant un auteur ou un mouvement en particulier

Pour répondre aux besoins de recherches précises, l'application permet aussi de préciser certains critères de recherche dits "fixes". Ainsi, il est possible de limiter la recherche à une époque, un style picturale, un artiste, ou une technique. Ces critères font référence à des données en dur, associées aux tableaux dans la base de donnée lors de leur insertion.

Notation et avis personnel sur les tableaux consultés

Une fois les tableaux trouvés suite à la recherche, une miniature de chacun est affichée à l'internaute. Il a la possibilité de connaître des détails concernant un tableau soit en passant la souris sur sa miniature soit en cliquant directement dessus. Le clic entraîne l'affichage du détail du tableau, avec l'image complète, ses informations, ses commentaires si il y en a, ainsi que ses critères évalués.

L'internaute peut modifier et enrichir les commentaires (comme il le ferait sur des encyclopédies en ligne comme Wikipedia), et peut indiquer sa propre évaluation par critères. Pour cela, il dispose des critères déjà créés par les autres membres de la communauté.

Un tel site est intéressant dans la mesure où il est en constante évolution. En effet, chaque tableau peut être mis à jour par chaque internaute, de plus les notations évoluent constamment ce qui fait que chaque recherche, chaque visite du site est différente d'une autre et répond ainsi parfaitement au principe de découverte du Data Mining.

3. Le clustering

Définition :

Le clustering (partitionnement de données) est une méthode [statistique](#) d'[analyse de données](#) qui a pour but de regrouper un ensemble de données en différents paquets homogènes, en ce sens que les données de chaque sous-ensemble partagent des caractéristiques communes, qui correspondent le plus souvent à des critères de proximité que l'on définit en introduisant des mesures de [distance](#).

L'utilisation du clustering nous a permis de répondre à deux problématiques :

- Comment offrir aux internautes la possibilité de se renseigner sur des tableaux répondant à leurs critères émotionnels ?
- Comment ajouter à ce nouvel outil une touche d'inconnue et de découverte ?

Pour cela nous avons utilisé la notion de flou. La notion de flou fait intervenir des calculs de similarités différents (similarité floue). Cela permet de classer les données selon des critères plus qualitatifs que quantitatifs. Dans notre cas, nous ne classons pas les tableaux avec 1 ou 2 critères mais avec un nombre n , indéterminé, dépendant du nombre de critères possible, ainsi que du choix de l'internaute. Faisant alors de la similarité un calcul plus ou moins complexe selon les souhaits de l'utilisateur.

La notion de flou est également représentée par le choix d'une valeur comprise entre 0 et 100 pour un critère émotion et non uniquement une réponse comme « oui » ou « non » :

Tristesse :

Les tableaux de notre base appartiennent alors à un espace de dimension n , espace dans lequel nous devons classer les tableaux. Un internaute pourra par exemple demander une classification répondant aux critères de joie, d'énergie, de paix et de chaleur, tandis qu'un autre s'intéressera plus à l'énergie et la tristesse.

Le cluster

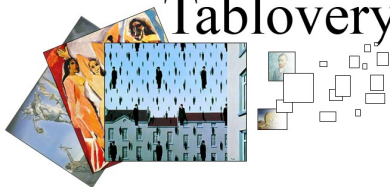
Le but :

- Découvrir les tableaux répondants aux mêmes critères, et de la même façon
- Visualisation de la distance qui sépare deux tableaux du même cluster

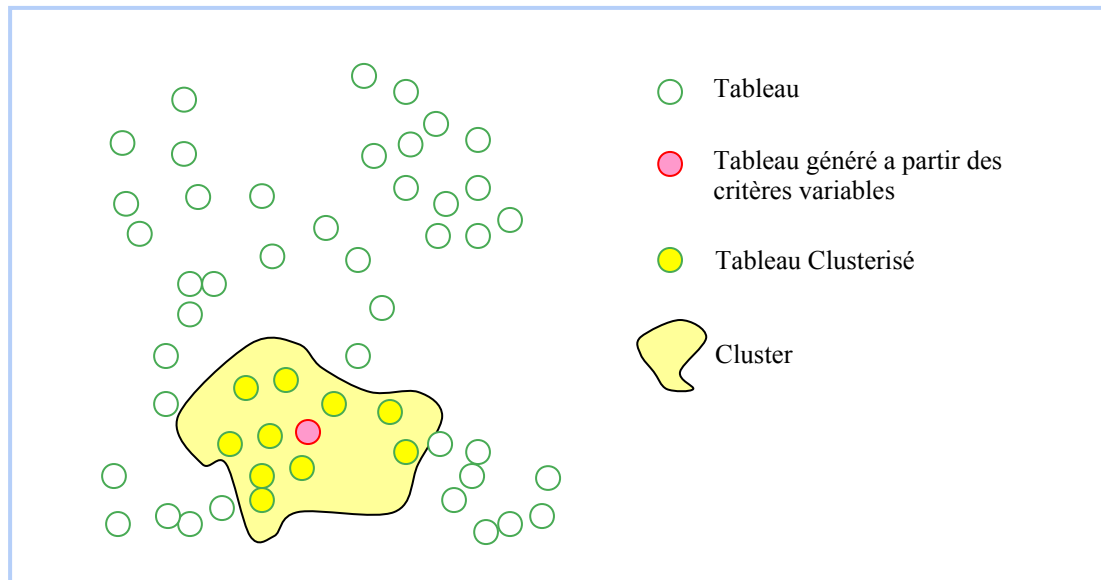
Cet algorithme se base uniquement sur les critères émotions sélectionnés par les utilisateurs. Les critères dits « fixes » auront permis d'appliquer au préalable un filtre sur les données.

Le fonctionnement de l'algorithme :

Tablovery



- **Pas 1** : À partir des n critères variables création d'un tableau dit « de référence » par rapport auquel seront comparés l'ensemble des tableaux de la base dans l'optique de créer le cluster.
- **Pas 2** : Pour chaque tableau de la base
 - Pas 3 : Calcul de la similarité du tableau avec le tableau généré sur les n critères de l'internaute
- **Pas 3** : On conserve les 8 tableaux les plus proches représentant notre cluster



Le cheminement

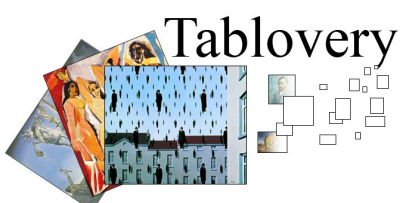
Le but :

- Rapprocher des tableaux qu'une classification conventionnelle n'aurait jamais rapprochés.
- Découvrir de nouveaux tableaux non semblables « à priori »
- Créer la notion de cheminement parmi l'ensemble des tableaux de la base.

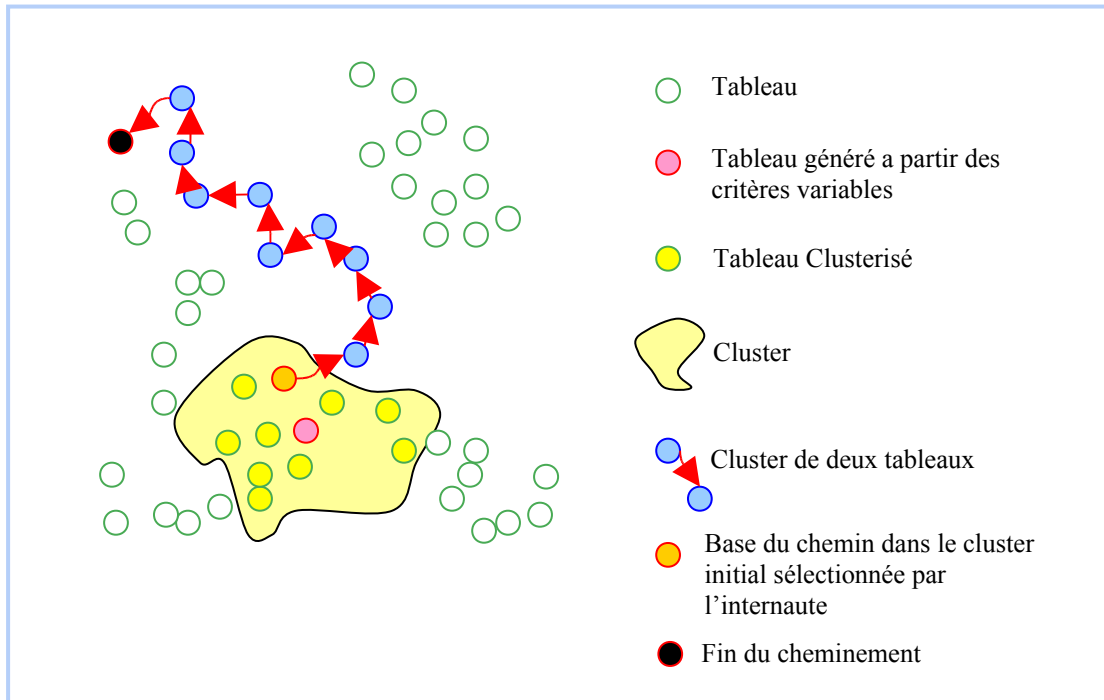
Ce principe de cheminement utilise également le principe de clustering vu précédemment. Cependant, son utilisation est différente. Là où chaque tableau était comparé à un tableau dit « de référence » dans l'algorithme précédent, chaque tableau est ici comparé à un tableau existant.

Le fonctionnement de l'algorithme :

- **Pas 1** : Sélection d'un tableau dans le cluster que l'on utilisera comme base du cheminement. L'ensemble des tableaux déjà présents dans le cluster sont gardés en mémoire pour ne pas les re-comparer avec la base.
- **Pas 2** : On applique l'algorithme vu précédemment à l'ensemble des tableaux restant avec comme tableau « de référence » le tableau choisi par l'internaute
- **Pas 3** : On conserve l'unique tableau ayant la plus grande similitude avec le tableau « de référence ».



- **Pas 4 :** On exclut l'ancienne base et le nouveau tableau obtenu devient à son tour le tableau « de référence ». On reprend au pas 2



4. Les évolutions envisagées

Une inscription

De manière à renforcer le principe de communauté il serait intéressant d'ajouter une phase d'inscription (avec login et mot de passe). Ceci également dans un souci de sécurité par rapports aux modifications apportées par les internautes.

Ajout de critères variables

Il faut encore ajouter la fonctionnalité d'ajout de critères variables qui permettra aux internautes de créer leurs propres critères variables afin de classer à leur manière les tableaux qu'ils consulteront. Ceci dans le but d'inciter la participation des internautes au contenu.

Ajout de tableaux

Il faut aussi développer la fonctionnalité d'ajout de tableaux qui va permettre aux internautes de soumettre de nouveaux tableaux, qui seront par la suite validés et insérés par les administrateurs du site. Ceci toujours dans le but d'inciter la participation des internautes au contenu.

Extension à d'autres formes d'art

Le site propose actuellement uniquement de classer et de consulter des tableaux de maîtres. Mais dans l'optique d'étendre cette encyclopédie de l'art qu'est Tablovery à l'ensemble de l'art, il serait intéressant d'étendre son contexte à la BD, la sculpture, l'architecture, etc.

5. Le bilan

Arrivé au terme du délai de réalisation du projet, il s'avère que certaines choses restent encore à développer, mais les fonctionnalités principales sont opérationnelles et permettent déjà d'exploiter l'encyclopédie d'art avec nos données de test.

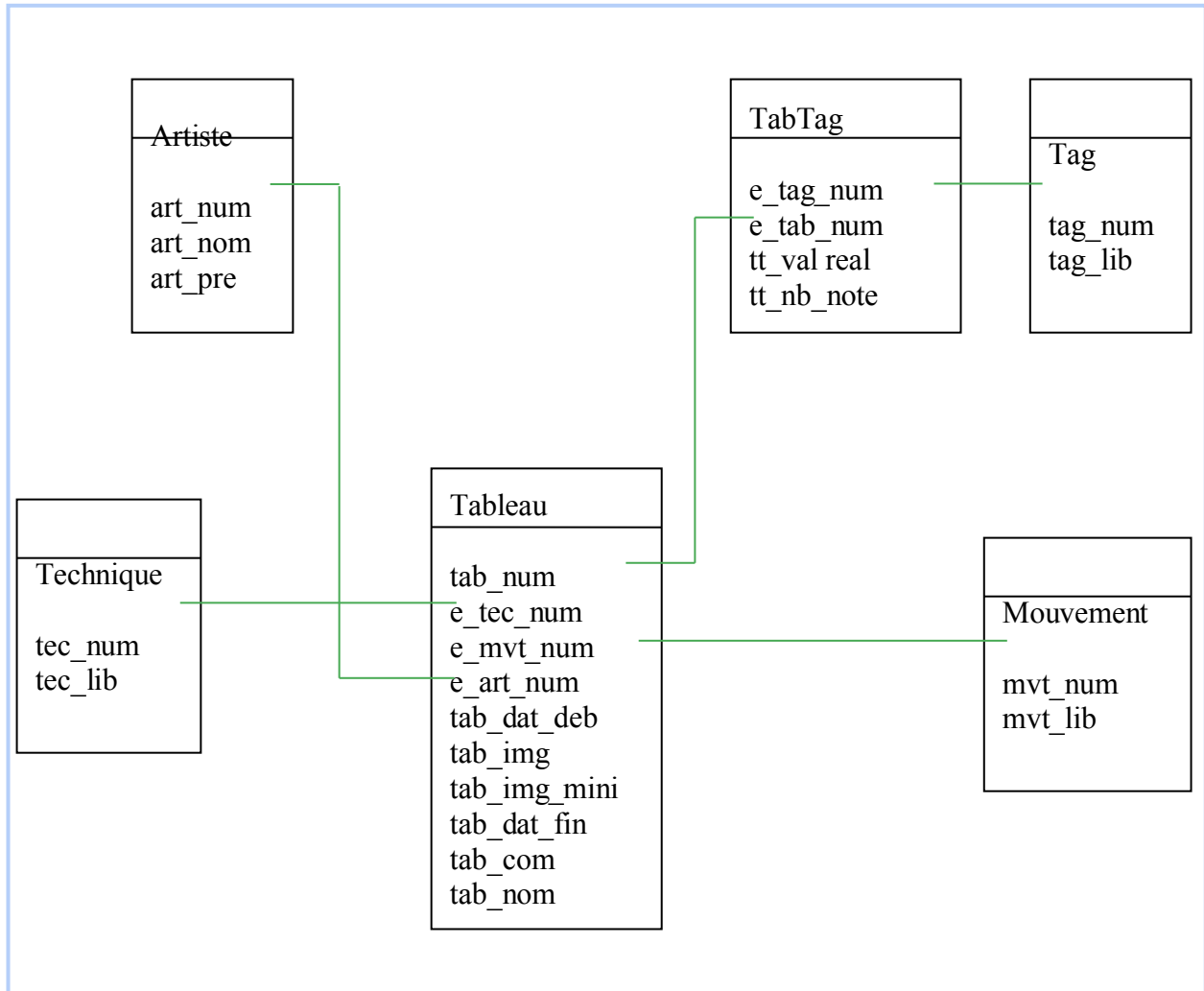
Le projet Tablovery résulte de deux principaux postulats, selon lesquels :

- Aujourd'hui apparaissent sur Internet de plus en plus d'outils d'apprentissage communautaires tels que Wikipedia, où chaque internaute a la possibilité de contribuer au savoir commun.
- Le data mining a une vocation de découverte de sens cachés, de similitudes et de rapprochements nouveaux et exploitables

Ces au regards de ces principes qu'est né l'idée de créer une encyclopédie communautaire de l'art, pour le moment limité à l'art picturale, mais qui vise à s'étendre à l'art en générale. C'est une première étape dans une manière nouvelle de classifier l'art, qui demandera l'implication des internautes dans un projet dont l'essor dépendra uniquement de leur motivation, et leur utilisation du site.

6. Annexe

Base de données :



```
--  
PostgreSQL database  
--
```

```
CREATE TABLE artiste (  
  art_num serial NOT NULL,           //numéro de l'artiste (p_key)  
  art_nom character varying NOT NULL, //nom de l'artiste  
  art_pre character varying NOT NULL //prénom de l'artiste  
);  
  
CREATE TABLE mouvement (  
  mvt_num serial NOT NULL,           //numéro du mouvement (p_key)  
  mvt_lib character varying NOT NULL //libellé du mouvement  
);
```

```

...

CREATE TABLE tableau (
  tab_num serial NOT NULL,           //numéro du tableau (p_key)
  e_tec_num integer,                //numéro de technique
(e_key)
  e_mvt_num integer,                //numéro du mouvement (e_key)
  e_art_num integer NOT NULL,       //numéro de l'artiste (e_key)
  tab_dat_deb integer NOT NULL,     //date de début de création
  tab_img character varying NOT NULL, //nom de l'image complète
  tab_img_mini character varying NOT NULL, //nom de l'image réduite
  tab_dat_fin integer,              //date de fin de création
  tab_com character varying,        //commentaire des internautes
  tab_nom character varying NOT NULL //nom du tableau
);

//La table tabtag regroupe le détail des critères variables
//définis pour chaque tableaux
CREATE TABLE tabtag (
  e_tag_num integer NOT NULL,       //numéro du tag (e_key)
  e_tab_num integer NOT NULL,       //numéro du tableau (e_key)
  tt_val real,                       //valeur moyenne des votes pour le
                                     //tableau associé
  tt_nb_note integer                //nombre de vote pour ce tag
);

//Un tag est un critère variable utilisé pour classer les tableaux
CREATE TABLE tag (
  tag_num serial NOT NULL,          //numéro du tag (p_key)
  tag_lib character varying NOT NULL //libellé du tag
);

CREATE TABLE technique (
  tec_num serial NOT NULL,          //numéro de la technique (p_key)
  tec_lib character varying NOT NULL //libellé de la technique
);

```