



# Discovery Hub

Explications graphiques des réponses fournies par un moteur de découverte d'informations

Commanditaires : Damien Legrand, Nicolas Marie (Inria/I3S, Wimmics)

S15 – Master 2 IFI  
Groupe 5  
Baptiste Comet  
Lucas Crivelli  
Emilie Palagi  
Mathieu Rul

CEIHM - Rapport intermédiaire  
Tuteur : Alain Giboin  
Lecteur : Anne-Marie Déry-Pinna  
Rendu le 8 décembre 2013

# Tables des matières

|      |  |           |
|------|--|-----------|
| I.   | <b>Introduction.....</b>   | <b>4</b>  |
| II.  | <b>Partie 1: Analyse de l'existant .....</b>                             | <b>5</b>  |
|      | Présentation de Discovery Hub .....                                      | 5         |
|      | Recensement des systèmes existants .....                                 | 10        |
|      | Discovery Hub .....  | 10        |
|      | RelFinder .....  | 11        |
|      | Pearltrees .....   | 12        |
|      | LinkedIn InMaps .....  | 13        |
|      | Identification des éléments à améliorer.....                             | 14        |
|      | Éléments à conserver .....   | 14        |
|      | Éléments à modifier .....  | 14        |
|      | Éléments à ajouter .....   | 15        |
|      | Remarques .....  | 15        |
| III. | <b>Partie 2: Procédures/techniques de récolte des informations .....</b> | <b>16</b> |
|      | Entretiens.....  | 16        |
|      | Procédures d'entretien .....   | 16        |
|      | Analyse des entretiens .....   | 16        |
|      | Questionnaires .....   | 17        |
|      | Rédaction du questionnaire .....   | 17        |
|      | Analyse du questionnaire .....   | 18        |
| IV.  | <b>Partie 3: Modélisation.....</b>                                       | <b>22</b> |
|      | Modélisation des utilisateurs .....                                      | 22        |
|      | Persona 1.....   | 22        |
|      | Scénarios du persona 1 .....   | 23        |
|      | Persona 2.....   | 23        |
|      | Scénarios du persona 2 .....   | 24        |
|      | Remarques .....  | 24        |

|     |  |           |
|-----|--|-----------|
|     | Analyse et modélisation de la tâche..... | 25        |
|     | Analyse de la tâche.....                 | 25        |
|     | Modélisation de la tâche.....            | 25        |
|     | Remarque.....                            | 27        |
|     | Modèle de tâche du système.....          | 28        |
| V.  | <b>Conclusion.....</b>                   | <b>29</b> |
| VI. | <b>Annexes.....</b>                      | <b>30</b> |
|     | Planning.....                            | 30        |
|     | Procédure entretien.....                 | 32        |
|     | Questionnaires.....                      | 33        |

## Introduction

---

- Auteurs: Emilie Palagi
- Relecteurs: Lucas Crivelli, Mathieu Rul

Discovery Hub est un moteur de recherche exploratoire qui permet à l'utilisateur d'élargir ses connaissances en proposant des résultats inattendus sur un éventail de domaines comme la culture (cinéma, musique, littérature), la politique, l'automobile, etc.

Discovery Hub est basé sur la base de triplets DBpedia fournie par Wikipédia, et permet également de rediriger l'utilisateur sur d'autres plates-formes comme YouTube et Deezer.

Ce moteur de recherche a la particularité de proposer une explication aux résultats affichés de trois manières différentes : des explications en catégorisant les résultats en fonction de leurs caractéristiques communes à la recherche ; en fournissant un cours résumé issu de Wikipédia ; en affichant une représentation graphique des résultats sous forme de graphe.

Dans le cadre de notre projet, nous nous intéresserons à la troisième explication, c'est-à-dire le graphe. Il s'agit ici de soulever les problèmes de cette interface graphique à l'aide d'entretiens utilisateurs et la diffusion d'un questionnaire. La première partie de notre dossier portera sur l'analyse de l'existant ; une deuxième partie sera composée des procédures d'entretien et questionnaire ainsi que l'analyse des résultats obtenus ; et enfin, la troisième partie présentera une description du modèle de la tâche, les personas et leurs scénarios, élaborés à partir des données recueillies des entretiens et du questionnaire.

## Partie 1: Analyse de l'existant

---

- Auteurs: Mathieu Rul
- Relecteurs: Lucas Crivelli, Emilie Palagi

Afin d'analyser l'existant, nous allons d'abord présenter Discovery Hub, puis nous le comparerons aux autres solutions existantes.

### Présentation de Discovery Hub

Discovery Hub est un moteur de découverte créé dans le cadre d'une thèse au sein de l'INRIA (*Institut National de Recherche en Informatique et en Automatique*). Il a été développé par Nicolas Marie (chercheur à l'INRIA) pour la partie moteur de recherche et par Damien Legrand (ancien étudiant de l'université de Nice Sophia Antipolis) pour l'interface graphique.

Ce moteur permet d'effectuer une recherche sur un ou plusieurs sujets différents afin d'en découvrir les points communs et obtenir des suggestions en rapport avec ces thèmes. Les données utilisées par Discovery Hub sont issues du projet universitaire DBpedia qui vise à structurer, organiser et formaliser de manière sémantique les données de l'encyclopédie libre Wikipédia. Pour cela, DBpedia se repose sur l'exploitation des données écrites dans le langage RDF (*Resource Description Framework*) développé par le W3C. Ce langage est un formalisme visant à décrire une ressource web. Plus précisément, une ressource web peut contenir plusieurs triplets RDF qui sont l'association d'un sujet, d'un prédicat et d'un objet. Le sujet indique la ressource à décrire, le prédicat définit les propriétés applicables à cette ressource et enfin l'objet représente la valeur d'une propriété.

Une première version de l'interface web de Discovery Hub développée en HTML5 a vu le jour en 2012, puis une refonte en une deuxième version a été lancée courant 2013. La version actuelle est accessible depuis un navigateur Internet récent, mais en revanche ne dispose pas encore de version optimisée pour les mobiles.

Le site propose une barre de recherche qui permet à un utilisateur de saisir le ou les thèmes qu'il souhaite aborder.

Discovery Hub.

Q south park

**South Park** Television Show  
South Park is an American animated sitcom created by Trey Parker and Matt Stone for the Comedy Central television network. Intended for mature audi...

**South Park: Bigger, Longer & Uncut** Film  
South Park: Bigger, Longer & Uncut is a 1999 American animated satirical comedy/musical film based on the animated television series South Park...

**South Park Mexican** Artist  
Carlos Coy (born October 5, 1970), better known by his stage name South Park Mexican, is an American rapper, founder of Dope House Records, and con...

**South Park (Colorado basin)** Other  
South Park is a high intermontane grassland basin, approximately 10,000 ft (3,000 m) in elevation, in the Rocky Mountains of central Colorado. It e...

**South Park (season 13)** Other  
The thirteenth season of South Park originally aired in the United States on Comedy Central between March 11 and November 18, 2009. The season was ...

**South Park Blocks** Park  
The South Park Blocks form a city park in downtown Portland, Oregon. The Oregonian has called it Portland's "extended family room", as Pioneer Cour...

**South Park (season 1)** Other  
Season one of South Park, an American animated television series created by ...

All Album Artist Band Book Film Other Park Place Software

Television Show  
South Parks Road  
South Parks Road is a road in Oxford, England. It runs east-west past the main ...

News Random

lopedia on the web,  
to find what you are  
you ! It allows  
s. As a hub, it proposes  
outube, Deezer and

Get recommendations on your interests

ation related to writers you like ? Want to discover bands at the crossroad of  
and rock record-labels you like ? Interested by more complex and composite  
sed on your deepest interests : a writer, a film and a band combination ? Or  
impler ? If you have a thirst for discovery and knowledge, Discovery Hub has  
answers for you.

Understand a recommendation

Figure 1 : Exemple de saisie sur Discovery Hub

Une fois la recherche lancée, les résultats suggérés apparaissent de manière regroupée par catégorie (par exemple : personne, géographie, entreprise ...), chaque élément de réponse est cliquable afin d'obtenir plus d'information.

Discovery Hub.

Q Start your exploration here...

News Random

**430 topics to explore**  
**16 kinds of topics**





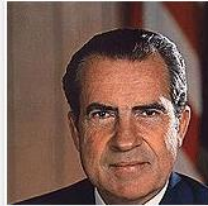

[Top](#)
[Building](#)
[Company](#)
[Election](#)
[Film](#)

[Government Agency](#)
[Governor](#)
[Military Person](#)
[Military Unit](#)

[Miscellaneous](#)
[Office Holder](#)
[Others](#)
[Politician](#)
[School](#)

[Ship](#)
[Writer](#)

★ **Top**

Handedness Of Presidents Of The United States   
 Thomas E. Stephens   
 Franklin D. Roosevelt   
 John F. Kennedy   
 Richard Nixon   
 George H. V

**Building**














Figure 2 : Résultats d'une recherche sur Discovery Hub

Le détail d'un résultat indique un résumé de Wikipédia, des vidéos YouTube, des musiques, des images ou encore des informations précises, par exemple pour un président : son successeur, son vice-président...

Discovery Hub.

Franklin D. Roosevelt

## Franklin D. Roosevelt

**Why Franklin D. Roosevelt is recommended ?**

★ Add to Favorites   ■ Add to a Collection   🔍 Run an exploration

**Vice President**

- Harry S. Truman
- Henry A. Wallace
- John Nance Garner

**Successor**

- Harry S. Truman
- Herbert H. Lehman

**Religion**

- Episcopal Church (United States)

**Child**

- Franklin D. Roosevelt
- Franklin Delano Roosevelt, Jr.
- Anna Roosevelt Halsted
- John Aspinwall Roosevelt
- Elliott Roosevelt
- James Roosevelt

**Source**

**Franklin Delano Roosevelt** also known by his initials, FDR, was the 32nd President of the United States (1933–1945) and a central figure in world events during the mid-20th century, leading the United States during a time of worldwide economic crisis and world war. The only American president elected to more than two terms, he facilitated a durable coalition that realigned American politics for decades. FDR defeated incumbent Republican Herbert Hoover in November 1932, at the depths of the Great Depression. FDR's persistent optimism and activism contributed to a renewal of the national spirit.

▶ Youtube's Videos For Franklin D. Roosevelt And Dwight D. Eisenhower

▶ Youtube's Videos For Franklin D. Roosevelt And Harry S. Truman

Figure 3 : Détail d'un résultat



De plus, il est possible de connaître pourquoi un résultat a été recommandé en fonction des éléments de recherche saisie. Cette information est représentée sous la forme d'un graphe, où les thèmes recherchés sont à l'extrémité de celui-ci. Les nœuds sont libellés avec les noms de l'entité qu'il symbolise et les arêtes relient les éléments recherchés en passant par les sujets connexes.

Discovery Hub.

Franklin D. Roosevelt

## Franklin D. Roosevelt

**Why Franklin D. Roosevelt is recommended ?**

★ Add to Favorites   ■ Add to a Collection   🔍 Run an exploration

**Dwight D. Eisenhower**  
Dwight David "Ike" Eisenhower was a five-star general in the United States Army and the 34th President of the United States, from 1953 until 1961, and the last to be born in the 19th century. During World War II, he served as Supreme Commander of the Allied forces in Europe, with responsibility for planning and supervising the successful invasion of France and Germany in 1944–45, from the Western Front. In 1951, he became the first supreme commander of NATO.

**Vice President**

- Harry S. Truman
- Henry A. Wallace
- John Nance Garner

**Successor**

▶ Youtube's Videos For Franklin D. Roosevelt And Dwight D. Eisenhower

Figure 4 : Explication d'une recommandation d'un résultat

Ce graphe relie les thèmes recherchés au résultat sélectionné (ces nœuds sont mis en évidence par la couleur rose). Les autres nœuds (en gris) sont les sujets connexes qui relient les nœuds précédents. Il est possible de survoler un nœud afin de mettre en évidence les liens directs. Pour ce faire, le nœud survolé est cerclé en orange et les nœuds auxquels il est directement lié en gris s'ils sont roses ou cerclés en rose s'ils sont gris.

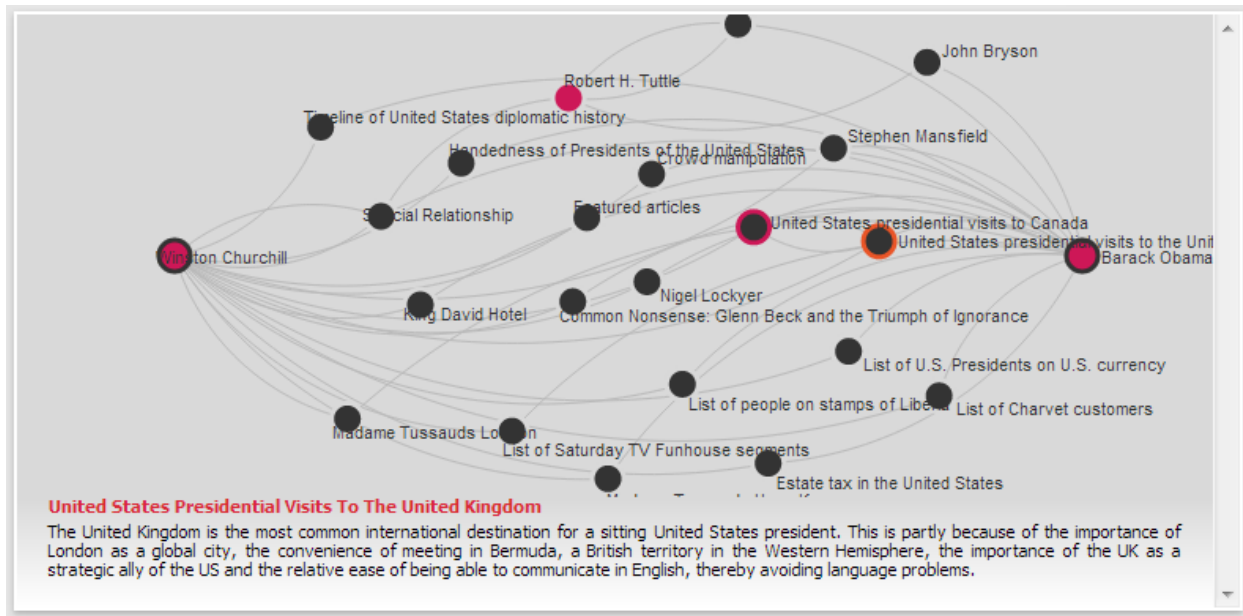


Figure 5 : Mise en évidence des liens directs

Notre projet a pour objectif d'identifier les problèmes de l'explication graphique, ainsi que de proposer des pistes d'amélioration de cette interface.

## Recensement des systèmes existants

Nous avons recherché des sites Internet ayant une interface semblable à celle de Discovery Hub. En effet, nous nous sommes concentrés sur les systèmes proposant eux aussi une représentation des données sous forme de graphe.

### Discovery Hub

Adresse web : <http://discoveryhub.co/>

But : proposer des liens entre différents sujets de recherche et suggérer des résultats connexes.

#### Fonctionnalités :

- Affiche l'ensemble des liens entre un ou plusieurs sujets en une réponse sous forme de graphe.
- Les nœuds peuvent être déplacés.

Éléments d'IHM : représentation graphique en 2D d'un réseau composé de thèmes ou d'objets où les nœuds peuvent être réorganisés.

#### Public visé :

- Curieux voulant découvrir de nouvelles choses.

- Étudiant désireux d'obtenir les liens entre plusieurs sujets.
- Chercheur souhaitant acquérir des ressources de travail.

#### Avantages :

- Les sujets recherchés sont mis en valeur par l'utilisation d'une couleur différente des autres nœuds.
- Les liens directs peuvent être mis en évidence.

#### Inconvénients :

- Les liens ne sont pas annotés, ce qui ne donne pas d'indication sur leurs natures.
- Lorsqu'il y a de nombreux éléments, le graphe devient brouillon et manque de clarté.

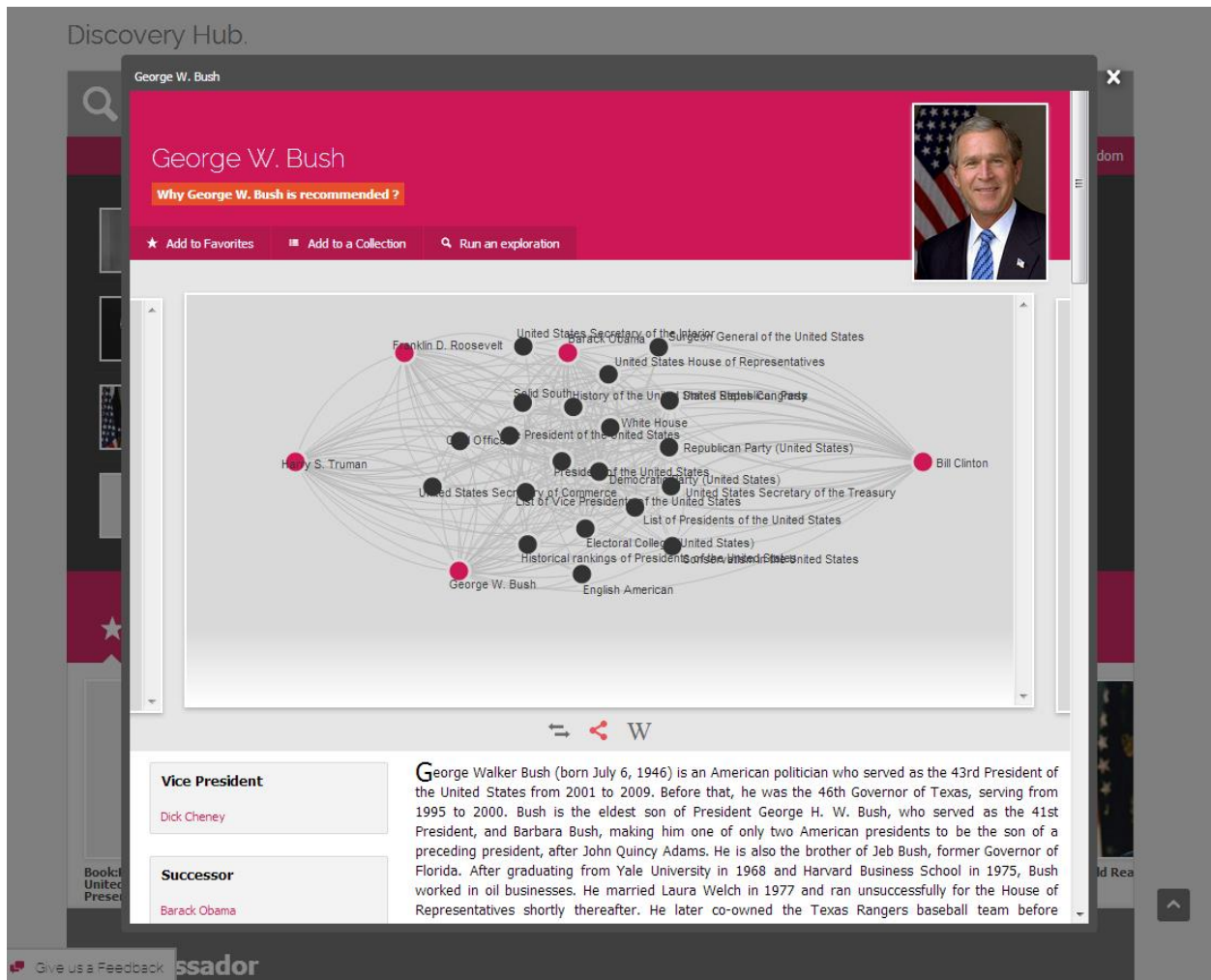


Figure 6 : Discovery Hub

#### RelFinder

Adresse web : <http://www.visualdataweb.org/relfinder.php>

But : mettre en évidence les relations entre deux éléments en se basant sur les données d'Internet décrit au format RDF (se base aussi sur DBpedia).

#### Fonctionnalités :

- Les liens entre les objets sont annotés.
- Affiche un ou plusieurs liens existants entre deux nœuds.

- Les résultats de recherche s'affichent au fur et à mesure de manière dynamique.

**Éléments d'IHM :** propose un affichage sous forme de graphe de liaison sémantique où tous les éléments peuvent être réorganisés.

**Public visé :** toutes les personnes souhaitant découvrir les liens sémantiques entre deux ressources recensées sur le web.

**Avantages :**

- Le système est efficace dans le sens où les liens entre les nœuds sont annotés et clairement explicités.
- Mise en évidence des liens et des nœuds reliant les entités.
- Le graphe est interactif.

**Inconvénients :**

- L'interface est un peu austère : il y a peu de contraste et de couleurs et pas beaucoup d'élément graphique mis à part les liens et les nœuds.
- Fonctionnalités limitées (par exemple : ne propose pas de lien vers Wikipédia).

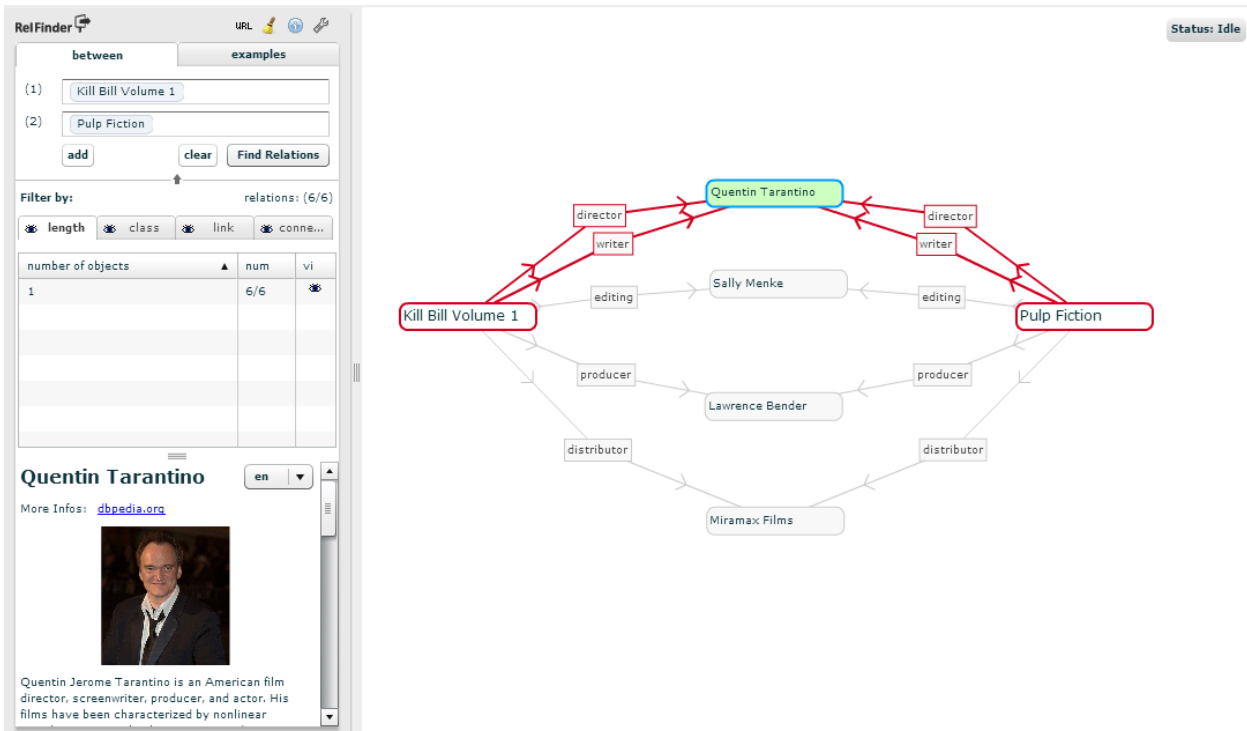


Figure 7 : RelFinder

**Pearltrees**

**Adresse web :** <http://www.pearltrees.com/>

**But :** permettre aux utilisateurs de collecter, d'organiser et de partager ce qu'il trouve sur Internet. Il peut s'agir de page web, de vidéo, de musique, etc. Chaque contenu trouvé constitue une perle.

**Fonctionnalités :**

- Classification des perles par centre d'intérêt.
- Suggestion de sujets connexes aux centres d'intérêt.

**Éléments d'IHM :** propose un affichage graphique sous forme de réseau de perles, où chaque perle peut être réorganisée.

**Public visé :** toutes les personnes voulant créer un réseau autour de leur centre d'intérêt.

**Avantages :**

- La représentation est très esthétique grâce à l'utilisation de photos, de couleurs et d'animations.
- Le graphe est interactif.

**Inconvénients :**

- Pas très ergonomique et peut devenir vite surchargé, notamment dû à une absence de filtrage.
- L'utilisation manque de clarté et n'est pas très intuitive.

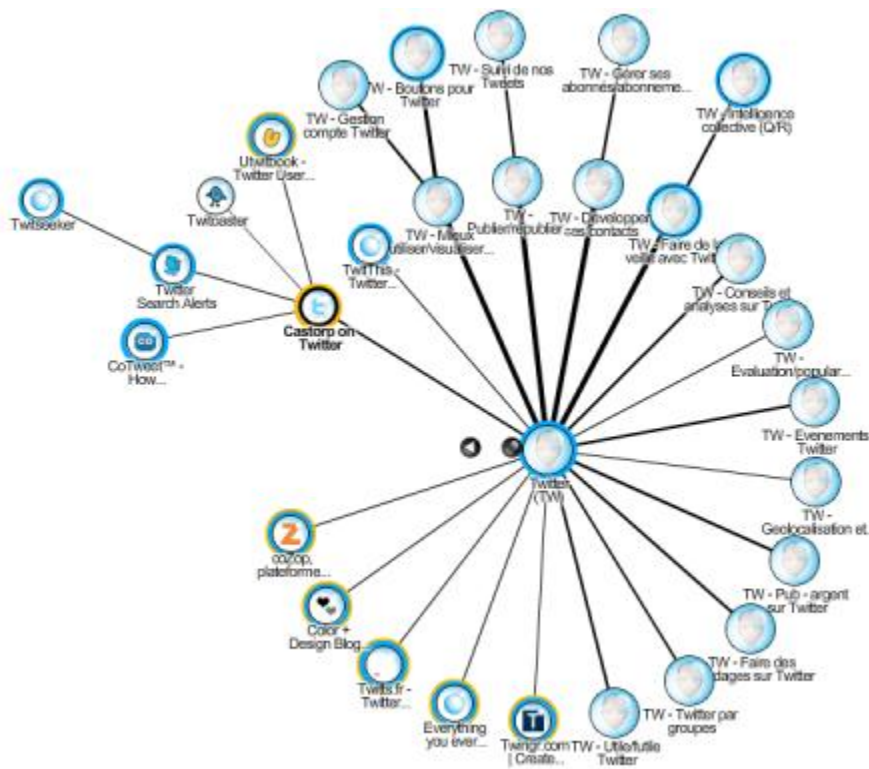


Figure 8 : Pearltrees

### LinkedIn InMaps

**Adresse web :** <http://inmaps.linkedinlabs.com/>

**But :** permettre aux utilisateurs du réseau social professionnel LinkedIn d'afficher leurs contacts sous la forme d'un graphe afin de voir les liens entre chaque personne du réseau.

**Fonctionnalités :**

- Affichage du nom du contact sous la forme d'un nœud
- Possibilité de zoomer et dézoomer sur le graphe
- Coloration des liens en fonction de leur nature (collaborateur, ancien camarade de promotion, autre relation professionnelle, etc.)

**Éléments d'IHM :** propose un affichage graphique sous forme de réseau composé de nœuds et de liens statiques.

**Public visé :** utilisateur de LinkedIn voulant visualiser de manière graphique leur réseau professionnel.

**Avantages :**

- Les résultats sont regroupés de manière cohérente.
- Le graphe est interactif.

- Une fois chargé, le système est réactif.
- Mise en évidence des informations importantes.

#### Inconvénients :

- Les résultats sont longs à charger lorsqu'il y a beaucoup de contacts.
- Le graphe devient trop brouillon et illisible lorsque les contacts sont nombreux.
- Les noms des contacts sont peu lisibles.

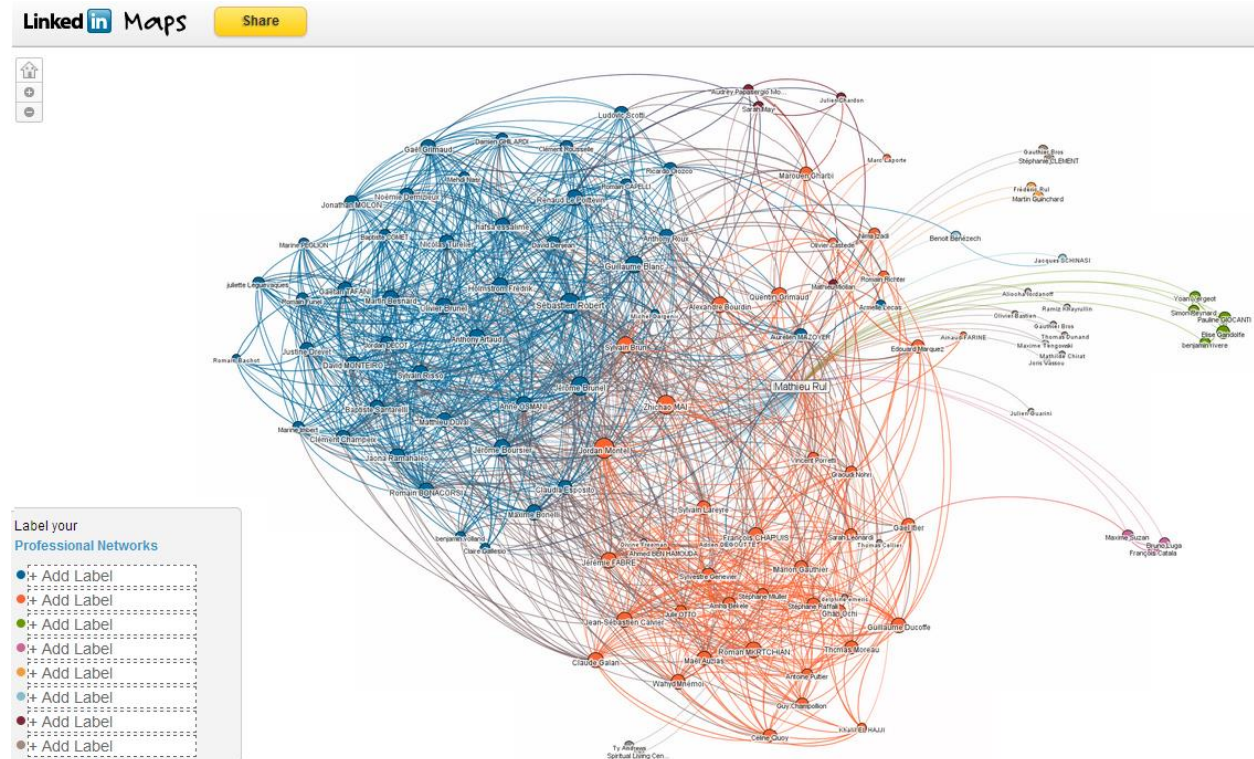


Figure 9 : LinkedIn InMaps

## Identification des éléments à améliorer

### Éléments à conserver

Les systèmes existants que nous avons vus se basent sur une représentation en réseau à deux dimensions. Le choix d'un graphe composé de nœuds reliés par des arêtes permet de représenter clairement les objets et de facilement identifier leurs liens. Cependant, plusieurs axes d'amélioration sont à explorer afin d'enrichir ce graphe.

### Éléments à modifier

Afin notamment d'enrichir les informations sur les liens entre les différents nœuds du réseau, l'ajout d'une annotation explicitant la nature du lien semble important.

### ***Éléments à ajouter***

Le recensement des interfaces existantes a permis d'identifier l'importance de la coloration des liens et des nœuds de même nature afin de grandement améliorer la lisibilité d'un graphe, comme sur LinkedIn InMaps par exemple. Il peut donc être intéressant de regrouper par couleur les différentes catégories du graphe (film, géographie, personne ...) en affichant bien sûr la légende correspondante. De plus, il serait judicieux de pouvoir filtrer les résultats par catégorie à partir de la légende.

Afin de pouvoir mesurer rapidement la force d'un lien entre deux objets, il serait bien de faire varier l'épaisseur du trait : ainsi, un trait fin représenterait une faible connexion, tandis qu'un trait épais montrerait un lien fort entre deux nœuds/entités.

Une autre piste d'amélioration est de rendre l'interface plus interactive, notamment pouvoir obtenir des informations plus complètes sur un nœud lorsque celui-ci est sélectionné, à la manière de RelFinder ou LinkedIn InMaps.

Afin de pouvoir faire varier le niveau de granularité des résultats et ainsi pouvoir obtenir plus de détails sur une partie du graphe, une fonction de zoom permettrait de rendre le graphe plus interactif et surtout d'en augmenter la lisibilité, tout en conservant la possibilité d'avoir une vision globale.

Enfin, il serait bien de pouvoir lancer une nouvelle recherche ou générer un nouveau graphe à partir d'éléments sélectionnés sur le graphe actuel.

### **Remarques**

Afin d'analyser l'existant, nous avons d'abord cherché à bien comprendre le site Discovery Hub existant. Pour cela, nous avons analysé les fonctionnalités et les éléments d'IHM du dispositif. Puis identifier les utilisateurs auxquels il s'adresse, en définissant leur but lorsqu'il utilise le dispositif et les tâches qu'ils peuvent réaliser. Enfin, nous avons analysé les interactions que l'utilisateur peut faire avec le système.

Puis dans un second temps, nous avons recensé les autres systèmes existants, afin de pouvoir les comparer à Discovery Hub et surtout d'en faire la critique afin d'en ressortir les avantages et les inconvénients. Cette analyse nous a permis d'identifier les éléments à conserver ou à améliorer à partir des points positifs des autres systèmes existants.

## Partie 2: Procédures/techniques de récolte des informations

---

- Auteurs: Baptiste Comet
- Relecteurs: Emilie Palagi, Mathieu Rul

Afin de mieux comprendre les problématiques de l'application Discovery Hub, notamment concernant la partie graphique sur laquelle nous travaillons, il est important d'apprendre à connaître l'utilisation qui en serait faite. Nous avons réalisé des entretiens avec des utilisateurs potentiels et diffusé un questionnaire afin d'identifier les forces et les faiblesses de la présentation graphique des résultats de recherche de Discovery Hub, et de mieux comprendre dans quel cadre l'application Discovery Hub sera utilisée.

### Entretiens

Interviewer des utilisateurs potentiels allait nous permettre d'avoir des retours directs sur l'application que nous présentons, en fonction de la personne interviewée (en fonction d'un type d'utilisateur). Il fallait pour cela interviewer des personnes correspondant aux utilisateurs types que nous avons imaginés à partir de l'analyse de l'existant, il s'agit ici d'étudiants, d'enseignants ou de chercheurs.

#### *Procédures d'entretien*

Tous nos entretiens suivaient une procédure préalablement établie, une méthodologie qui nous a permis de rester efficaces dans la requête d'informations dans un premier temps, puis dans la comparaison de ces informations.

Nous devons commencer par mettre en confiance l'interviewé en lui présentant clairement notre projet ainsi que la motivation de l'entretien en cours, puis en le laissant se présenter, afin d'instaurer un dialogue serein et de récolter des informations utiles. Nous demandons ensuite à l'interviewé de réaliser quelques tâches qui nous aideraient à comprendre l'utilisation qui est faite de Discovery Hub, et à identifier des problèmes ou des gênes rencontrés par l'interviewé.

Ces tâches mettent notamment en scène les différences d'utilisation entre l'outil Discovery Hub et d'autres outils de recherche de connaissance tels que Google ou Wikipédia. La tâche finale consistant à étudier le graphe de liaisons de Discovery Hub afin de retrouver les liens qui existent entre plusieurs entités d'une recherche (mots-clés et résultats). Le détail de la procédure des entretiens peut être trouvé en annexe.

#### *Analyse des entretiens*

Nous avons réalisé 5 entretiens, en interviewant 2 étudiants et 3 enseignants chercheurs. Voici les principales tendances qui se dégagent des informations que nous avons recueillies :



*À noter : il nous est impossible sur un tel échantillon d'interviewés de pouvoir dégager des tendances différentes pour les étudiants d'un côté et les chercheurs de l'autre, l'analyse de ces entretiens est alors générale.*

→ Tous les interviewés effectuent des recherches sur Internet, très majoritairement avec le moteur de recherche Google. La plupart d'entre eux (4 sur 5) s'intéressent également à la découverte de nouvelles informations liées à leurs centres d'intérêt. On peut alors en conclure que l'outil Discovery Hub répond bien à un besoin des utilisateurs, dont il faudra considérer les retours.

→ Dans l'ensemble, les interviewés ont trouvé l'interface du graphe de Discovery Hub trop peu intuitive. En effet, ils la qualifient de "trop brouillon" et déplorent que les liens entre les entités présentées ne soient pas assez explicites. On observe que seulement 2 des 5 interviewés ont su retrouver, tout en étant satisfaits, le lien entre deux entités de recherche dans le graphe, bien que ce lien existait dans tous les cas. Ce premier diagnostic souligne le fait que si l'outil Discovery Hub propose des solutions, il peine à les rendre accessibles auprès de ses utilisateurs.

→ Les interviewés ont pu suggérer divers axes d'améliorations dont les graphes de Discovery Hub pourraient bénéficier, les plus importants sont : la possibilité de "zoomer" sur un graphe afin de le rendre moins chargé et plus lisible ; la mise en valeur de certains liens grâce à une coloration spécifique en fonction de leur thème, donc la possibilité de filtrer par thème ; l'augmentation du contrôle par la souris, afin de réorganiser les données, les filtrer ou accéder à plus d'informations.

## Questionnaires

Les entretiens que nous avons réalisés nous ont fourni une première approche de l'utilisation du système de découverte Discovery Hub. Être capable d'observer l'utilisateur pendant la tâche et avoir une discussion directe avec lui ont été d'importants avantages pour la récupération d'informations relatives à notre projet. Cependant, cette méthode est évidemment limitée par son coût en temps, par la disponibilité et la proximité des interviewés, elle ne permet donc pas d'obtenir un grand nombre de résultats. Afin d'approfondir nos connaissances sur l'utilisation de Discovery Hub, nous sommes partis des premiers résultats obtenus par les entretiens et avons rédigé puis diffusé un questionnaire en ligne.

### *Rédaction du questionnaire*

Diffuser un questionnaire allait nous permettre d'identifier et quantifier dans un premier temps les types d'utilisateurs potentiels de Discovery Hub et leurs utilisations, l'objectif étant de modéliser l'utilisateur et la tâche. Pour cela, la première partie du questionnaire permet d'identifier la personne en fonction de son âge, sa catégorie socioprofessionnelle, mais également de ses habitudes face aux nouvelles technologies.

Une seconde partie nous permet de faire le lien entre l'utilisateur et la tâche, car elle nous apporte des informations plus précises quant à la curiosité des utilisateurs et leur comportement face aux moteurs de

recherche et à la recherche sur le web en général. Enfin, une dernière partie fait la présentation de graphes visibles sur Discovery Hub, et récolte des retours sur la compréhension et l'ergonomie apparente de ces graphes. Elle propose également au panel d'exprimer des remarques ou des idées d'améliorations sur la présentation graphique.

### ***Analyse du questionnaire***

Les retours du questionnaire permettent de réaliser une analyse plus quantitative des utilisateurs et des tâches. Il est intéressant d'analyser les comportements de nos lecteurs vis-à-vis de l'informatique et de la recherche sur le web.

#### **→ Le panel**

75% des réponses nous proviennent de jeunes étudiants (âgés de 20 à 29 ans) et 17% de chercheurs. Ces chiffres sont évidemment à mettre en corrélation avec l'environnement de notre groupe de projet constitué d'étudiants, mais il est important de les noter afin d'analyser la suite des informations recueillies. Bien qu'il soit d'un niveau informatique varié (les réponses couvrent équitablement l'ensemble des niveaux proposés), le panel affirme passer plus de 3h par jour sur un ordinateur (92% des réponses). On en déduit que même s'il n'est pas constitué d'experts, le panel est déjà très ouvert à l'utilisation de l'informatique

#### **→ Le comportement face aux recherches web**

Le panel nous montre que ses thèmes de recherches sur le web sont très variés, mais également que ces recherches se font autant dans un cadre professionnel que personnel. Aucune forte tendance ne se dégage donc de ces premiers résultats si ce n'est que le domaine de tâches pour des recherches web est très vaste.

Lors de recherche sur Google, 83% du panel privilégie une recherche rapide à une recherche approfondie, et 75% du panel souhaite obtenir uniquement des résultats en relation directe avec le sujet de la recherche. On aurait alors tendance à en déduire que la recherche "ouverte", la recherche de découverte proposée par Discovery Hub n'intéresse pas grandement notre panel. Cependant, ces observations sont liées à l'environnement de l'utilisateur, qui réalise une recherche sur Google, dont les objectifs diffèrent de ceux d'une recherche sur Discovery Hub. En effet, on note pour la question suivante que 75% du panel porte un intérêt aux liens existant entre une entité de recherche et les résultats proposés (des liens qu'ils pourront tenter de découvrir avec Google (pour 53%) ou Wikipédia (pour 35% d'entre eux)).

### → L'étude des graphes

Une recherche sur Discovery Hub a été faite, avec comme entités de recherche “Barack Obama” et “George W. Bush”, un résultat proposé était “Bill Clinton”. On a ouvert le graphe d'explication de ce résultat et obtenu cette représentation :

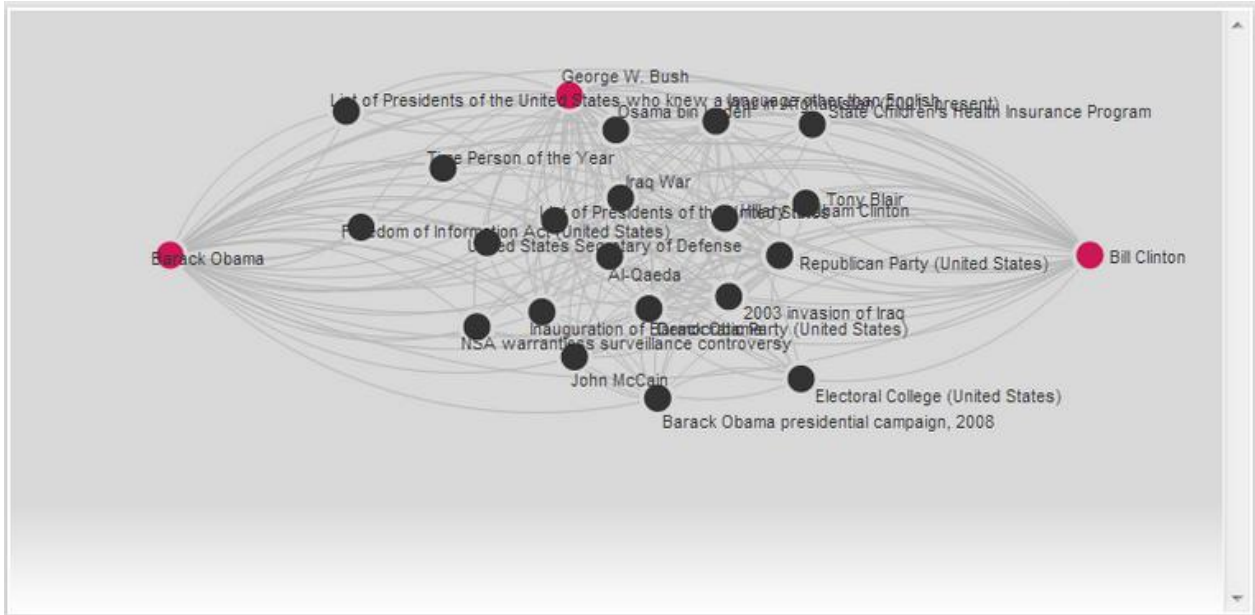


Figure 10 : Graphe du résultat pour “Bill Clinton”

En se basant sur cette représentation, 67% du panel trouve que le graphe est trop peu compréhensible, 50% du panel jugent également la pertinence des données comme mauvaise (auxquels s'ajoutent 42% du panel qui jugent cette pertinence “moyenne”). Enfin, selon plus de deux tiers du panel, c'est aussi la compréhension des relations entre les nœuds qui est difficile. Ces chiffres parlants sont la preuve que le système existant ne convient pas aux utilisateurs en termes d'ergonomie. Et si ces chiffres élevés semblent aussi être liés au fait qu'on ne présente qu'une capture d'écran statique et non la page de l'application elle-même, on sait que l'interaction avec la souris sur le graphe est faible. On juge alors que les appréciations d'utilisateurs seraient identiques.

On choisit ensuite de sélectionner le nœud “Democratic Party (United States)”, le graphe devient le suivant :

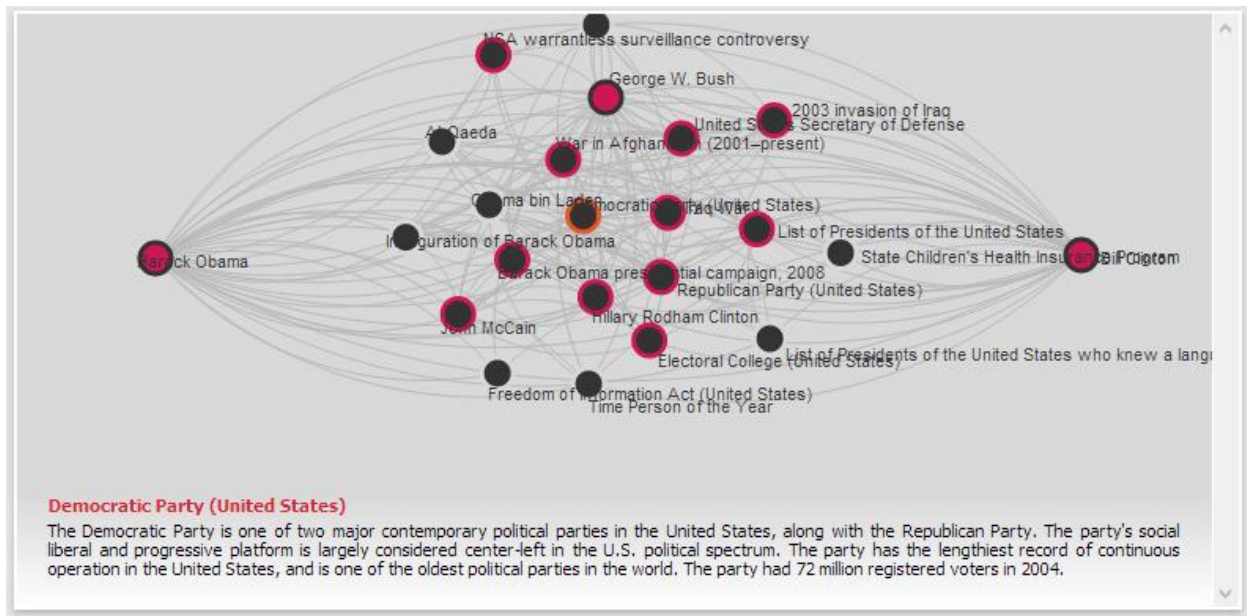


Figure 11 : Résultat avec le nœud “Democratic Party (United States)” sélectionné

En posant des questions sur la compréhension de la coloration des nœuds qui est apparue, on apprend maintenant que si 83% du panel fait bien la distinction graphique des différents nœuds, 67% n’en saisissent pas la signification totale (la compréhension du code couleur n’est pas complète). Globalement, le panel trouve la représentation des nœuds peu ou moyennement explicite. On remarquera cependant que l’impossibilité de jouer avec le graphe (par exemple sélectionner différents nœuds et comparer les colorations qui s’en suivent) et un réel frein pour sa compréhension. L’affichage d’une capture d’écran statique a pu être trop éloigné de la réalité et enlever un peu de crédibilité aux chiffres retrouvés.

On a demandé au panel d’identifier les problèmes et de proposer des améliorations (en questions semi-ouvertes) concernant ces graphes de liaison de résultats Discovery Hub. Les problèmes les plus relevés sont la superposition de labels empêchant la lecture de certains d’entre eux, la taille trop petite du graphe provoquant une sensation d’étroitesse et le manque de distinction entre les liens qui se juxtaposent ou se croisent en grand nombre. L’identification de ces problèmes a été aidée par des propositions de réponse présentes dans le questionnaire, et ces propositions trouvent leur origine dans les entretiens passés au préalable. De la même manière, le panel a proposé des améliorations. Quatre d’entre elles ont été particulièrement suggérées : un élargissement du graphe, une mise en valeur des liens entre les entités, une coloration des différents nœuds en fonction de leur nature, et enfin l’affichage de moins de contenu dans le graphe.

À l'heure de donner une note d'appréciation au graphe (sur 10), le panel se montre relativement difficile, mais cohérent avec les réponses globalement apportées dans l'ensemble des questions : 83% du panel propose une note entre 3 et 6, la note moyenne étant de 4,33 sur 10.

Ce questionnaire nous a permis d'identifier les problèmes et les exigences des utilisateurs quant à la représentation graphique des liens entre les résultats de recherche sur Discovery Hub. Surtout, il nous permet également d'identifier les profils d'utilisation (pour quoi, par qui et comment) de l'application Discovery Hub, ce qui nous amène à la modélisation de l'utilisateur.

## Partie 3: Modélisation

### Modélisation des utilisateurs

- Auteurs: Baptiste Comet, Lucas Crivelli
- Relecteurs: Emilie Palagi, Mathieu Rul

Pour permettre une compréhension claire et précise d'un système, il est nécessaire de modéliser, entres autres, ses utilisateurs ainsi que ses principaux cas d'utilisation. La diffusion d'un questionnaire et la réalisation de plusieurs entretiens nous ont permis de concevoir deux *personas* et de leur associer chacun deux *scénarios d'utilisation*.

Nous avons tenté de couvrir l'ensemble des profils utilisateur du système Discovery Hub en concevant deux personas bien distincts. En effet, les données récoltées du questionnaire et des entretiens ont mis en évidence deux types d'utilisateurs différents.



Figure 12 : Persona 1 : Benoit Maitre

#### *Persona 1*

**Prénom et nom :** Benoit Maitre

**Âge :** 22 ans

**Profession :** étudiant en commerce à Paris

**Situation familiale :** célibataire

**Niveau informatique :** moyen

**Caractère :** extraverti, dynamique, impatient

**Activités domestiques et de loisirs :** Benoit sort avec ses amis, va au cinéma, regarde des séries TV américaines, partage ses centres d'intérêt et ses sorties sur les réseaux sociaux.

**Usage des technologies :** Benoit a un PC portable, un iPhone et un iPad. Il se sert du PC pour écrire des essais, prendre des notes pour ses cours et réaliser des présentations (grâce à la suite Microsoft Office). Il préfère utiliser ses terminaux mobiles pour se connecter à ses réseaux sociaux, lire des journaux, se divertir sur le web.

**Attitude à l'égard des technologies :** relativement à l'aise avec la technologie, Benoit s'intéresse aux derniers modèles sortis dans certaines gammes (smartphones, tablettes, TV).

**Centres d'intérêt :** passionné de cinéma (notamment regardant sur les musiques de film), de publicité.

**Objectifs** : élargir ses connaissances dans le domaine du cinéma, afin de découvrir de nouveaux films ou réalisateurs intéressants. Mieux connaître les personnalités du cinéma lui permettra de mieux choisir les affiches qu'il souhaitera voir en salle.

**Citation** : "New is always better"

### **Scénarios du persona 1**

1. Benoit a récemment vu au cinéma *Gatsby le Magnifique* avec Leonardo DiCaprio. Impressionné par le rythme du film, son esthétique et sa bande-son, il souhaite en savoir plus sur les personnes responsables de cette "réussite artistique" selon lui. Un ami lui a rapidement parlé du site Discovery Hub, il va donc essayer cette application pour la première fois. Il recherche "The Great Gatsby" et "Baz Luhrmann", et espère pouvoir trouver d'autres éléments intéressants en parcourant le graphe de données dont il a entendu parler.
2. Dans le cadre d'un de ses cours, Benoit doit réaliser une présentation sur un courant artistique de son choix. Connaissant bien le cinéma américain des années 70, Benoit cherchera des exemples concrets de films acteurs ou réalisateurs pour appuyer ses idées. Il utilise le site Discovery Hub (qu'il a déjà utilisé avant, pour des recherches personnelles) afin de trouver une liste d'entités du cinéma des années 70, et recréer un réseau de connaissances pour son devoir. Sa recherche se porte sur "Cinema of the United States" et "Seventies". Il espère trouver parmi les résultats proposés un film typique de ce mouvement, afin de resserrer le domaine du graphe résultat de Discovery Hub.

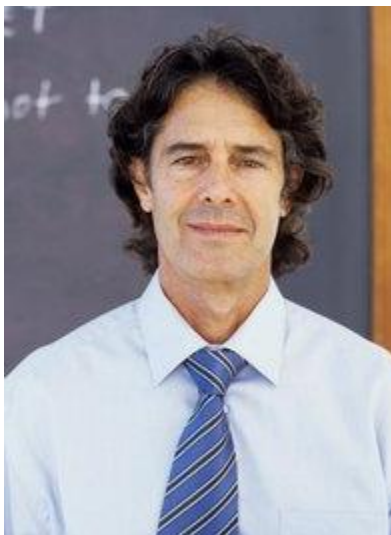


Figure 13 : Persona 2 : Jean Villard

### **Persona 2**

**Prénom et nom** : Jean Villard

**Âge** : 41 ans

**Profession** : enseignant chercheur à l'EPU

**Situation familiale** : marié, 2 enfants

**Niveau informatique** : moyen

**Caractère** : curieux, patient, attentif

**Activités domestiques et de loisirs** : lecture, cinéma, ballade nature, voyage.

**Usage des technologies** : usage à la fois personnel et professionnel pour toute sorte d'usage.

**Attitude à l'égard des technologies** : niveau expert en informatique, au courant des dernières technologies.

**Centres d'intérêt :** passionné d'histoire, de littérature, de nouvelles technologies et de géographie.

**Citation :** "Découvrir le monde qui nous entoure en quelques clics"

**Objectifs :** Jean a pour objectif d'élargir ses connaissances sur ses thèmes favoris. De plus, il est curieux d'explorer le web de données et d'en découvrir la puissance. Enfin, il aimerait pouvoir valider certaines de ses hypothèses basées sur ses connaissances.

### *Scénarios du persona 2*

1. Jean aimerait approfondir ses connaissances sur la bataille de Waterloo (acteurs, enjeux, etc.). Pour se faire, il pense qu'il serait intéressant d'utiliser Discovery Hub pour avoir accès à des informations qu'il ne pourrait pas trouver de lui-même en passant par un moteur de recherche classique. Après avoir accédé à Discovery Hub, il saisit le mot-clé "Battle of Waterloo" et lance la recherche. Son attention s'arrête sur le résultat "Napoleon" et il souhaite connaître l'ensemble des liens qui existent entre ce personnage et la bataille de Waterloo. Il accède donc au graphe pour visualiser les liens. Une fois arrivé sur le graphe, il le parcourt rapidement puis il tente de voir l'ensemble des nœuds qui relient le nœud "Napoléon" à celui de "Battle of Waterloo". Jean tente alors de passer sa souris sur le nœud "Napoleon".
2. Jean, passionné de littérature, aimerait vérifier les liens qui existent entre 2 écrivains et ainsi valider l'hypothèse qu'il a sur la question. Pour se faire, il pense tout de suite à Discovery Hub qui lui permettrait d'obtenir directement l'ensemble des liens existant entre ces 2 écrivains. Il lance donc une recherche sur les noms des 2 personnages. Par la suite, il parcourt la liste des résultats tout en cliquant sur chacun d'eux pour afficher le graphe de liens et ainsi avoir une représentation globale de tous les liens existants. Il tente alors de trouver le/les liens existants en identifiant les nœuds intermédiaires.

### *Remarques*

À la suite de cette phase de modélisation de l'utilisateur, nous avons remarqué que les données issues du questionnaire et des entretiens étaient insuffisantes.

Il est possible de se demander si les personnes interviewées, ou qui ont répondu au questionnaire, étaient trop semblables d'un point de vue du profil utilisateur. Une plus grande diffusion du questionnaire et un plus grand panel de personnes interviewées nous auraient peut-être amené à une plus grande diversité des profils utilisateurs et personas lors de leur modélisation.

De plus, il nous a été difficile de rédiger des scénarios pour les personas, car les cas d'utilisation de Discovery Hub n'ont pas pu être clairement identifiés lors des phases d'entretien. Dans des conditions réelles, nous aurions pu continuer les phases d'entretien jusqu'à obtention de données pertinentes et significatives. Par



conséquent, par manque de données, la rédaction des scénarios des personas a été fortement influencée par notre propre expérience/vision du système.

## Analyse et modélisation de la tâche

- Auteurs: Emilie Palagi
- Relecteurs: Lucas Crivelli, Mathieu Rul

### *Analyse de la tâche*

Afin d'analyser la tâche et la modéliser, nous avons utilisé la méthode HTA (*Hierarchical Task Analysis*) et plus précisément la *méthode analytique de description de tâche* (MAD).

Nous avons utilisé la technique du pourquoi/comment en nous basant sur les résultats obtenus lors des entretiens utilisateurs.

La technique du pourquoi/comment nous a permis de distinguer les différentes tâches qu'effectuent les utilisateurs (à travers les questions « pourquoi ? »), leurs sous-tâches s'il y en avait (à travers les questions « comment ? »).

Cette technique d'analyse de la tâche a la particularité de faire ressortir également les différentes relations qu'entretiennent les tâches entre elles et leur organisation.

De cette analyse de la tâche, nous avons pu formaliser notre modèle de tâches et le représenter en arbre de tâches.

### *Modélisation de la tâche*

Lorsqu'une recherche sur Discovery Hub est lancée sur le moteur de recherche, il apparaît à l'utilisateur une liste de résultats liés à sa recherche, de là commence notre modèle de tâches (il s'agit d'un prérequis de notre modèle de tâches).

Le modèle de tâches que nous proposons s'applique aux différents personas définis précédemment de manière identique.

L'objectif de notre modèle de tâche est la compréhension de la raison pour laquelle un résultat X, qui paraît inattendu au premier abord, est présenté lors d'une recherche sur Discovery Hub, et ce en utilisant la représentation graphique que le moteur de recherche propose.

Le modèle de tâches présente, en premier lieu, quatre parties différentes (qui correspondent à quatre tâches séquentielles distinctes) :

1. La sélection d'un résultat X inattendu dans la liste des résultats proposés par Discovery Hub
2. L'ouverture de la représentation graphique
3. La recherche de la relation entre la recherche et le résultat X
4. La découverte de la relation

Il est possible de définir deux manières différentes d'aborder les résultats du graphique :

- la première correspond à la recherche simple des relations entre le résultat et la recherche (tâche nommée « *Découvrir les liens* » dans le modèle de tâches),
- la seconde convient à la vérification d'une hypothèse de relation que l'utilisateur a déjà en tête et voudrait vérifier dans le graphique (tâche nommée « *Vérifier une hypothèse* » dans le modèle de tâches).

Il s'agit de deux tâches alternatives strictes.

Pour effectuer ces deux tâches différentes, l'utilisateur sera amené à effectuer les mêmes actions sur le graphique : dans notre modèle de tâches nous les avons nommées « *Expliciter les liens* » et « *Repérer un lien probable entre les nœuds* » (deux tâches alternatives non strictes).

La tâche « *Expliciter les liens* » se décompose en deux sous-tâches (alternatives non strictes) :

- *Modifier la position des nœuds roses* : en déplaçant la position des nœuds roses dans le graphique.
- *Bouger la position des nœuds gris* : les nœuds gris ont la particularité d'être des nœuds fixes. Les utilisateurs peuvent néanmoins les déplacer en les sélectionnant avec le curseur, mais les nœuds gris reviendront à leur place d'origine. Il s'agit donc d'une modification de leur position momentanée.

La tâche « *Repérer un lien probable entre les nœuds* » se décompose également en deux sous-tâches (alternatives non strictes) :

- *Bouger le curseur pour « surligner » les nœuds qui sont directement reliés entre eux* : en déplaçant le curseur sur le graphique et en l'arrêtant sur un nœud, certains nœuds directement reliés à lui seront entourés de gris pour les nœuds roses et en rose pour les nœuds gris.
- *Laisser le curseur sur un nœud et lire le texte qui apparaît* : lorsqu'on laisse le curseur sur un nœud, un texte court apparaît sur le graphique décrivant le nœud sélectionné. Ce texte peut permettre à l'utilisateur d'avoir une hypothèse de relation et de la vérifier par la suite, ou de comprendre la représentation du résultat X.

La tâche « *Chercher la relation entre la recherche et le résultat X* » est une tâche itérative : elle s'effectue en boucle tant que la relation n'est pas découverte.

« *Trouver la relation entre la recherche et le résultat X* » correspond à la clôture de notre modèle de tâches : l'utilisateur a compris, à travers la manipulation du graphique et les informations qu'il propose, la relation entre sa recherche et le résultat X inattendu proposé par Discovery Hub.

### **Remarque**

À partir des analyses des entretiens et des questionnaires, les difficultés rencontrées par les différents utilisateurs concernent la tâche « *Rechercher la relation entre la recherche et le résultat X* » de notre modèle de tâches. En effet, il en ressort une mauvaise compréhension du graphique et une méconnaissance des différentes possibilités de le manipuler.

Il s'agit donc des éléments qu'il faudra améliorer et modifier par la suite : c'est-à-dire une modification de la présentation graphique de Discovery Hub.

Il est également possible de remarquer, à travers les entretiens, que la tâche « *Repérer un lien probable entre les nœuds* » est moins utilisée qu'« *Expliciter les liens* ».

### Modèle de tâche du système

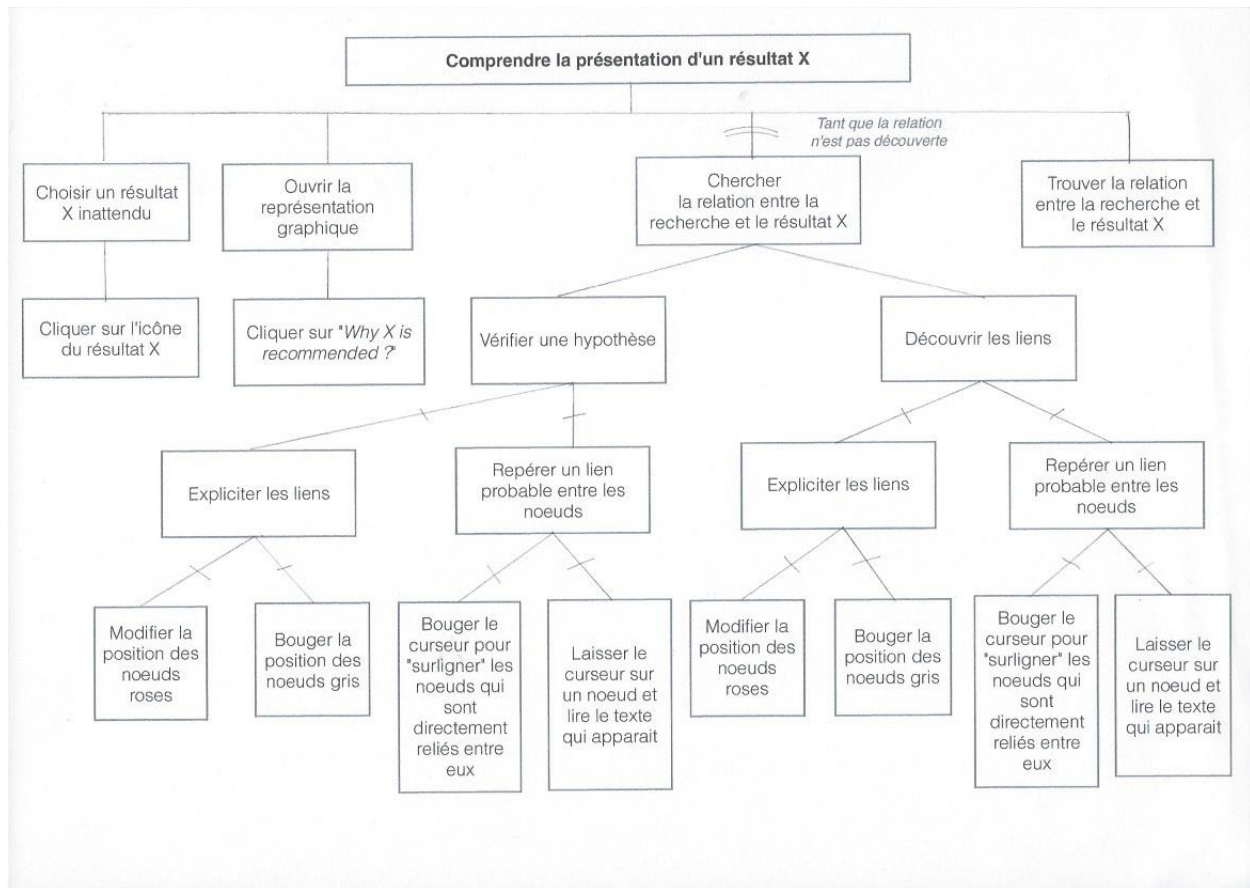


Figure 14 : Modèle de tâche du système

## Conclusion

---

Ce document décrit les 3 premières phases de notre projet: analyse de l'existant, récolte d'information à l'aide d'entretien et questionnaire, et enfin modélisation de l'utilisateur et des tâches.

Chaque phase du projet a été discutée et validée par l'ensemble des membres de l'équipe.

Concernant la rédaction du rapport, chaque partie a été rédigée par une/plusieurs personnes(s) (auteur(s)) et relue par un/plusieurs autres(s) membres(s) de l'équipe (relecteur(s)). Pour éviter de contraindre les membres de l'équipe, chacun a travaillé sur sa partie en parallèle. Enfin, chaque partie a été ajoutée à un document commun (visible et modifiable par tous) pour relecture et correction par l'équipe. La mise en page définitive du rapport a été faite par Mathieu Rul.

## Annexes

---

### Planning

- **Semaine 1 (30/09/2013)**
  - Feuille de route
    - Création du groupe et choix du projet : Discovery Hub
  - Livraisons : Liste des membres du groupe
- **Semaine 2 (07/10/2013)**
  - Feuille de route
    - *Équipe : Analyse de l'existant* (tous les membres de l'équipe ont travaillé sur la même tâche)
      - **Compréhension initiale du sujet**

Pour cette étape, nous avons regardé en détail l'interface existante. De plus, nous avons recensé l'ensemble des fonctionnalités de celle-ci.
      - **Recensement des systèmes existants : description et critique des systèmes de représentation graphique de données**

Nous avons recensé 3 outils existants potentiellement intéressants :

        - **RelFinder** : Visualisation du web de données à travers une interface web (réseau 2D)
        - **Pearltrees** : Création d'un réseau autour de centres d'intérêt
        - **LinkedIn InMaps** : Visualisation du réseau LinkedIn
      - **Analyse : éléments à conserver, modifier et ajouter**

Enfin, pour la phase d'analyse, nous avons dressé un tableau récapitulatif des avantages et inconvénients de l'ensemble des systèmes (dont celui à réaliser). À partir de ce tableau, nous avons décidé des éléments à conserver, à modifier et ceux à ajouter.
- **Semaine 3 (14/10/2013)**
  - Feuille de route
    - Collecter des informations
    - Modéliser le système: utilisateur, scénario, etc.
    - Préparer l'entretien
- **Semaine 4 (21/10/2013)**
  - Feuille de Route
    - Entretiens avec 5 utilisateurs
    - Synthèse des entretiens
- **Semaine 5 (28/10/2013)**
  - Feuille de Route
    - Analyse des entretiens

- Préparation du questionnaire en ligne
- **Semaine 6 (04/11/2013)**
  - Feuille de Route
    - Rédaction du questionnaire
  - Livraisons : Mise en ligne du questionnaire
- **Semaine 7 (18/11/2013)**
  - Feuille de Route
    - Répartition des parties de chacun sur le rapport
    - Rédaction du rapport
- **Semaine 8 (09/12/2013)**
  - Livraison du rapport intermédiaire

## Procédure entretien

1. **Présentation**
  - a. de l'équipe
  - b. du projet/cadre
  - c. de la motivation de l'entretien
2. **Présentation de l'interviewé**
3. **Recenser les thèmes de recherche favoris de l'interviewé**

### Poser les questions

⇒ Êtes-vous intéressé par la découverte de nouvelles informations liées à vos centres d'intérêt ? Si oui, pour quelles raisons (curiosité, travaux de recherche, etc.) ?

⇒ Utilisez-vous des moteurs de recherche qui permettent d'obtenir des réponses inattendues à partir de mots-clés ? Si oui, lesquels ?

4. **Lancer une recherche sur un des thèmes favoris de l'interviewé**

### Poser les questions

⇒ Vous sentez-vous « submergé » par le nombre de résultats proposés ?

⇒ Êtes-vous intéressé par les raisons des résultats proposés ?

5. **Identifier les résultats obscurs**

### Poser les questions

⇒ Voyez-vous des réponses vous semblant impertinentes/inattendues ?

Voulez-vous savoir pourquoi ces résultats vous sont proposés ?

6. **Demander à l'interviewé de choisir un résultat inattendu**
7. **Demander à l'interviewé de trouver la relation sur Google**

### Poser les questions

⇒ Avez-vous trouvé le lien ? Si oui en combien de temps ?



## 8. Demander à l'interviewé de trouver la relation sur Discovery Hub

### Lancer la recherche sur le mot-clé et les résultats

#### Poser les questions

⇒ *Trouvez-vous le/les liens(s) ? Si oui en combien de temps ?*

⇒ *Le lien est-il explicite ?*

⇒ *Que pensez-vous de la représentation graphique ?*

*Est-ce simple à comprendre et à utiliser ?*

⇒ *Avez-vous des suggestions sur la représentation graphique ? Image ? Taille ? Éléments manquants ? Disposition ? Etc.*

## 9. Remercier l'interviewé

### Questionnaires

Le questionnaire est hébergé sur Google Documents à cette adresse :

<https://docs.google.com/forms/d/14HrnYxleyj3pZIWVZgdXQ8lipc81fDW0V0CXTM8MsGI/viewform>