

## Inspirer le concepteur : WebGallery

La présentation d'exemples a une influence sur les productions réalisées par les concepteurs [Smith 1993, Marsh 1993, Bonnardel 1999, 2009]. Plusieurs travaux récents [Lee 2007, Lee 2010, Ritchie 2011] se sont intéressés aux galeries web comme sources d'inspirations. Ces travaux proposent aux concepteurs de naviguer selon un critère de similitude entre les pages référencées. Cette similitude est sensée porter sur des aspects liés au style des pages. Cependant, aucune mesure n'a été faite quant à la perception des similitudes de style qu'ont réellement les utilisateurs. Dans ce chapitre, nous présentons une étude qui vise à en savoir plus sur la perception effective des similitudes de style par l'humain et nous explorons des algorithmes visant à évaluer cette similitude.

### 12.1 Principes

La consultation d'exemples est une pratique courante en conception [Eckert 2000, Buxton 2007, Herring 2009]. Elle permet aux concepteurs de suivre les modes, tendances, avancées technologiques, etc. et permet de réutiliser des parties des exemples pour constituer ses propres artefacts. Cela est particulièrement vrai dans le domaine de la conception de sites web parce que : 1/ les évolutions des technologies du web et du style des pages web sont rapides ; en conséquence, les concepteurs doivent en permanence se tenir à jour des tendances, 2/ un site web est constitué d'un ensemble assez restreint d'éléments clefs qui sont déclinés et répétés sur tous les sites (entête, barre de navigation, paragraphe, etc.) ; aussi la réutilisation est-elle favorisée, et enfin 3/ il est assez facile d'accéder au code des sites web dans le but de le réutiliser. Ces trois aspects ont été mis à profit pour favoriser l'apprentissage et la conception de sites web par la présentation d'exemples [Lee 2010].

En pratique, l'exploration du web - qu'elle soit conduite pour trouver des exemples ou des informations - se fait via l'utilisation d'un moteur de recherche. Si les moteurs sont performants pour rechercher des pages web sur la base de leur contenu textuel, ils ne sont en revanche pas (encore) capables de rechercher des pages sur la base de leur présentation, de leur style graphique (couleurs dominantes, mise en page, etc.). Ainsi, en l'état actuel, ils sont d'une aide limitée à un concepteur qui souhaiterait explorer des variantes de styles de pages données.

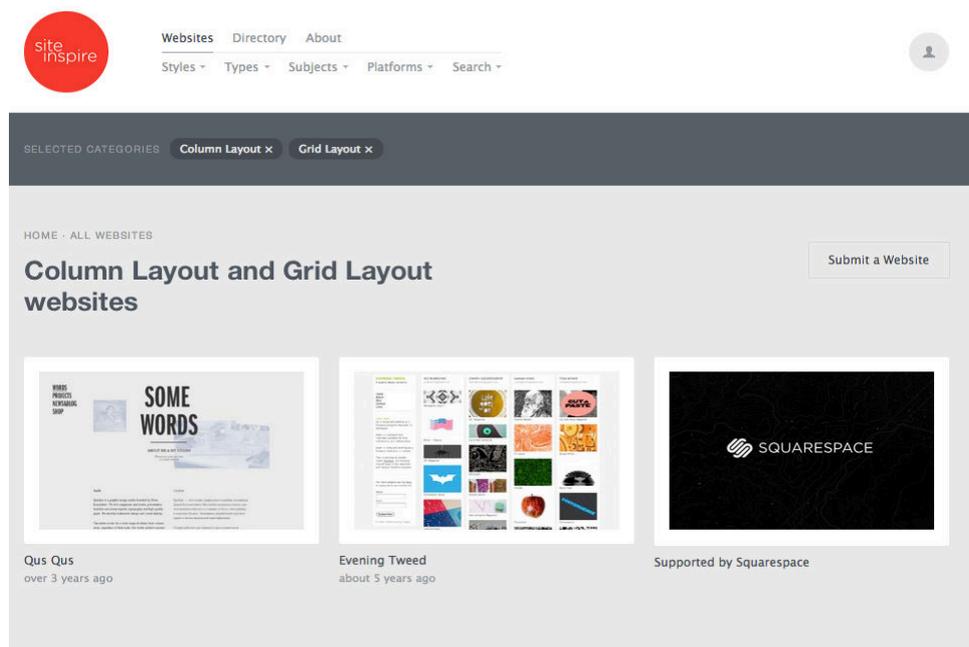
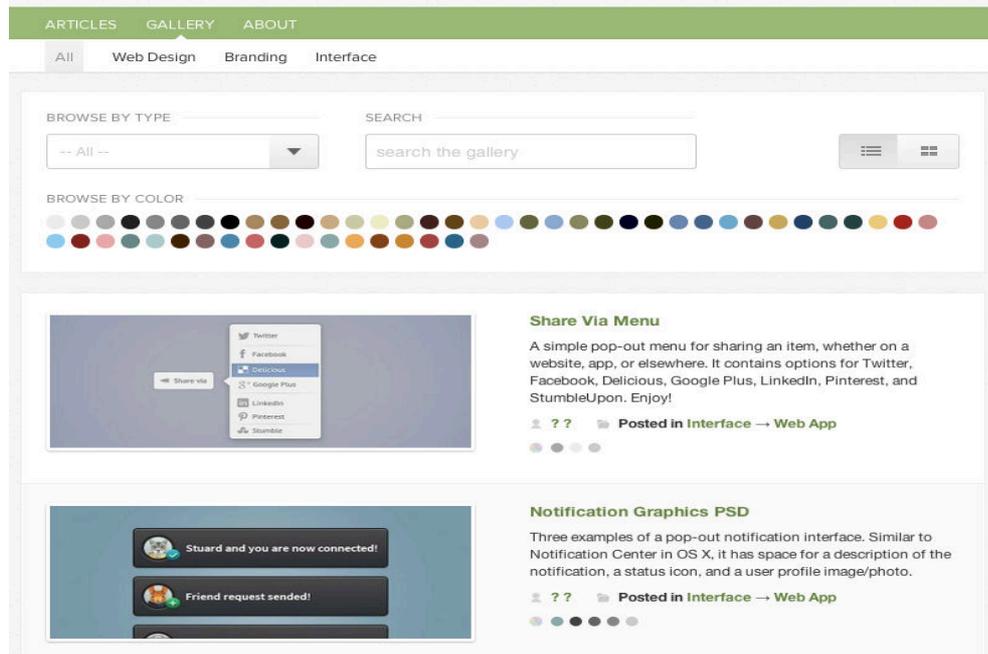
### 12.1.a Galeries web

C'est pour répondre à cette problématique que les concepteurs se sont orientés vers l'archivage de pages web, que ce soit de manière personnelle ou collective, avec la mise en place de galeries. Une galerie web est une collection sélectionnée et organisée de sites web. Tous les jours, des centaines de nouveaux sites web sont mis en place par des concepteurs à travers le monde. Certains, jugés intéressants, seront repérés, collectés et annotés par des concepteurs afin de les placer en vitrine sur une ou plusieurs galeries. Des blogs, spécialisés en web design, rassemblent les meilleures de ces galeries dans des listes destinées à inspirer les concepteurs selon des classements par thèmes, tendances ou technologies.

En général, les galeries ressemblent à leurs équivalents physiques : ce sont de simples catalogues que le concepteur peut feuilleter pour y chercher l'inspiration. Très souvent, les possibilités d'interaction s'arrêtent là. Parfois, des fonctionnalités de recherche avancée sont mises en place, par exemple une recherche par catégories, couleurs et même parfois types d'éléments présents (logo, navigation) ou mise en pages (logo, images, etc.). Certaines galeries, comme [siteinspire.com](http://siteinspire.com), proposent même de naviguer sur la base de la similitude du design. Cependant, rien n'est dit sur la manière donc cette navigation est réalisée.

En tout état de cause, les galeries jouent un rôle important en amont et en aval de la conception de sites. Je résume ici leur influence dans le cycle de vie que j'organise en trois phases :

1. **Inspiration** : Le designer cherche de l'inspiration pour un certain type de site (par exemple banque, e-commerce, site personnel, etc.), un style donné (minimaliste, flat-design, etc.), un type de mise en page, un élément particulier (menu, bouton, titre, etc.), un aspect esthétique (couleur, texture, etc.), soit sur des blogs de conception, des galeries web ou via ses contacts sur les réseaux sociaux ;
2. **Production** : Le designer produit un site web et le met en ligne ;
3. **Diffusion** : Le designer en parle sur les réseaux sociaux (Twitter, Pinterest, Google+, Facebook, etc.). Il le met sur son portofolio, le soumet à des galeries web ou utilise alternativement un service payant pour le diffuser dans de nombreuses galeries (e.g. [thecssgallerylist.com/](http://thecssgallerylist.com/) gère une liste de 100 sites, [galleryrush.com/](http://galleryrush.com/) en recense plus de 120) pour une somme comprise entre 25 et 50 dollars US.



La diffusion est très importante pour le designer : c'est par elle qu'il gagnera ou fidélisera des clients. Il se fait connaître des autres designers, des sites spécialisés et par conséquent gagne des rétroliens (des sites webs connus ont des liens vers sa page/portfolio) ce qui améliore son référencement (page-rank) sur les moteurs de recherche (Google search, Bing, etc. ). Les gardiens du domaine [Csikszentmihalyi 1997] entrent alors en jeu. Des twitteurs influents retweetent les meilleurs designs des gens qu'ils suivent (resp. partagent sur Google+ et Facebook, épinglent sur Pinterest, 3-98

etc). Les blogueurs spécialisés dans la conception web produisent régulièrement des listes des meilleurs designs dans une catégorie ou pour un style donné. Ainsi les galeries sont aussi un outil social d'acceptation par la communauté. En cela, elles constituent à la fois une source d'inspiration mais aussi un filtre puisque seuls les sites cooptés y apparaissent, reflétant ainsi la pression de la communauté.

En amont de la production, les galeries sont une source d'inspiration à l'intérieur de laquelle les concepteurs vont puiser de manière plus ou moins structurée, recherchant des pages ou des éléments d'un certain style. Nous nous intéressons ici justement à cette notion de style qu'il nous semble cruciale de bien définir afin de pouvoir proposer, au sein des galeries, des mécanismes de navigation adéquats.

### 12.1.b La notion de style

La notion de style est définie par le TLF comme l'«Ensemble des traits caractéristiques des œuvres (d'un artiste, d'une époque, d'une civilisation)». La notion de style est contruite par les gardiens du domaine. En pratique ce sont donc les concepteurs qui apprécient et définissent les tendances de l'évolution du style des pages web. Ainsi, la plupart des blogs de conceptions web publient des articles sur les tendances actuelles dans le domaine. Par exemple pour l'année 2010, les styles dominants des sites webs peuvent être groupés en fonction des éléments graphiques utilisés (illustrations et cartoons, grandes images et arrière-plans photographiques ou photoréalistes) ou sur les textures et palettes de couleurs utilisées (motifs, bichromatique ou imprimé papier) (cf. [smashingmagazine.com](http://smashingmagazine.com)<sup>11</sup> and [webdesignledger.com](http://webdesignledger.com)<sup>12</sup>).

En 2012, les styles dominants ont évolué en mise en page et en interaction sous l'impulsion certaine des avancées technologiques en HTML5, CSS3 et responsive design (introduction d'une partie des idées de la plasticité au design web). La navigation en position fixe (la barre de navigation est fixe par rapport à l'écran et ne change pas de place lorsqu'on défile la page) (cf. [smashingmagazine.com](http://smashingmagazine.com)<sup>13</sup> et [webdesignledger.com](http://webdesignledger.com)<sup>14</sup>) ou encore les menus sur plusieurs colonnes sont des exemples d'évolution majeure. De nouvelles formes et des effets modernes sont facilement réalisables comme les cercles, les rubans fourchus, les bordures en zigzag ou en pointillé ainsi que l'animation des éléments grâce au triptyque

---

<sup>11</sup>[www.smashingmagazine.com/2010/05/04/web-design-trends-2010/](http://www.smashingmagazine.com/2010/05/04/web-design-trends-2010/)

<sup>12</sup>[webdesignledger.com/tips/web-design-trends-in-2012](http://webdesignledger.com/tips/web-design-trends-in-2012)

<sup>13</sup>[www.smashingmagazine.com/2010/05/04/web-design-trends-2010/](http://www.smashingmagazine.com/2010/05/04/web-design-trends-2010/)

<sup>14</sup>[webdesignledger.com/tips/web-design-trends-in-2012](http://webdesignledger.com/tips/web-design-trends-in-2012)

javascript/CSS3/HTML5. D'autres éléments de style sont des marqueurs de cette période, par exemple la typographie avec « letterpress », les polices de caractères personnalisées ou encore les illustrations typées XIX<sup>ième</sup> siècle et les éléments skeuomorphiques (c'est-à-dire un élément de design dont la forme n'est pas directement liée à la fonction, mais qui reproduit de manière ornementale un élément qui était nécessaire dans l'objet d'origine).

Comme on peut le constater, la notion de style est complexe et difficile à appréhender. Cela rend incertaine l'utilisation de la similitude entre styles comme moyen de navigation dans les galeries web. Malgré cela, des techniques ont été développées pour explorer les galeries d'exemples. Faceted Metadata Browsing [Yee 2003] est une technique efficace pour les galeries de grande taille. Chaque élément de la galerie est étiqueté de plusieurs facettes comme sa couleur, le type de media, la date, etc. L'exploration se fait en deux phases : à l'ouverture, l'ensemble des facettes est présenté. Durant l'exploration, l'interface est découpée en deux parties : d'une part, l'affichage des résultats correspondant aux facettes déjà choisies et, d'autre part, le reste des facettes qu'il est possible de choisir pour préciser la recherche. A chacune de ces facettes restantes est associé le nombre d'éléments qui correspondent à cette facette. Cependant, cette technique nécessite l'extraction de caractères qui peuvent être facilement compris et étiquetés manuellement par des concepteurs.

Une alternative à l'exploration par facettes est l'exploration adaptative. Dans ce cas, l'utilisateur de la galerie se voit présenter un ensemble de designs dont le style est proche de ceux qu'il est en train d'observer. Ce mécanisme est analogue au mécanisme de recommandation de films que l'on trouve sur les sites de vidéo à la demande. L'utilisateur indique un certain nombre de films qui lui plaisent et le site lui propose d'autres films qu'il pourrait voir. La recommandation peut s'appuyer sur de nombreux critères comme le genre, le réalisateur ou les acteurs. Contrairement à la recherche par facettes, le mécanisme calcule une mesure de similarité en boîte noire et présente les meilleurs résultats. [Lee 2010] a montré que pour la conception de sites web, l'exploration de galeries d'exemples par similarité est plus efficace qu'une exploration aléatoire.

Ces galeries adaptatives peuvent être divisées en deux catégories : d'une part, les galeries qui génèrent de nouveaux exemples à partir des exemples sources, et d'autre part, les galeries qui calculent une mesure de similarité à partir d'une base d'exemples existants.

Design Gallery [Marks 1997] est un visualiseur de scènes 3D qui génère et organise automatiquement des variantes d'une même scène 3D en modifiant les lumières présentes dans la scène. Design Gallery propose une organisation hiérarchique des scènes basée sur le rendu. Magellan [Masson 2010], présenté dans le chapitre 13, génère automatiquement des variantes d'interfaces utilisateurs correspondant à un modèle des tâches. Ces exemples sont créés en recombinaison et en faisant évoluer des éléments d'interfaces (les boutons, listes et autres interacteurs) ainsi que leurs paramètres (couleurs, taille, positionnement).

Les galeries telles que Adaptive Interface [Lee 2007] ou D.tour [Ritchie 2011] déduisent la similarité entre exemples d'après un jeu de caractéristiques extraites des sites web manuellement ou automatiquement. Avec Adaptive Interface, les concepteurs explorent un ensemble d'exemples suivant deux variantes : soit la galerie est composée d'éléments qui sont vus comme similaires à un exemple cible, soit la galerie est composée d'éléments qui maximisent la variété selon un aspect du style. Le style est déterminé par des méta-données assignées manuellement à chaque site web, et reste limité à la couleur du fond, la police de caractère du document, le nombre de colonnes et la densité visuelle. Avec D.tour, le style est automatiquement calculé à partir de la structure de la page HTML. Il est défini par quatre aspects majeurs : la complexité de la page, calculée à partir du nombre d'éléments dans la page et leur positionnement dans l'arbre HTML, la densité visuelle basée sur la quantité d'images et de texte dans le document, la couleur exprimée en température, luminosité et saturation, et finalement la typographie limitée à la taille de la police. Bien que plus sophistiquée que dans Adaptive Interface, cette définition du style manque encore des aspects critiques du style comme la mise en page, les éléments dynamiques ou esthétiques. Finalement si plusieurs éléments sont sélectionnés, leurs caractéristiques sont agrégées en faisant la moyenne. Ainsi, sélectionner un site blanc et un site noir résultera en l'affichage de sites gris, plutôt que d'un ensemble de sites noirs et blancs.

Ce tour d'horizon de la notion de style nous montre combien cette notion est multidimensionnelle, encore mal définie et abordée d'un point de vue surtout système dans les galeries actuelles puisque focalisées sur le facilement calculable plutôt que ressenti réel des utilisateurs. Nous proposons ici d'attaquer le problème à la racine en étudiant ce que signifie la similitude de style du point de vue de l'humain. Nous pensons que cette étape est nécessaire pour établir une vérité terrain et ainsi pouvoir comparer plus objectivement les différentes approches et futurs algorithmes.

### 12.1.c La perception humaine de similitude de styles

Le principe général [Masson 2014] est de collecter un ensemble de pages web illustrant la diversité des designs puis de les présenter par groupes de trois (triplets) à des utilisateurs afin qu'ils identifient les deux plus proches.

La notion de style étant vague, nous avons précisé aux participants qu'ils devaient se concentrer sur la mise en page et non les couleurs par exemple. Nous nous attendions à ce que tous les triplets ne fassent pas **consensus** parmi les participants, voire à ce que, confrontés plusieurs fois au même triplet, certains utilisateurs ne soient pas **cohérents** et changent d'avis d'une fois sur l'autre. Le protocole que nous avons mis en place reflète ces considérations.

Pour l'expérimentation, le corpus était composé de 211 captures d'écran des pages d'accueil de sites référencés par des galeries web existantes. Les captures d'écran ont été collectées automatiquement par un programme informatique. La liste complète des sites web et les captures d'écran peuvent être trouvées respectivement en annexes de cette thèse et à l'adresse [iihm.imag.fr/demeure/Dimitri/](http://iihm.imag.fr/demeure/Dimitri/). 164 participants (89 hommes, 46 femmes et 29 non précisés) ont rempli un questionnaire en ligne.

Les participants devaient remplir leur nom, âge, sexe, profession et expertise en conception de sites web (sur une échelle de Likert à 5 choix). Les participants étaient ensuite informés qu'ils devaient dans cette expérimentation évaluer la similarité de sites web selon leur mise en page. L'explication de la notion de mise en page suivante leur était fournie : « la structure graphique du site et l'organisation spatiale des zones ». Il leur était aussi précisé que « ni la couleur ni le contenu (le sujet du site) ne doivent être considérés ».

Avant que les participants ne commencent l'expérimentation, l'interface utilisateur de l'expérimentateur leur était présentée.

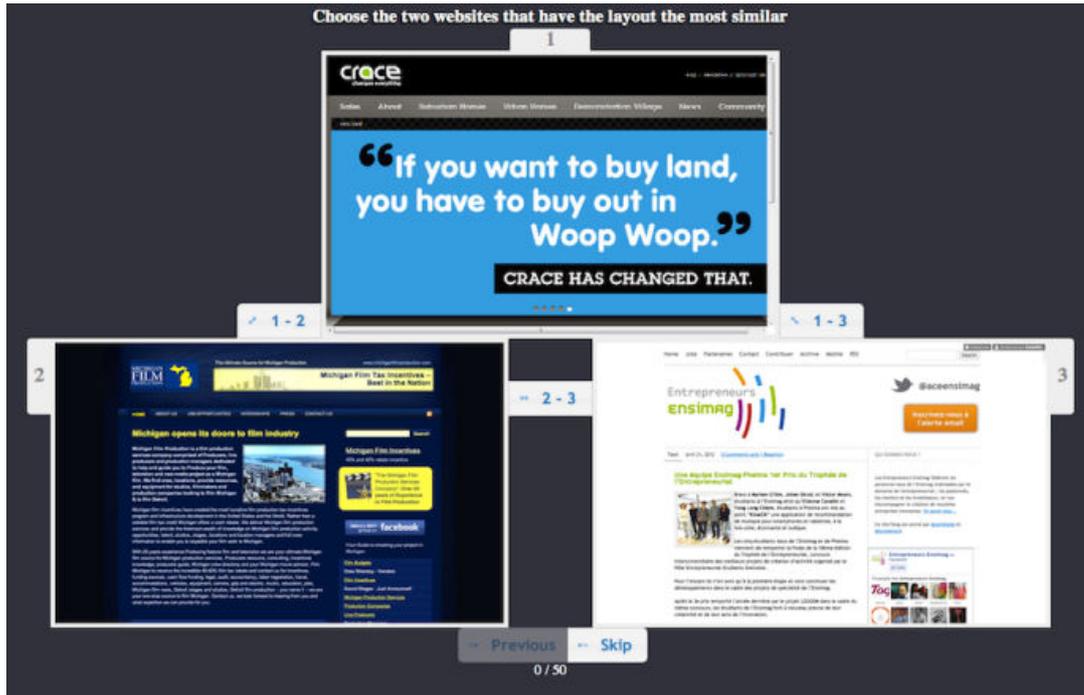


Figure 33. Interface de l'évaluateur.

Les sites web étaient présentés par groupes de trois (appelés triplets par la suite). Des boutons permettaient aux participants de choisir le couple de sites le plus similaire parmi le triplet, mais aussi de passer au triplet suivant en cas d'indécision. Il était possible de revenir aux triplets précédents en cas de changement d'avis ou d'erreur.

Durant l'expérimentation, les participants étaient exposés à 50 triplets. Les 5 premiers triplets, identiques pour tous les participants, ont été sélectionnés pour entraîner les participants et évaluer le consensus entre participants. Les triplets suivants étaient choisis de sorte que chaque triplet soit évalué par plusieurs participants pour évaluer le **consensus**. L'ordre de présentation des sites web dans chaque triplet était équilibré au travers des répétitions. De plus, 5 des triplets étaient présentés deux fois à un intervalle d'au moins 10 triplets dans l'expérimentation pour évaluer la **cohérence** des réponses des participants. Tous les dix triplets, les participants devaient justifier en quelques mots leur choix.

L'analyse des 693 triplets jugés par au moins 7 participants donne seulement 216 (31%) triplets présentant un consensus d'au moins 5/7. Parmi eux, seuls 5 sont des consensus sur le fait qu'il n'y avait aucune similarité parmi les 3 sites web. Cela signifie que les participants étaient capables d'évaluer la similarité de mise en page et qu'il y avait une part significative d'évaluation consensuelle. Cependant, même dans le cas où les mises en page étaient dissemblables, les participants ne sont pas passés

au triplet suivant. Une explication possible est que les participants se sont concentrés sur des aspects locaux des sites (par exemple, 2 sites sont plus similaires si l'on regarde leur barre de menu, mais si l'on s'attache à la zone centrale c'est un autre couple qui est le plus similaire).

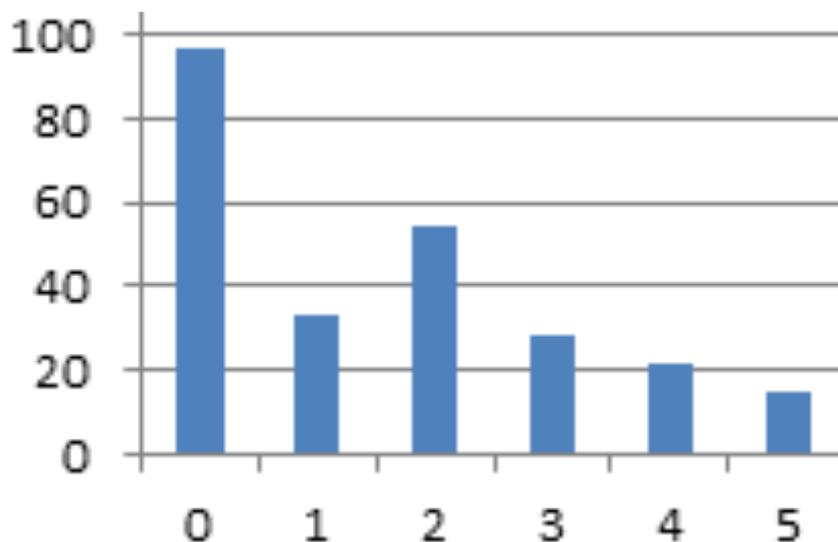
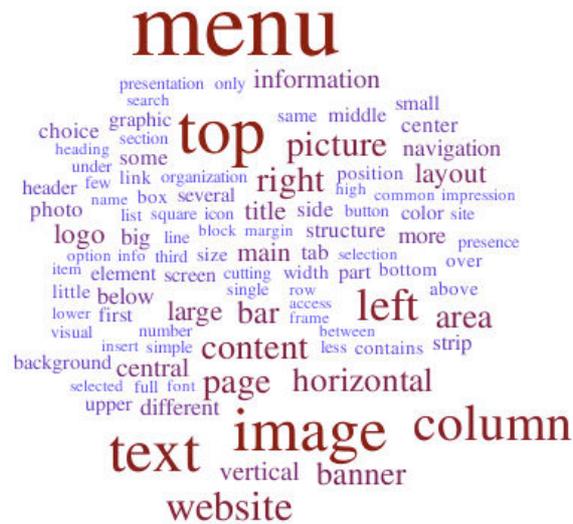


Figure 34. Nombre de participants par nombre d'incohérences sur les 5 triplets répétés.

En ce qui concerne les 5 triplets évalués par tous les participants, 3 ont un consensus supérieur à 75% (82%, 76% et 75%), 1 reçoit un consensus de 64% et le dernier aucun consensus. Puisque nous avons soigneusement choisi ces 5 triplets, nous étions étonnés que le consensus ne soit pas plus élevé. Cela suggère qu'évaluer la similarité de mise en page est loin d'être trivial.

On peut observer des incohérences dans les réponses des participants lorsqu'ils évaluaient deux fois le même triplet. Ce n'est pas un cas rare, puisque 152 des 164 utilisateurs ont fait au moins une réponse incohérente (93%). Sur un ensemble de 1102 triplets, 630 ont été testés pour l'incohérence (62%) et 341 ont levé une incohérence chez au moins un participant (54%). Sur les triplets ayant au moins 7 évaluations, 253 des 465 testés ont levé une incohérence. Cela semble indiquer que les participants n'ont pas changé leur évaluation au cours de l'expérimentation mais qu'ils étaient plutôt peu sûrs de leurs choix. En effet, parmi les 143 triplets avec un consensus d'au moins 5/7 et testé pour l'incohérence, seulement 53 ont levé une incohérence (37% < 54%).



**Figure 36 Nuage de mots des justifications**

Enfin la Figure 36 illustre sous la forme d'un nuage de mots la justification des choix des participants. Ce nuage met en avant l'importance de la position des éléments (top, left, right) mais aussi l'importance du type d'éléments tels que le menu, la navigation, les images ou le texte..

Il aurait été souhaitable de tester les algorithmes de l'état de l'art [Lee 2007, Lee 2010, Ritchie 2011] sur notre base mais du fait de la difficulté à obtenir ces algorithmes, nous n'avons pas pu le temps imparti : ceci constitue une perspective à la thèse. Nous n'avons pu tester que des algorithmes développés en interne, le but était de voir si des mesures relativement simples pouvaient rendre compte des choix des utilisateurs.

## 12.2 WebGallery en action

Nous avons implémenté quatre algorithmes relativement simples ayant pour but de donner une mesure de similitude entre deux sites par rapport à leur mise en page. Pour chacun des sites considérés, deux images sont produites : la première est une capture d'écran convertie en niveaux de gris (Figure 37), la seconde est basée sur une transformation du site de sorte à remplacer les boîtes englobantes des éléments DOM de la page par des rectangles blancs translucides (Figure 38). La génération de ces images a été faite automatiquement à l'aide d'une extension au navigateur Chrome que j'ai programmée.

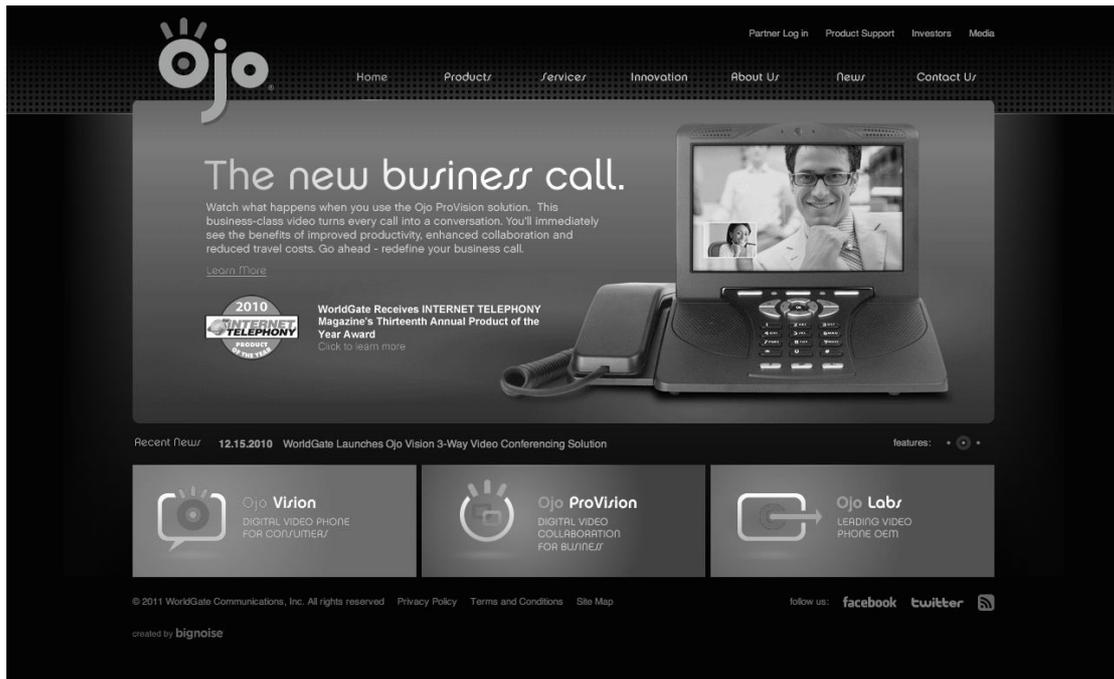


Figure 37. Capture d'écran d'un site web



Figure 38. Capture d'écran du même site avec le filtre DOM-Depth

Le fonctionnement de l'extension est le suivant : tant qu'il reste des sites web à capturer, l'extension ouvre une fenêtre contenant, dans des onglets, un sous-ensemble de pages web issues de la liste. Cette parallélisation est utile pour réduire le temps de capture d'une grande quantité de sites. Le chargement des ressources nécessaires (scripts, feuilles de styles, images etc.), la construction de l'arbre DOM (Document Object Model) ainsi que l'exécution des scripts sont chronophages et ne nécessitent pas d'écran pour être exécutés. Le nombre d'onglets ouverts par l'extension est limité à 15 pour des raisons de ressources physiques (temps de calcul, mémoire<sup>15</sup>, capacité de téléchargement).

Dès qu'une page est complètement chargée, l'extension change le focus sur son onglet. Le navigateur web calcule alors le rendu graphique de l'arbre DOM. Dès que le rendu est terminé, l'extension demande au navigateur de lui fournir une copie du rendu. L'extension encode l'image dans un format texte et l'envoie à un serveur web qui se charge d'ajouter l'image à la base de sites web. L'extension exécute chacun des scripts de sa liste et recommence le processus de capture et d'envoi au serveur pour chacun d'entre eux. Une fois tous les scripts exécutés, l'extension ferme l'onglet et passe à la page chargée suivante ou attend qu'une page finisse de charger.

Une fois que le dernier onglet est fermé, la fenêtre se ferme. S'il reste des sites à collecter, l'extension recommence le processus. Cette fermeture systématique des onglets et des fenêtres réduit l'empreinte mémoire de l'extension. L'API des extensions étant encore expérimentale et sujette aux fuites de mémoire, cela augmente la fiabilité de la capture.

Le serveur est un serveur apache qui exécute un simple script cgi en TCL qui décode le format base64 et enregistre l'image sur le disque dur. Le script ajoute aussi dans la base de données le site traité, le script utilisé et l'image résultante.

Une fois toutes les images obtenues, la similarité de mise en page est calculée de deux manières. La première calcule la distance euclidienne entre deux images (c'est-à-dire la différence pixel à pixel) ; la seconde applique un algorithme basé sur DTW [Sakoe 1978] qui aligne et déforme un signal avant de mesurer la distance étendue à deux dimensions (l'algorithme est fourni en annexe). Ainsi, en combinant les deux techniques de génération d'images avec les deux algorithmes de comparaison, nous obtenons quatre mesures de similarités :

---

<sup>15</sup> une page fait entre 10Mo et 100Mo

1. Captures d'écran en niveau de gris et mesure de la distance euclidienne ;
2. Captures d'écran en niveau de gris et mesure DTW 2D ;
3. Image DOM-depth et mesure de la distance euclidienne ;
4. Image DOM-depth et mesure DTW 2D.

Afin de se faire une idée de la pertinence a priori de ces mesures, nous les avons testées sur des exemples construits à la main. Les Figure 39 et Figure 40 résument les résultats obtenus. Ils renforcent l'idée d'une possible pertinence de l'algorithme 4 (Image DOM-depth et mesure DTW 2D) pour des cas simples.

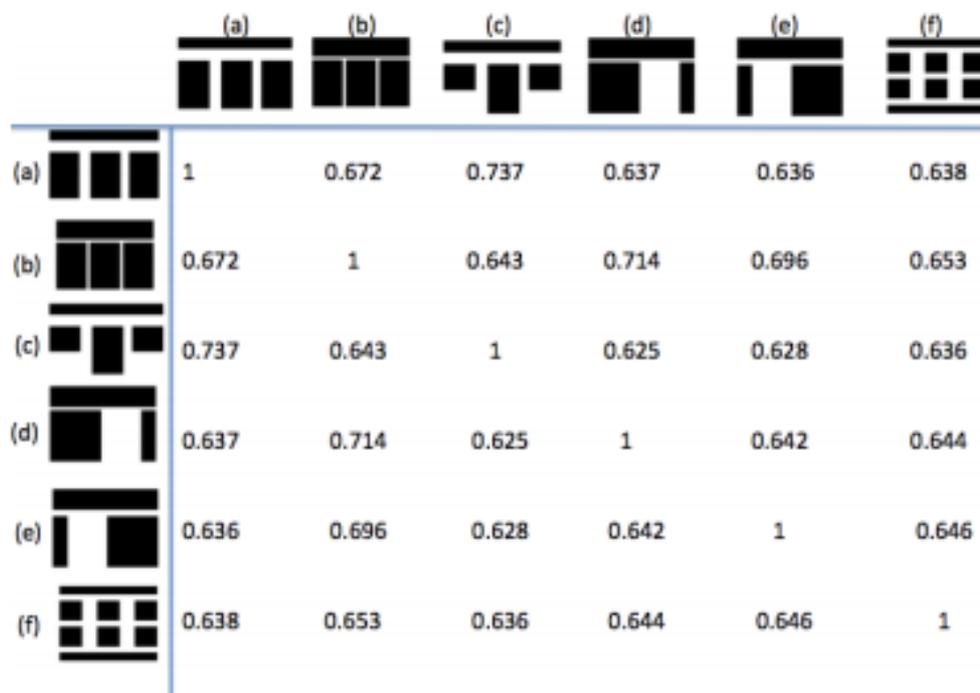


Figure 39. Similarité euclidienne

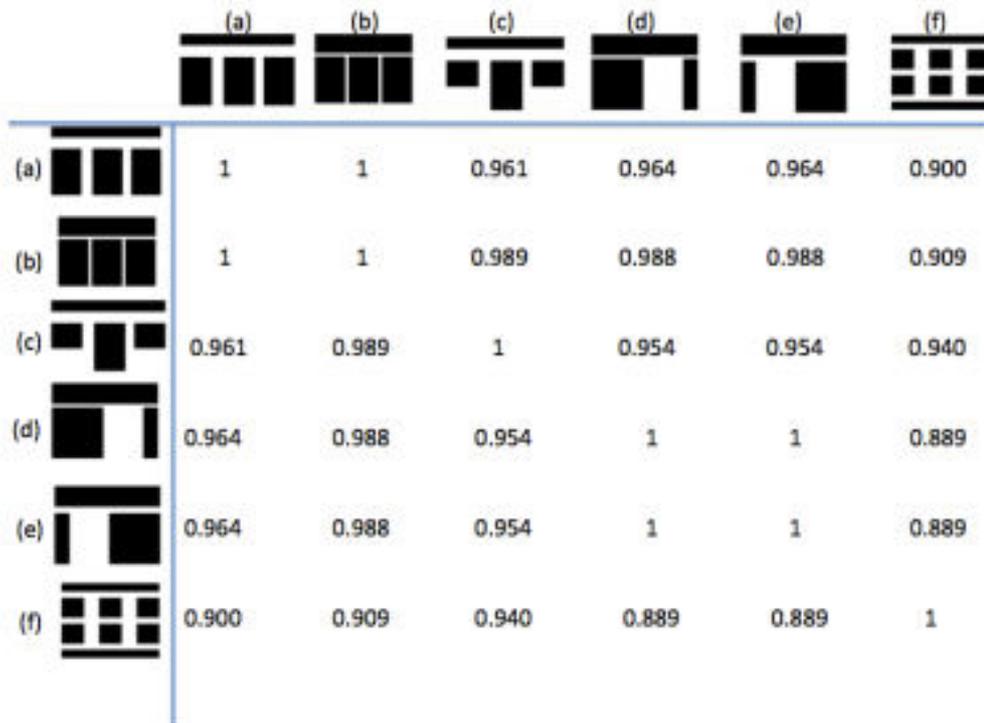


Figure 40. Similarité DTW 2D

### 12.3 Discussion

Nous avons ensuite fait évaluer les triplets de l'expérimentation précédente par ces quatre mesures de similitude. Les résultats obtenus ne diffèrent pas significativement du hasard (même en ne prenant en compte que les triplets pour lesquels les humains avaient établi un consensus). Ces résultats confirment que la mesure de similarité de mise en page entre sites web est complexe. Il serait intéressant de tester les algorithmes de l'état de l'art sur notre base et ainsi savoir si le problème vient de la base et/ou des algorithmes. Dans les deux cas, la base de test [Masson 2014] doit être étendue et les approches multipliées pour espérer approcher les résultats humains.

Une fois des algorithmes efficaces trouvés, on pourra automatiquement créer des galeries de sites web, en extraire des tendances et alors analyser leur évolution (apparition, suppression, diffusion d'éléments clefs, icônes, etc.). Ces galeries serviront de source d'inspiration à tout moment de la conception : sketch, prototype, site web existant, etc. On pourrait même alors automatiquement valider des productions par rapport à des objectifs cibles (voir chapitre 13), avec une application immédiate en enseignement.

## **Chapitre 13.**

# **Explorer l'espace de conception : Magellan**

Nous avons vu dans l'état de l'art qu'il existe des galeries web ayant pour but d'inspirer les concepteurs d'IHM en leur présentant des exemples issus du travail d'autres concepteurs. Cela permet d'explorer une partie de l'espace de conception en suivant les chemins tracés par d'autres, que ce soit dans le même domaine (par exemple la conception d'une page personnelle) ou dans d'autres domaines. Le point essentiel ici est que l'exploration de l'espace de conception se limite alors à une partie de ce qui a déjà été produit par d'autres concepteurs humains. Cela présente l'avantage de ne présenter que des produits éprouvés mais présente le désavantage de cacher des pans entiers de l'espace de conception non encore explorés par la communauté.

C'est de ce constat qu'est née l'idée de Magellan. Le principe est d'explorer l'espace de conception d'une façon différente de celle des galeries web. L'exploration se fait ici parmi des produits générés par le système sur la base d'une description fonctionnelle de la page et de règles de transformations plutôt que sur les produits déjà existants. En ce sens, Magellan propose une approche complémentaire à celle des galeries existantes.

### **13.1 Principes**

Magellan [Masson 2010] combine les AGI et les approches basées modèles. Les individus manipulés sont des transformations produisant des IHM à partir d'un modèle de concepts et de tâches. Ces IHM sont de type WIMP, codées en HTML, s'appuyant en grande partie sur la bibliothèque jQuery-ui. Le modèle des concepts et des tâches représente le « problème à résoudre » de l'AGI. Initialement, ce modèle ne présage pas de la façon selon laquelle les éléments qui le composent seront représentés. Ceci permet d'éviter la fixation du designer sur une première IHM.

L'évolution des transformations s'appuie sur une base de connaissances structurée en un réseau sémantique [Demeure 2006]. Ce réseau permet d'associer à chaque tâche ou opérateur entre tâches les IHM possibles ainsi que les paramètres de personnalisation de ces IHM. Par exemple, l'opérateur entre tâches « entrelacement » peut être représenté par des interacteurs de type « accordéon », « onglet » ou encore par une IHM organisant les éléments entrelacés de façon linéaire. En particulier, la base contient : trois types de conteneurs (simple div HTML, frame, frame rétractable), quatre types d'entrelacement (linéaire, par accordéon, par tabulation, matriciel), trois

types de choix dans une liste (combo-box, radio-boutons, menu déroulant), etc. Chacune de ces IHM peut, en outre, être paramétrée. Par exemple, l'entrelacement linéaire peut être vertical ou horizontal, la couleur des onglets peut être configurée, etc. A l'aide de cette base de connaissances, Magellan peut substituer des interacteurs

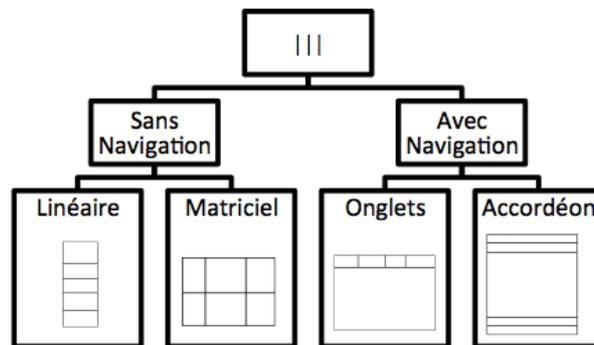


Figure 41. Illustration de la base de connaissances.

au cours du processus d'évolution (par exemple, un entrelacement par onglet devient linéaire) et varier les paramètres qui leur sont associés (par exemple, la couleur de fond des onglets).

Les transformations générées par Magellan peuvent s'appuyer sur les concepts associés aux tâches pour associer des éléments d'IHM dans un but esthétique et/ou de guidage. Le processus d'enrichissement peut regrouper des tâches qui partagent des concepts similaires.

Magellan possède une architecture en workflow, modulaire et reconfigurable, deux propriétés essentielles pour une production intensive d'exemples. Cela permet aux concepteurs d'imaginer et de construire leur propre AGI. Il existe cinq types de composants : d'une part, les quatre opérations classiques en programmation génétique, à savoir initialisation de la population, opérations génétiques, évaluation des individus et sélection des individus ; d'autre part, le contrôle de flux par branchements dans l'organigramme (branchement conditionnel, union de branches, etc.).

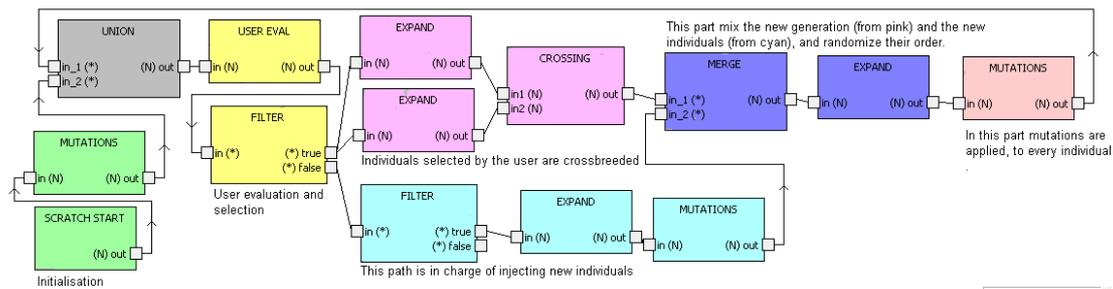


Figure 42. Architecture de l'algorithme génétique utilisé dans l'évaluation

Grâce à cette architecture, le concepteur peut, par exemple, créer le flux suivant : à l'initialisation, la boîte « scratch start » (en vert) produit une population de taille N. Cette population est transmise à la boîte « Mutations » qui applique des mutations à chaque individu, les rendant ainsi uniques. La boîte « union » insère cette nouvelle population dans la boucle. Les individus sont transmis à la boîte « User Eval » qui applique les transformations de chaque individu sur le modèle des tâches cible et affiche les CUI produites dans une interface d'évaluation. Une fois l'évaluation terminée par le concepteur humain qui utilise Magellan, un filtre est appliqué : les individus ayant reçu une évaluation positive vont suivre le chemin violet ; les autres le chemin bleu. Le chemin violet duplique la population (branchement), duplique aléatoirement des individus pour que chaque branche ait de nouveau N individus, puis croise aléatoirement les individus de chaque branche de sorte qu'en sortie du chemin violet, on ait de nouveau une population de N éléments issus de deux individus sélectionnés. Le chemin bleu crée une population de N nouveaux individus pour apporter des éléments nouveaux au concepteur à chaque itération. En effet, tous les individus sont supprimés par la succession du filtre False et True. « Expand » recrée une population de N nouveaux individus qui sont immédiatement mutés. Les deux flux sont ensuite combinés avec un ratio que le concepteur peut faire varier pour se concentrer sur une exploration en largeur (en favorisant le chemin bleu) ou en profondeur (en favorisant le chemin violet). La boîte « expand » suivante mélange les individus, puis des mutations sont appliquées à chaque individu. Le cycle recommence alors.

### 13.2 Magellan en action

Nous utilisons comme cas d'étude un logiciel de messagerie instantanée. Les raisons de ce choix sont les suivantes : les logiciels de messagerie sont assez répandus et similaires. Il sera donc plus facile d'observer l'influence des exemples produits par Magellan. Ils sont également plus riches que le simple contrôleur de Quiroz [Quiroz 2007]. Enfin, ils font intervenir plus d'interacteurs que les pages web produites par

Monmarché [Monmarché 1999]. En plus de l'échange de messages textuels, le logiciel considéré permet de gérer une liste de contacts ainsi que le statut de l'utilisateur (occupé, disponible, etc.). L'utilisateur peut gérer les conversations avec ses contacts. Pour les besoins de l'expérience, le logiciel est démarré avec une liste de cinq contacts. L'utilisateur est Gaëlle. Son profil contient sa photo, son nom et son statut. Deux conversations sont ouvertes : l'une avec Bob ; l'autre avec Dimitri. De plus amples informations sur le cas d'étude peuvent être trouvées dans [Masson 2010].

La Figure 43 montre comment les messages écrits par Gaëlle peuvent être précédés d'éléments de son profil comme sa photo par exemple. Ceci est rendu possible par le fait que, dans le modèle de concepts et de tâches, les messages et éléments du profil partagent un même concept : Gaëlle. L'action de suppression d'un contact partage le concept de "contact" avec chacun des contacts de la liste. Aussi, une mutation possible est l'insertion d'un bouton « supprimer » spécifique à chaque contact plutôt qu'un bouton général. Ces aspects sont décrits plus précisément dans [Masson 2010].

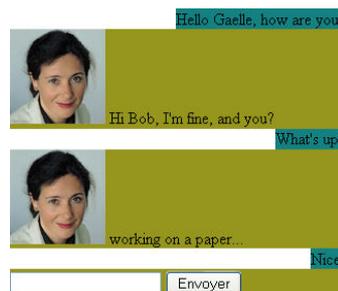


Figure 43 : IHM enrichie. Chaque message de Gaëlle est précédé de sa photo, élément du profil.

### 13.3 Discussion

Nous avons procédé à une évaluation de Magellan pour savoir si la navigation dans l'espace de conception ouvert par Magellan peut inspirer les designers d'IHM dans une première phase de conception. Nous voulions aussi cerner comment les AGI peuvent être plus généralement utiles lors de cette phase. Le but de l'expérience n'était pas de mesurer un gain de créativité par l'usage de Magellan (il n'existe pas d'outil pour une telle mesure) mais plutôt le ressenti des designers vis-à-vis de cet outil pendant leur activité.

L'IGA de Magellan était configuré de la façon suivante : la population initiale de seize designs était générée aléatoirement selon des paramètres permettant aux premiers exemples d'être assez complexes et variés. Le designer sélectionnait ensuite les individus qui lui semblaient intéressants. Ces individus pouvaient ensuite

participer à la reproduction. Un nouveau groupe de huit individus était alors généré par croisement des individus sélectionnés. Huit nouveaux individus, étrangers au processus de sélection, étaient ensuite ajoutés à la population. Les raisons de ce processus sont les suivantes :

- Une population de seize éléments est suffisante pour avoir de la variété sans trop surcharger cognitivement le designer ;
- Les individus sélectionnés ne sont pas conservés d'une génération à l'autre (pas d'élitisme) car présenter une seconde fois les mêmes individus n'apporte pas de nouvel élément au designer ;
- Le croisement des individus sélectionnés favorise la pensée convergente. On favorise l'apparition d'IHM regroupant des traits intéressants ;
- L'injection de nouveaux individus soutient la pensée divergente et apporte à chaque itération des éléments nouveaux pour inspirer le designer.

Nous avons utilisé le logiciel Magellan et la méthode des entretiens semi dirigés. Nous nous sommes entretenus avec 11 personnes présentant des rapports différents à la conception (Figure 44).

1.1. E0	1.2. Programmeur	1.3. Doctorant, conception web
1.4. E1	1.5. Designer	1.6. Product Designer
1.7. E2	1.8. Expert IHM	1.9. Maître de conférence
1.10. E3	1.11. Expert IHM	1.12. Doctorant
1.13. E4	1.14. Expert IHM	1.15. Doctorant, 20 ans d'expérience industrielle
1.16. E5	1.17. Programmeur	1.18. Senior programmeur
1.19. E6	1.20. Débutant	1.21. Etudiant en IHM
1.22. E7	1.23. Débutant	1.24. Etudiant en IHM
1.25. E8	1.26. Programmeur	1.27. Webmaster
1.28. E9	1.29. Débutant	1.30. Etudiant en IHM
1.31. E10	1.32. Débutant	1.33. Designer graphique

Figure 44 : Participants à l'expérimentation.

Chaque entretien est structuré en trois temps. Tout d'abord, nous demandons aux sujets de remplir un formulaire décrivant leur pratique de conception et comment ils perçoivent la créativité et l'innovation dans leur activité. Nous sommes particulièrement intéressés par les sources d'inspiration utilisées par les concepteurs. Nous souhaitons aussi connaître les éléments que les participants estiment créatifs lorsqu'ils cherchent de l'inspiration.

Dans un deuxième temps, nous présentons le logiciel de messagerie instantanée à l'aide d'un modèle des tâches et de captures d'écrans de plusieurs logiciels similaires (MSN, Yahoo messenger, etc.). Nous leur demandons de produire un maximum de croquis d'IHM pendant 45 minutes. Les sujets ont à leur disposition des feuilles A4 ainsi que des stylos, des crayons et des feutres de couleurs. Cette partie de l'expérience joue un double rôle : 1) placer les participants en situation de conception et les familiariser avec le logiciel de messagerie ; 2) confirmer que les participants n'explorent que peu de pistes. A la fin des 45 minutes, les participants doivent présenter leur production.

Finalement, nous leur présentons Magellan et les laissons l'utiliser en leur demandant de verbaliser leurs impressions. Une fois cette étape terminée, nous questionnons les sujets en suivant une grille d'entretien sur les avantages et inconvénients qu'ils voient à utiliser un logiciel comme Magellan. Pour enrichir la discussion, nous leur présentons aussi le logiciel Picbreeder [Secretan 2008] comme illustration des AGI appliqués au domaine de la génération d'images.

Nous structurons cette synthèse selon les trois temps de l'expérimentation. L'analyse des résultats fait l'objet de la section suivante.

**Première phase.** Les participants ont mentionné s'inspirer d'IHM existantes avant et pendant les phases de conception. Nous distinguons deux types d'exploration de ces IHM :

Une exploration des sites web ou IHM similaires au produit à développer (E1, E7, E8, E9) ou des blogs spécialisés en conception (E0, E6, E8). L'objectif de ce type d'exploration est d'acquérir des connaissances du domaine, de cerner les tendances, d'identifier les problèmes et les solutions retenues. Plusieurs participants ont exprimé l'importance de considérer les tendances car, dans certains cas, la production d'un design trop original par rapport à ces tendances peut amener les utilisateurs finaux à se détourner du produit.

Le second type d'exploration est plus informel. Il s'apparente à une recherche sans but précis menée dans la vie de tous les jours (A1, E3, E4). Par exemple, (E1) proposa au cours de l'expérimentation une IHM à base de très gros icônes pour représenter la liste de contacts. Il expliqua ensuite qu'il s'était inspiré en cela « des gros boutons comme ceux des cadrans de téléphone ». Les designers professionnels (E1, E10) ont aussi mentionné avoir recours à des bibliothèques d'images collectées par leurs soins tout au long de leurs différents projets. Les programmeurs nous ont dit réutiliser souvent des parties de code extraites d'autres projets. Le plus expérimenté d'entre eux (E5) a même précisé qu'il s'était constitué une base de codes pour ne pas avoir à « réinventer la roue » à chaque nouveau projet.

Les participants semblaient particulièrement intéressés par l'esthétique des éléments d'IHM lors de la phase de consultation des exemples de logiciels de messagerie instantanée (logos, images, widgets, etc.). Au contraire, la mise en page n'a été mentionnée comme innovante que par cinq participants.

**Seconde phase.** La plupart des sujets ont suivi un processus de conception itératif. Ils ont commencé par une analyse du problème qui les a conduits à identifier chacun un ou deux problèmes des logiciels de messagerie instantanée existants. Par exemple, (E0, E2, E3, E4) considèrent que le nombre de fenêtres dans ce type de logiciel est trop important.

En général, les participants n'ont produit qu'un seul design lors de cette phase et ont déclaré avoir terminé avant les 45 minutes imparties. Seul (E1) a produit plus d'un design et a déclaré avoir manqué de temps pour explorer tout ce qu'il avait en tête. (E1) est le seul des deux designers à être vraiment rentré dans le jeu et à s'être montré motivé par l'expérience. Ceci peut expliquer pourquoi (E1) a produit de nombreux dessins alors que (E10) n'en a produit que deux. En effet, d'après [Amabile 1983], un concepteur ne peut être vraiment créatif qu'à condition d'avoir de bonnes connaissances du domaine exploré, des capacités créatives et enfin d'être motivé. Il semble que les étudiants, bien que motivés par l'expérience, aient eu quelque appréhension à subir un jugement négatif bien que nous leur ayons à plusieurs reprises précisé que la qualité des dessins fournis ne serait pas prise en compte. Enfin, seuls les deux designers nous ont dit utiliser Photoshop pour affiner leurs idées avant de passer à la production finale. Tous les autres participants ont indiqué qu'ils codaient directement leurs idées, ce qui explique sans doute qu'ils n'en aient produit qu'une seule chacun.

Lors de cette phase, seul (E4) s'intéressa vraiment à l'esthétique en utilisant plusieurs couleurs pour réaliser ses croquis. Les participants ont concentré leurs efforts sur la mise en page générale de l'application et sur les interacteurs. Cette observation correspond bien à ce qui est attendu dans une phase de croquis [Buxton 2007].

**Troisième phase.** En général, les participants ont trouvé Magellan « intéressant mais perfectible » (E1). La critique la plus courante portait sur l'ajout, à chaque génération, de nouvelles IHM générées indépendamment de la génération précédente. Cela créait de la confusion chez les participants qui se demandaient pourquoi certaines propositions faites par Magellan ne correspondaient pas aux choix qu'ils avaient faits à la génération précédente. Cependant, lors de la discussion, il est apparu que ce n'était pas vraiment la génération aléatoire de nouvelles IHM qui posait problème mais plutôt le fait que ces IHM n'étaient pas immédiatement identifiables comme telles.

Une nouvelle utilisation d'un tel outil nous a été proposée par le designer (E1) : celui-ci utiliserait un tel outil pour initier la discussion avec des clients afin de mieux cerner leur demande. Le logiciel serait utilisé comme générateur de produits et permettrait de matérialiser les idées discutées avec le client. Ainsi, (E1) imaginait pouvoir converger petit à petit avec le client sur une définition plus précise du produit à concevoir en modifiant les données du problème et les paramètres des solutions.

Les participants n'ont pas repris naturellement l'activité de dessin comme nous pensions qu'ils le feraient. Ils se sont concentrés sur l'observation des IHM proposées et ont commencé à les critiquer. Plusieurs participants avaient des questions relatives à l'interaction avec les IHM générées. Par exemple : « que se passe-t-il si on ouvre une nouvelle conversation ? ». Même si Magellan n'apportait pas de réponse à ce type de question (l'arbre des tâches du logiciel de messagerie n'étant relié à aucun noyau fonctionnel), ceci permettait tout de même aux participants d'envisager les problèmes potentiels de certaines IHM.

Bien que la mise en page des IHM n'ait été mentionnée comme innovante que par cinq participants lors de la première phase de l'entretien, tous les participants ont sélectionné les IHM proposées par Magellan en se basant sur ce critère. Les participants ont beaucoup critiqué la mise en page mono colonne (mise en page standard en HTML) car elle aboutissait souvent à des IHM difficiles à lire. De même, les IHM n'établissant pas de claire distinction entre les différentes tâches étaient très

critiquées. Les IHM séparant clairement les différentes parties étaient unanimement préférées.

Les participants étaient généralement réticents à sélectionner une IHM dont une partie leur déplaisait vraiment, quand bien même elle contenait d'autres aspects qui leur semblaient intéressants. Cette réticence les amenait à réclamer la possibilité de sélectionner les sous-parties d'IHM qui resteraient fixes au cours de l'évolution et/ou celles sur lesquelles devrait porter l'évolution.

Finalement, nous avons constaté que les participants essayaient souvent de retrouver dans un premier temps des IHM proches de celles qu'ils avaient proposées lors de la première phase de l'expérimentation.

Cette évaluation nous permet de confirmer que l'exploration de nombreuses pistes est difficile à mettre en œuvre pour les designers, en particulier pour les débutants : seul un des participants, designer professionnel, a produit plusieurs IHM lors de la seconde phase de l'expérimentation.

**Niveaux de précision des exemples.** Compte tenu de l'état de l'art des galeries d'exemples, nous pensions initialement qu'en fournissant des IHM précises et non des croquis, les designers seraient d'autant plus inspirés pour explorer de nombreuses pistes. Au terme de notre évaluation, le résultat est mitigé. D'une part, les participants ont apprécié de pouvoir appréhender les différents interacteurs dans le cadre d'exemples de l'IHM à concevoir mais, d'autre part, les couleurs ont plutôt joué un rôle perturbateur dans l'évaluation qu'ils faisaient des IHM proposées par Magellan. En effet, si les IHM « en noir et blanc » apparaissaient rédhitoires aux participants, nous avons noté que les couleurs étaient utilisées pour rejeter des IHM plutôt que pour les sélectionner. Par exemple, les thèmes à dominante rose étaient unanimement rejetés. Ainsi, un bon thème de couleurs n'était pas suffisant pour sélectionner une IHM mais l'était pour la disqualifier.

Cependant, la gestion des couleurs reste intéressante pour les designers si elle n'évolue pas au cours du processus ou si elle résulte de leur choix. Plusieurs propositions furent faites par les participants quant à la gestion des couleurs dans Magellan. Nous reportons ici les trois principales : 1) l'utilisateur devrait pouvoir fixer le thème des couleurs pour toutes les IHM proposées. (E1) et (E10) expliquent en particulier que le thème de couleurs est souvent imposé par le client ; 2) l'utilisateur devrait pouvoir essayer différents thèmes de couleurs pour une IHM

donnée (E7, E8) ; 3) les thèmes de couleurs devraient être éditables pour chaque IHM (E2).

Nous avons constaté que les participants se focalisaient sur la mise en page des IHM et peu sur l'esthétique. Ce constat est valable à la fois pour la phase de croquis sur papier et d'utilisation de Magellan. Cette observation remet en cause le choix de présenter des IHM très détaillées lors des premiers cycles d'évolution au moins. Il serait intéressant d'explorer, dans des travaux futurs, la présentation d'exemples relativement peu détaillés, présentant surtout des mises en pages. Les exemples pourraient ensuite se préciser (en termes d'interacteurs, couleurs, etc.) au fur et à mesure du processus d'évolution.

**Contrôle utilisateur.** En nous inspirant de l'état de l'art des AGI, nous avons donné un contrôle relativement limité à l'utilisateur sur les exemples produits par Magellan (sélection des individus jugés intéressants). Nous pensions que l'impossibilité de modifier directement les exemples pousserait les designers à repasser de temps en temps sur papier pour y coucher de nouvelles pistes d'IHM inspirées par Magellan. Au terme de notre évaluation, il apparaît très clairement que cette approche n'est pas la bonne. Les participants ont unanimement fait part de leur frustration de ne pas pouvoir éditer les exemples produits par Magellan. Outre les couleurs, les participants ont été frustrés de ne pouvoir éditer directement les exemples pour corriger des erreurs mineures ou intégrer leurs propres idées dans les exemples. Les participants identifiés comme designers (E1 et E10) pensaient que la créativité devait venir d'eux mêmes. (E10) en particulier voyait Magellan ou Picbreeder [Secretan 2008] comme des limites posées à sa créativité du fait que l'espace de conception était contraint.

Ce besoin de contrôle sur les IHM présentées s'est aussi manifesté par le fait que les nouvelles IHM injectées à chaque génération étaient au mieux ignorées ou au pire jugées perturbatrices. De plus, les participants ont tous cherché, dans un premier temps, à sélectionner des IHM se rapprochant des croquis qu'ils avaient produits dans la phase précédente. Il semble donc que les designers aient tendance à se focaliser sur la réalisation de la piste d'IHM qu'ils ont en tête. Si le système ne leur permet pas de l'atteindre, cela génère de la frustration. Un besoin de contrôle plus fin sur le processus d'évolution est aussi apparu dans le fait que les participants étaient réticents à sélectionner des IHM dont une partie leur déplaisait. Une évaluation plus fine des IHM proposées, permettant d'identifier précisément les parties intéressantes ou déplaisantes, a été réclamée par les participants.

Une explication possible du fait que les participants aient ressenti le besoin de modifier directement les IHM proposées par Magellan et n'aient pas utilisé le papier pour matérialiser leurs idées relève peut-être de l'aspect dialectique de l'activité de sketching telle qu'elle est vue par [Goldschmidt 1991]. D'après elle, le processus de sketching peut être vu comme un dialogue entre le concepteur et le croquis. Le designer matérialise ce qu'il a en tête et est, en retour, inspiré par l'interprétation qu'il fait de cette matérialisation. Dans ce cadre, la matérialisation et l'interprétation sont indissociables. Dans notre expérience, les IHM produites par Magellan pouvaient être interprétées par les participants ; ceux-ci avaient ensuite la possibilité de matérialiser leur inspiration sur papier mais l'esquisse produite n'était plus alors intégrable dans Magellan. Nous pensons que cette dissociation est à l'origine du blocage ressenti par les participants.

**Inspirer, éliminer.** Selon [Buxton 2007], l'activité de design est une succession de phases de génération et de sélection d'idées. Nous avons constaté que Magellan permettait aussi d'éliminer des pistes que les participants avaient en tête. Plusieurs participants se sont rendu compte que leur idée posait problème en la voyant matérialisée par Magellan. Par exemple, de nombreux participants ont constaté que l'ouverture de nouvelles conversations pouvait poser problème pour leur design en raisonnant à partir des exemples produits par Magellan.

Plus généralement, ces résultats peuvent être rapprochés de ceux obtenus par [Tohidi 2006] Ces derniers montrent que la présentation de plusieurs alternatives de designs permet de mieux critiquer ces dernières. Ainsi, dans la même veine que ce qui a été suggéré par (E1), Magellan peut servir de support à la discussion dans le cadre de travaux en groupes.