

SAMUEL HURON

LA VISUALIZATION CONSTRUCTIVE

UN PARADIGME DE DESIGN DE VISUALIZATION QUI PERMET D'ASSEMBLER DES
REPRÉSENTATIONS VISUEL DYNAMIQUE POUR PERSONNES NON EXPERTES.
SYNTHÈSE EN FRANÇAIS. (VERSION COMPLÈTE EN ANGLAIS EST ACCESSI-
BLE EN LIGNE.)

Copyright © 2015 Samuel Huron

Licensed under the Apache License, Version 2.0 (the “License”); you may not use this file except in compliance with the License. You may obtain a copy of the License at <http://www.apache.org/licenses/LICENSE-2.0>. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an “AS IS” BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

First printing, 2014 July 29

Second printing, 2014 September 15

Abstract

Durant les vingt dernière années, la recherche en visualisation d'information a créé de multiples techniques pour supporter les analyses de données pour la science, l'industrie et les gouvernements. Cela a permis à un grand nombre de tâches d'analyse d'être exécutées, des tâches variant selon les types et les volumes de données impliqués. Cependant, la majorité de la recherche s'est concentrée sur des données statiques, et les visualisations de données on tendances a être utilisées uniquement par des utilisateurs experts. Plus récemment, des changements sociaux, ainsi que des avancées technologiques ont produit des corpus de données des plus en plus dynamiques, et sont utilisés par une audience de plus en plus large. Un exemple de ce type de flux de données dynamiques inclu les e-mails, les mises à jour de statuts sociaux, les flux RSS ¹, les systèmes de versioning, les réseaux sociaux et bien d'autres.

Ces nouveaux types de données sont utilisés par une population de personnes qui ne sont pas forcément formées à la visualisation d'informations. Certaines de ces personnes peuvent être considérées comme des utilisateurs occasionnels, alors que d'autres peuvent être profondément impliqués avec les données. Dans ces deux cas, ces personnes n'ont pas reçu de formation formelle en visualisation d'information. Pour la simplicité, dans cette dissertation, je ferai référence à ces personnes (utilisateurs occasionnels, novices, experts des données) en utilisant le terme non-expert.

Les changement sociaux et technologiques ont soulevé de multiples défis, d'autant plus que la plupart des modèles et techniques de visualisation ont été conçus pour des experts et des corpus de données statiques. Peu d'études ont été conduites qui explorent ces défis. Dans cette thèse, j'adresserai cette question: *Peut-on autonomiser des non-experts dans leur utilisation de la visualisation afin de leur permettre de contribuer à l'analyse de données dynamiques autant qu'à la création de leur propre modèle de visualisation?*

Le premier pas pour répondre a cette question est de déterminer si cette population non entraînée à la visualisation de données et la sciences des données, peut conduire des tâches utiles d'analyse de données dynamiques en utilisant des systèmes de visualisation d'information adaptés. Dans cette première partie de la dissertation, je me suis concentré sur plusieurs scénarios et systèmes où différentes tailles (de 20 a 300 puis de 2000 à 700 000) de population de non experts utilisent des visualisations d'informations dynamiques afin

¹ RSS est un format de syndication de contenu sur le Web, cette acronyme correspond a "Really Simple Syndication", <http://validator.w3.org/feed/docs/rss2.html>

d'analyser des données dynamiques.

Un second important problème est le manque de principe de conception générique pour l'encodage visuel de visualisation dynamique. Dans cette dissertation je présente la conception, la définition et l'exploration d'un espace de design afin de représenter des données dynamiques pour non-experts. Cet espace de design est structuré par des jetons visuels représentant des éléments de données, ces derniers sont le matériel de construction de base qui permet l'assemblage dans le temps de différentes visualisation, des plus classiques, aux plus innovantes. Jusqu'à nos jours, la recherche dans l'encodage visuel s'est majoritairement concentré sur des données statiques, pour des tâches spécifiques, laissant des approches génériques et dynamiques non explorées et non exploitées.

Dans cette thèse, je propose la construction comme un nouveau paradigme de création de visualisation simple et dynamique a l'intention des population non expertes en infovis. Ce paradigme est inspiré par des théories établies en psychologie du développement autant que des pratiques passées et présentes de création de visualisation avec des objets tangibles. Je décris les composants conceptuels et les processus sous-jacents à ce paradigme. Ces description permettront de rendre plus facile l'étude et le support de ces processus pour un grand nombre de visualisations. Finalement, J'utiliser ce paradigme ainsi que des jetons tangibles afin d'étudier *si* et *comment* une population de non-experts créer des encodages visuels sans logiciel. En résumé cette thèse contribue a la compréhension des visualisations dynamiques pour les utilisateurs non experts en infovis.



Figure 1.2: Exemple d'un sablier. Some Rights Reserved ("running out of time" by Rui Malheiro) <http://goo.gl/1rLU30>

d'un récipient à l'autre. Pendant ce temps, un observateur extérieur peut constamment comparer le volume de sable dans chaque récipient. Cette comparaison visuelle entre les deux piles de sable permet d'évaluer le temps écoulé entre le retournement du sablier et le moment où le sablier finira de s'écouler. Dans ce cas, la représentation visuelle est mise à jour constamment par l'écoulement du sable. Pour cette raison, un sablier est une représentation dynamique. Cependant, de nos jours, il semble que nous utilisons plus souvent des outils numériques que des sabliers pour représenter visuellement des informations abstraites.

Les outils numériques jouent un rôle important dans notre capacité à produire, collecter et stocker des données. Ce phénomène est en augmentation rapide et en plus d'augmenter en volume, il augmente aussi en fréquence de mise à jour, en diversité de données et en variété de publics.

Il y a quelques décennies, la publication de données était asynchrone, et stockée dans des médiums non connectés. En raison de la démocratisation des interfaces informatiques, mais aussi de l'Internet et du Web, l'activité de publication des données est devenue de plus en plus synchrone, connectée et produites dans des formats numériques hétérogènes.

Par conséquent, la fréquence de publication de données croît rapidement, si bien que, dans certains domaines, on peut constater l'existence de flux de données constamment mis à jour. Ces flux de données peuvent être mises à jour en temps réel à différents rythmes, comme observé pour les systèmes tels que Twitter, les flux RSS, les salons de discussions etc. Néanmoins, nos capacités cognitives individuelles restent constantes et limitées et n'évoluent pas au même rythme que la production de ces données. De plus, le traitement de ces données pour l'analyse devient de plus en plus difficile et coûteux. Ce traitement exige souvent un haut degré d'expertise, de calcul, de temps et d'énergie. Une réponse possible à ces limitations pourrait être de créer des outils qui permettent à chacun d'augmenter leur propre capacité à traiter les données de manière synchrone.

Au cours des dernières années, certains systèmes visuels tels que Many Eyes [Viégas et al., 2007], et plus tard Tableau Public [Tableau, 2014] ont commencé à explorer la façon dont Internet pouvait être utilisé afin de démocratiser [Viégas et al., 2007] la visualisation de l'information. Ces systèmes permettent d'augmenter l'accessibilité des données à ceux qui ont des connaissances restreintes dans le domaine de la visualisation. Ces plateformes permettent à toute personne intéressée par l'analyse de données, de créer, de publier et de discuter des visualisations de leurs propres données. D'autres systèmes plus récents, tels que Comment Space [Willett et al., 2011] permettent à des populations similaires, d'annoter, de taguer et de commenter des visualisations afin d'aider l'analyse collective d'autres types de données.

Ces systèmes permettent d'extrapoler un avenir dans lequel la visualisation serait un nouveau langage visuel permettant l'analyse

de données. Et où tout un chacun pourrait être en mesure de lire et écrire à la volée des visualisation de données, à la même vitesse et fluidité que nous le faisons pour les langues naturelles.

Cependant, ces outils pour créer, publier et discuter des visualisation d'informations sont le plus souvent statiques et ne répondent pas aux exigences spécifiques de traitement de flux de données constamment mises à jour. De plus ces mêmes outils sont souvent limités à des visualisations établies, et définis à l'avance dans le programme. Ils ne permettent pas à des personnes non expertes de composer leur propre représentation visuelle.

Ce problème de la difficulté d'accès des outils de visualisation aux population non expertes a été identifiée par plusieurs chercheurs comme l'un des enjeux majeurs dans le domaine de la visualisation d'informations. Par exemple, il est écrit dans un rapport de la NSF sur les défis de la recherche : " Un grand défi est de créer [...] un système qui, tout en n'étant manifestement pas exhaustive et tout-puissant, puisse aider des utilisateurs non-experts à effectuer des tâches [...] dans un laps de temps raisonnable. [...] L'objectif est de faire de la visualisation un outil omniprésent qui permet aux gens ordinaires de penser visuellement dans des activités quotidiennes " [Johnson et al., 2006]. De même des chercheurs du domaine tels que Heer et Shneiderman soulignent en particulier la nécessité de créer de nouvelles interfaces pour les spécifications de visualisation: " Créer de nouvelles interfaces pour spécifier des visualisation reste un besoin. [...] De nouveaux outils nécessitant peu ou pas de programmation devrait placer le design de visualisation personnalisée dans les mains d'un plus large public." [Heer and Shneiderman, 2012].

Les question que j'adresse dans cette thèse est: **Comment pouvons-nous améliorer l'autonomisation de personnes non-expertes dans leur utilisation de la visualisation afin qu'un large public puisse contribuer à l'analyse de flux de données, tout en créant leurs propres visualisations?**

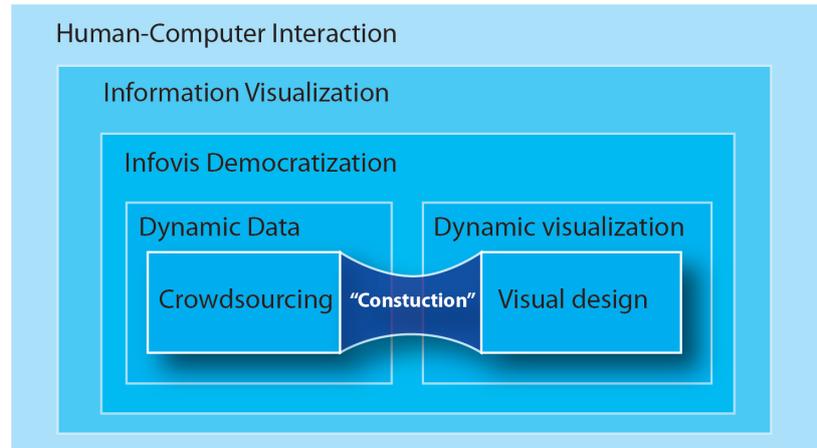
Cette question implique différents niveaux de travail: le niveau social et le niveau de l'interface homme-machine. Au niveau **social**, l'un des principaux défis est de savoir comment engager les non-experts dans un processus d'analyse dynamique utile. Au **niveau de l'interface homme-machine**, nous avons besoin de créer des méthodes et des outils qui permettent aux non-experts de lire et de manipuler des représentations de données qui soient à la fois simples, dynamiques et expressives.

Pour introduire cette thèse je présenterai d'abord le domaine de recherche sous l'angle de la (**démocratisation de représentations visuelles dynamiques**) pour des données dynamiques dans Section 1.1. Ensuite, je définirai les problèmes dans Section 1.2. Ceci nous conduira à la description de notre approche et des méthodes de recherche utilisé pour cette thèse Section 1.3. Enfin les contributions de ce travail seront résumées dans Section 1.5, finalement, Sections 1.4 and 1.6 contiendra un bref aperçu de la structure de la thèse.

"Lorsque les chercheurs en visualisation parlent de passages à l'échelle, cela signifie généralement faire une visualisation avec un grand ensemble de données. [...] Nous considérons un point de vue alternatif: Au lieu de considérer la mise à l'échelle de la taille des données, qu'est ce qui se passe lorsque nous augmentons la taille des utilisateurs?"
Fernanda Viégas et al. [2007]

1.1 Cadre de recherche

Figure 1.3: Research Scope.



Le sujet de recherche de cette thèse se situe dans le domaine de la visualisation de l'information (InfoVis). L'InfoVis est un domaine qui est lié à l'étude de l'interaction homme-machine (??) qui s'efforce de comprendre et de fournir des solutions technologiques pour l'analyse visuelle et le traitement des données. La recherche dans le domaine de la visualisation d'information est fondée sur le fait que les capacités cognitives de vision chez l'être humain sont particulièrement efficaces pour traiter des informations en un temps extrêmement court. L'InfoVis est maintenant un vaste domaine de recherche interdisciplinaire, incluant différentes disciplines telles que la psychologie cognitive [Treisman, 1985, Healey, 2007, Ware, 2004], les statistiques [Beniger and Robyn, 1978, Wilkinson et al., 2006], l'informatique [Card et al., 1999, Spence and Press, 2000, Shneiderman, 1996], le design [Vande Moere and Purchase, 2011, Neurath, 1939, Bertin, 1973, Tufte and Graves-Morris, 1983] et l'infographie.

La visualisation d'information est définie par Stuart Card [Sears and Jacko, 2007] comme "un ensemble de technologies qui utilisent les capacités préattentive de la vue pour amplifier les capacités cognitives humaines dans la manipulation d'informations abstraites". Durant les vingt dernières années la visualisation d'information a majoritairement été utilisée par des experts pour représenter, traiter, analyser et explorer des données majoritairement pour des organisations industrielles, scientifiques, et gouvernementales. Comme ces technologies ont montré leur efficacité elles commencent à être adoptées par une audience de plus en plus large et venant de domaines disciplinaires différents.

Nous pouvons de nos jours rencontrer des visualisations sur des blogs personnels, dans des oeuvres artistiques, ou dans la presse papier et télévisuelle ainsi que dans les communautés de quantification de soi. Cette adoption de la visualisation par différents groupes amène de nouveaux besoins applicatifs spécifiques à cette population. Le Design et la recherche doit maintenant créer de nouvelles technologies pour adresser ces nouvelles populations.

Il n'existe pas encore un sous domaine de la visualisation d'information qui adresse ce problème spécifique. Des pionniers tel que Fernanda Viegas Viégas *et al.* [2007] utilise le terme "démocratisation" pour définir l'utilisation de technologies d'analyses visuelles par des personnes non expertes en Infovis. L'objectif de ces outils est d'augmenter les capacités de tous afin de les aider à produire du sens à partir de leur données en utilisant des techniques de visualisation. Le défi de la création de nouveaux outils de visualisation d'informations, qui sont facilement accessibles à tout le monde, a été soulignée précédemment par plusieurs autres chercheurs [Heer and Shneiderman, 2012, Moorhead *et al.*, 2006].

Par coïncidence, en même temps que la démocratisation de la visualisation devient un phénomène de plus en plus présent, les services dynamiques de flux de données deviennent de plus en plus populaires. Les données dynamiques, peuvent être définies comme des séquences d'objets typés mise à jour au fil du temps. Ce type de données est très commun, par exemple les mises à jour de réseaux sociaux ou les tweets, les courriels, les flux RSS, les flux vidéo, etc. Aussi à un niveau plus technique on peut prendre comme exemple les mises à jour des journaux d'événement (log) dans les réseaux ou encore les événements des systèmes de contrôle de version distribués. Ces données présentent de multiples défis pour la visualisation car ils *apparaissent* à des moments imprévisibles, *s'accumulent* au fil du temps, et peuvent être très *hétérogène*, contenant par exemple différents médias tels que textes, images et vidéos.

Durant cette dissertation nous porterons notre attention sur la **conception de visualisations dynamiques pour des personnes non-expertes dans le domaine de la visualisation d'informations** (Figure 1.4). Mon étude va adresser trois différents aspects de cette approche:

(1) Si oui, et comment, une personne non-experte peut analyser des flux de données dynamiques. Notre motivation derrière ce premier point est de démontrer que l'utilisation des visualisations dynamiques simple pour des tâches d'analyse de données peut permettre la récolte de métadonnées qui pourront être utiles pour des analyses futures.

(2) Comment, peut-t-on créer des techniques de visualisation dynamique afin d'augmenter les capacités de gérer des données dynamiques pour les personnes non-expertes en InfoVis. Ce deuxième aspect est motivé par le manque de bonnes pratiques, de techniques et de savoir dans ce domaine. Effectivement, peu de techniques existent pour guider les personnes non expertes en Infovis dans leurs activité de lecture et d'écriture de de visualisation dynamique.

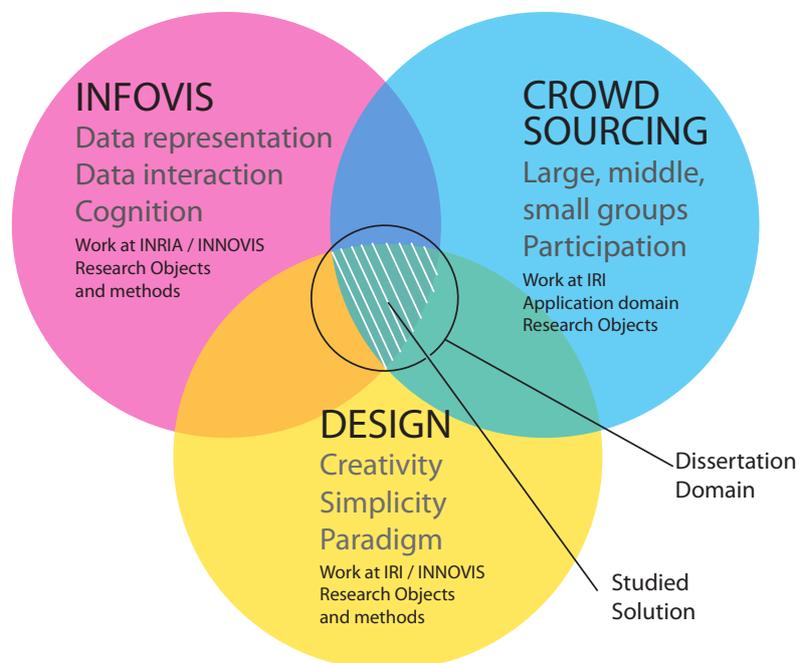
(3) Comment, peut-on fournir des moyens aux personnes non expertes pour créer leurs propres représentations visuelles dynamiques. Ce dernier aspect est motivé par le manque d'outils existant pour cette population et ce types de données.

1.2 Problem statement

La plupart d'entre nous utilisons des flux de données dynamique dans notre vie quotidienne, à une échelle individuelle autant qu'à une échelle collective. Par exemples les mises à jours de statuts dans les réseaux sociaux, la presse, les courriels, et des données plus personnelles comme nos données bancaires composées de ligne de crédit et débits.

L'ensemble de ces données dynamique est mis à jour à différentes fréquences, de plusieurs fois par seconde à juste une fois par mois. D'autres sont plus ou moins publiques et peuvent être utilisées collectivement ou individuellement, certains sont privés et personnels. Les tâches de traitement des ces flux de données peuvent varié en complexité. Par exemple, taguer, aimer, ou annoter dans un réseaux social est plus simple qu'opérer la mise à jour visuelle d'une représentation.

Figure 1.4: Domain Scope.



La question adressée ici est: Comment des populations différentes peuvent manipuler des données dynamiques? Afin d'adresser cette question nous allons porter notre attention sur deux problèmes:

PROBLÈME 1 *Est-il possible d'engager un public non expert en visualisation dans un processus dynamique d'analyse de données?*

Cette recherche commence par étudier comment on peut crowdsourcer l'analyse de données dynamiques lors d'un événement avec le soutien de systèmes de visualisation d'informations. Ce problème correspond à un besoin applicatif de l'IRI: rassembler et exploiter des métadonnées spécifiques lors d'un événement. Cela implique plusieurs défis et problèmes à prendre en compte: (P1.1) d'abord, est-il possible de tirer parti d'une foule de non-experts en infovis pour analyser des données dynamiques gratuitement? et (P1.2) deuxièmement, comment peut-t-on représenter visuellement des données dynamiques pour ce public spécifique?

PROBLÈME 2 *Comment peut-on démocratiser la création de visualisation dynamique?* Les solutions découvertes pour le (PROBLÈME 1) nous ont permis de définir une technique de visualisation qui permet à des designers ou ingénieurs d'engager un public non expert dans la production de métadonnées. Cependant, nous souhaitons cibler notre recherche aussi sur la possibilité de trouver des solutions qui permettent au public non experts de créer eux-mêmes leurs propres visualisations dynamiques. Cela implique différents sous problèmes: (PROBLÈME 2.1) Quels sont les défis pour créer un outil de création de visualisation pour personnes non-expertes? (PROBLÈME 2.2) Avec quels moyens peut-on aider le processus d'encodage visuel pour cette population? (PROBLÈME 2.3) Et plus que tout, peut-on mieux comprendre comment l'encodage visuel est généré ?

1.3 Approche de recherche

La visualisation d'information est un sous domaine de l'informatique. Ce domaine de recherche implique l'étude d'artefact et de phénomènes partiellement créés par les personnes qui l'étudient. Cette particularité fait de ce domaine une science de la conception (design) [Simon, 1969].

Les méthodes de conception (design) [Howard et al., 2008] dans le domaine de la visualisation d'information commencent à peine à être discutées [Munzner, 2009, Vande Moere and Purchase, 2011, Sedlmair et al., 2012, Goodwin et al., 2013], de plus la conception créative [Howard et al., 2008] et les méthodes d'idéation ne sont pas

définies dans ce domaine spécifique.

Afin de répondre à mes besoins durant les phases creatives de conception d'artefact, j'ai importé des méthodes de différents domaines. Ces méthodes sont l'usage de la métaphore [Gentner, 1983, Blackwell, 2006], la mise en place et l'exploration d'espaces de conception (design space) [Pugh, 1991, Laseau, 2001, Greenberg et al., 2011], l'utilisation de combinaisons et de mutations de systèmes [Cross, 1997] et aussi l'inspiration [Howard et al., 2008, Osborn, 1953].

En plus des méthodes de conception créatives appliquées à mes recherches, j'ai aussi appliqué un ensemble de différentes méthodes de recherche comportementale. Runkel and MacGrath [1972], McGrath [1995] caractérise les différentes stratégies de recherche pour étudier le comportement humain en utilisant un modèle à plusieurs dimensions: (1) des opérations de recherche les plus intrusives aux moins intrusives, (2) des observations les plus généralisables au moins généralisables. Ce modèle décrit également les différentes implications liées aux méthodes d'observation: la représentativité des acteurs, la précision des mesures, la possibilité de systématiser, ainsi que le contexte. Chacune de ces stratégies offre différents avantages et limites. Voici la liste des différentes méthodes que j'ai utilisées pour la conception (*en italique*) et pour l'observation de phénomènes:

Collection. Le processus de créer des collection afin de faciliter l'inspiration est une méthode de conception créative répandue [Herring et al., 2009, Howard et al., 2008]. Ce processus permet d'identifier les caractéristiques ou encore les formes d'intérêts ainsi que des fonctionnalités souhaitées qui n'existent pas et correspondent à certains besoins. J'ai utilisé cette méthode Section 1.6 afin d'inspirer mes choix de design, notamment l'utilisation d' "élément individuel et discret". J'ai aussi utilisé cette méthode comme support du paradigme que je présente dans cette thèse Section 1.6.

Expérience de terrain. Selon McGrath [Mcgrath, 1995] les expériences de terrain sont définies comme "travailler au sein d'un système naturel en cours aussi discrètement que possible, à l'exception de la manipulation d'une caractéristique majeure de ce système.". Cette approche fournit une validation externe forte, mais se limite à un contexte et un système étroit. Dans cette recherche, je souhaitais étudier si le public qui assiste à un événement peut exécuter des tâches d'analyse dynamiques des données. Afin d'étudier ces phénomènes au plus près de la population ciblée, j'ai choisi de mettre en place des expériences de terrain comme une stratégie de recherche plus discrète. Les raisons justifiant ce choix étaient qu'il est trop complexe de simuler l'engagement d'une foule dans une expérience de laboratoire, et qu'observer le phénomène en situation réelle apporte de multiples bénéfices.

J'ai utilisé cette méthode pour Section 1.6 et Section 1.6. Dans ce premier cas, j'ai utilisé une approche semi formelle, et dans le second une approche informelle.

Combinaison. Selon Cross [1997]: “La conception créative peut être génée en combinant différentes propriétés de différents designs existant en une nouvelle combinaison ou configuration.”.

La combinaison permet à chacun d’augmenter un système existant en lui ajoutant des propriétés ou fonctionnalité d’un autre, résultant un troisième nouveau système qui contient les propriétés des deux systèmes parents. Cette méthode permet d’étendre un système en augmentant ses propriétés. J’ai utilisé cette méthode dans différentes partie de cette dissertation tel que l’application de simulations physiques à un système de jetons dans Section 1.6, ou encore la combinaison de systèmes de jetons et de diagramme de surface dans Section 1.6.

Metaphore. Gentner [1983] décrit le processus comme le couplage entre des propriétés de deux structures (??). Les Métaphores sont fréquemment utilisées dans le domaines des interactions humain machine. Blackwell [2006] le relate en détail en proposant une histoire des métaphores dans ce domaine. Une métaphore peut être utilisée pour deux choses, comme une technique générative pour aider la conception créative [Cross, 1997], ou comme un support d’explication des propriétés d’un système [Carroll et al., 1987]. J’ai utilisé cette technique pour ces propriétés génératives et explicatives dans Section 1.6.

par observation. Ce type d’étude est une approche qualitative de l’analyse d’une situation, ces méthodes sont décrites par Creswell [2012].

Nos études par observation ont été conduites dans une configuration de laboratoire avec des contraintes déterminées. Ce type d’étude ne permet pas d’avoir des mesures précises ou des résultats généralisables. Néanmoins, elle offre un moyen d’identifier des phénomènes, des observations riches en détail, ainsi que des directions pour des recherches futures. Nous avons utilisé cette méthode afin d’obtenir une meilleure compréhension des procédures d’encodage visuel dans le paradigme de la visualisation constructive. Les détails de cette recherche sont décrits dans Section 1.6.

1.4 Parcours de recherche et collaboration

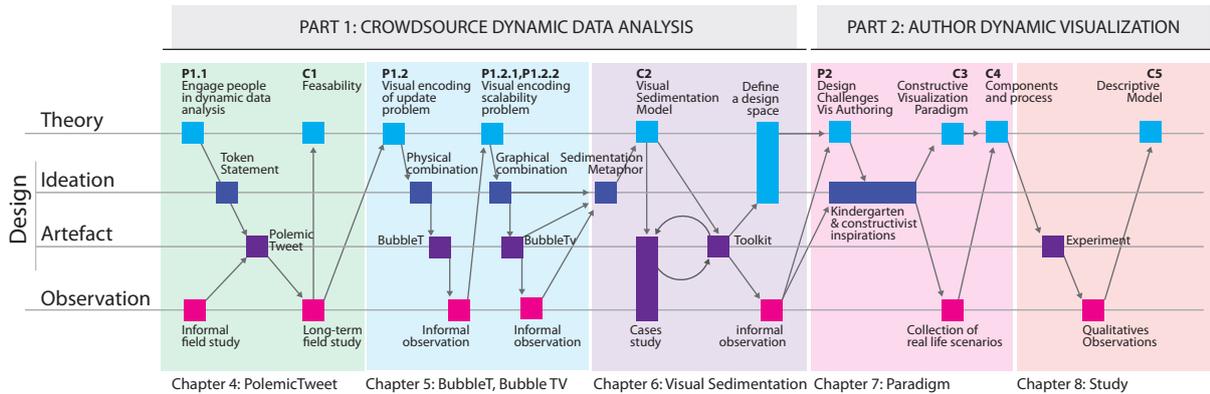


Figure 1.5: Vue d'ensemble de mon parcours de recherche au travers d'une version adaptée du modèle de triangulation de MacKay et Fayard. Chaque boîte est un phénomène et chaque ligne correspond à un mouvement d'un phénomène à un autre.

Dans cette section je présente les collaborations que j'ai eu le plaisir de mener durant cette recherche. Je présente aussi mon parcours de recherche et d'investigation qui m'a amené à faire les projets présentés dans cette dissertation.

1.4.1 A propos des collaborations

Durant ces trois dernières années, j'ai eu le plaisir de travailler et d'échanger avec des personnes aux sensibilités, formations et éducation très différentes.

Ces collaborations et échanges ont apporté à mon travail de recherche d'incalculables éléments de réflexions et d'inspiration. A chaque début de chapitre je note dans la marge le nom des collaborateurs avec qui j'ai travaillé et le nom de la publication concernée. Cependant, cette thèse reflète mon point de vue personnel et individuel sur le sujet. Cette thèse est écrite à partir de ma perspective en tant que chercheur et individu. Chaque projet a été initié par moi, et dans chacun de ces projets, j'ai tenu les rôles de designer et de chef de projet. Dans la section suivante, je vais décrire le rôle de chaque collaborateurs ainsi que le mien pour chaque projet. Je souhaite que le lecteur garde à l'esprit que ces collaborations ont été le lieu d'échange intellectuels extrêmement fructueux où l'ensemble est toujours supérieur à la somme des parties.

1.4.2 Parcours de recherche

La recherche dans le domaine des interactions humain machine implique différentes disciplines, telles que le design, les sciences humaines, l'informatique, et les sciences expérimentales. Durant cette recherche, j'ai itéré au travers de plusieurs cycles suivants: identifier un problème, générer des idées, concevoir des artefacts, introduire ces artefacts dans la situation visée, observer l'impact de ces artefacts, et finalement théoriser sur l'effet de cet impact. Afin de décrire

le parcours que j'ai fait parmi ces multiples activités, j'ai étendu le modèle de triangulation de Mackay and Fayard [1997].

Cette extension me permet de représenter les activités d'observations, conceptions, et théorisations dans un seul modèle. Afin d'étendre ce modèle, j'ai séparé la dimension "design"(conception) en deux branches "design d'artefact" et "idéation". Idéation se réfère au processus de conception créative [Cross, 1997, Howard et al., 2008], de génération d'idées [Pugh, 1991, Laseau, 2001, Greenberg et al., 2011], et de raffinement de nouveaux designs. Dans cette section je vais utiliser ce modèle pour décrire les relations entre ces activités et mon processus de recherche.

Sur Figure 1.5 nous pouvons observer que le processus est divisé en deux colonnes dédiées à chacun des problèmes présentés dans la section précédente. La première partie est sur la conception et le déploiement de différents systèmes de "crowd-sourcing" qui nous ont permis d'évaluer la possibilité d'engager des utilisateurs non expert dans l'analyse de données dynamiques. La seconde partie est concentrée sur la définition et l'étude du paradigme de la *visualization constructive* dans le contexte de la création de visualisation dynamique par des utilisateurs non-experts. Ces deux parties sont elles-mêmes divisées en phase de recherche qui correspondent chacune à un chapitre de cette thèse et sont décrits ci dessous.

J'ai commencé ce travail à partir de la question de recherche initiale (Figure 1.5,P1): "*Peux-t-on engager des personnes non-expertes en infovis dans un processus d'analyse de données dynamique?*" Afin d'étudier ce problème trop général, j'ai réduit cette question à un cas d'application spécifique: l'annotation et le marquage de vidéos de conférence. A partir de différentes observations informelles, j'ai conçu un premier système de visualisation basé sur le concept de jeton "PolemicTweet". Ensuite j'ai ¹ déployé, étudié et évalué ce système pendant une longue période. A partir de ces observations j'ai identifié des limitations dues à l'encodage visuel de la mise à jour de données (Figure 1.5,P1.2). Afin d'explorer les solutions de ce problème, j'ai combiné un encodage visuel à base de jeton avec une simulation physique, le résultat de cette combinaison est le design du système "Bubble-T" (Section 1.6). Ce système est une application publique fournissant une représentation visuelle des tweets arrivant durant une courte période de temps. Après avoir déployé cette application sur le web j'ai ² reçu des retours positifs qui nous ont encouragés à poursuivre dans cette approche. Ce système nous a aussi permis d'identifier quelques limites liées à cet encodage visuel. Les problèmes majeurs sont les suivants :(Figure 1.5,P1.2.1,P1.2.2): (i) la scalabilité de la représentation visuelle, une représentation à base de jeton est limitée à un encodage de un pour un. Cette propriété limite le volume de données que peut représenter un tel encodage dans un espace limité.

D'un autre côté des représentations visuelles telles que des diagrammes de surface permettent l'encodage visuel de différentes échelles. Afin d'adresser cet enjeu J'ai ³ décidé de combiner ces deux systèmes d'encodage visuel: jeton et diagramme de surface. Le résultat

¹ en collaboration avec, pour l'implémentation: of Raphael Velt, Karim Hamidou, Yves Marie Haussonne, pour le déploiement: Florence Duc, Nicolas Sauret, Vincent Puig, et pour l'analyse et l'évaluation: Petra Isenberg, Jean Daniel Fekete.

² avec la collaboration pour le design et développement de: Raphael Velt, Romain Vuillemot, Yves Marie Haussonne

³ en collaboration avec: Nicolas Sauret, Raphael Velt and Romain Vuillemot

tat de cette combinaison a donné un deuxième système “Bubble-TV”.

⁴ avec la collaboration pour le design et l’analyse, Romain Vuillemot and Jean-Daniel Fekete.

Afin de généraliser les encodages visuels de ces deux systèmes nous avons conçus une métaphore basée sur les jetons: “Visual Sedimentation” (Figure 1.5,C2). J’ai ⁴ raffiné cette métaphore en développant différentes études de cas et une boîte à outils. Cette boîte à outils permet à des personnes ayant des connaissances en programmation de créer des visualisations dynamiques de flux de données.

Cette première partie de mon travail était concentrée sur deux problèmes majeurs: **PROBLEME1.1** comment engager des personnes dans un processus d’analyse de données dynamiques, et **PROBLEME1.2** comment visuellement représenter des mises à jours de données dynamiques. Dans la première partie de cette recherche, j’ai exploré différentes visualisations basées sur des jetons avec pour objectif de créer des systèmes d’analyse de données dynamiques pour non-expert.

⁵ avec Sheelagh Carpendale, et la collaboration de Alice Thudt, Anthony Tang et Michael Maurer.

Dans la seconde partie de cette recherche j’ai généralisé l’approche basée sur les tokens à la création de visualisation dynamique (**PROBLEME2**). Premièrement j’ai ⁵ définis les défis de conceptions relatifs aux outils de création de visualisation pour utilisateur non-expert (Figure 1.5,P2). Je me suis inspiré de théories de la psychologie du développement [Piaget, 1989] ainsi que des outils pédagogiques développés durant l’invention de l’école maternelle [MacCormac, 1974]. A partir de ces inspirations j’ai tout d’abord défini le paradigme de *visualization constructive* (Figure 1.5,C3) et collecté des observations sur des exemples historiques ou de la vie courante. A partir de ces exemples j’ai extrait les composants et processus de la *visualization constructive* (Figure 1.5,C4).

⁶ en collaboration avec Yvonne Jansen et Sheelagh Carpendale

Afin d’étudier ce paradigme j’ai ⁶ conçu un protocole expérimental afin d’observer les comportements que ce paradigme peut engager. Les observations récoltées lors de cette expérience nous ont permis de produire un modèle des processus d’encodage visuel par assemblage de jetons chez les utilisateurs non experts (Figure 1.5,C5).

1.5 Contributions

Cette recherche a produit des contributions de différentes sortes: la **première phase de recherche** (Part II) est concentrée sur un scénario d’engagement d’un groupe dans une activité d’analyse de données dynamiques: concevoir et déployer des applications durant des événements avec l’intention d’engager le public dans des processus d’analyse de données dynamiques. La **seconde phase de recherche** (Part III) est concentrée sur la généralisation des principes de conception trouvés dans la première partie.

- **Part II:** Analyse de données dynamiques sociale

- J’ai rassemblé des preuves expérimentales qui supportent le fait que: (A) il est possible d’engager des personnes lors d’un événement afin qu’elles prennent en charge des tâches d’analyse sur des données dynamiques (B) que ces types d’analyse sur des

données dynamiques peuvent être efficaces, (C) que les données produites peuvent être utiles par la suite. Ces contributions sont des éléments importants pour la résolution du **PROBLEM 1.1** comme il est décrit dans Section 1.6.

- J’ai identifié les besoins pour, et ai conçu plusieurs systèmes de “crowdsourcing” en temps réel par la participation du public assisté de visualisation d’information. Ces systèmes ont été déployés dans différentes situations, de groupes de taille moyenne (20 personnes), à des groupes de plus grandes tailles (300 personnes), à de très grand groupes (700 000 personnes). Une partie de cette contribution est aussi d’avoir fait un pas dans le sens de la scalabilité des visualisation et du scénario utilisé (**PROBLEM 1.1** et **PROBLEM 1.2**). Ceci est discuté dans Section 1.6 et Section 1.6.
- J’ai utilisé une métaphore pour créer une nouvelle famille de techniques de visualisation qui permet l’encodage visuel des mises à jour de données dynamiques à plusieurs échelles de temps. Cette métaphore est générative et me permet d’offrir une classe de nouvelles visualisations. Cette contribution adresse le **PROBLEM 1.1** et est majoritairement discutée dans Section 1.6 et Section 1.6.
- **Part III:** La démocratisation de la conception de visualisation dynamique
 - J’ai identifié un paradigme de visualisation constructive. Ce nouveau paradigme de conception fournit un moyen pour les non experts de créer leurs propres visualisations dynamiques. La communauté InfoVis a déjà reconnu que la création d’un outil simple pour les utilisateurs non-experts est important, mais l’application de principes constructifs à la conception de visualisation dynamique n’a jamais été examinée auparavant.
 - J’ai créer une description opérationnelle du paradigme de *visualisation constructive* qui peut être utilisée dans la conception et construction de nouveaux systèmes des visualisation d’informations. J’ai créé un modèle conceptuel basé sur l’étude de comment des personnes non expertes construisent des visualisations à partir de jetons. Ce modèle défini les composant majeurs et le processus général que suit un auteur qui fait une visualisation constructive. J’ai aussi décrit l’utilisation de ces composant et processus dans quatre scénarios de la vie réel. Cette contribution est en réponse au **PROBLEM2.2**, et discuté dans Section 1.6.
 - J’ai rassemblé des preuves empiriques soutenant que des personnes non expertes peuvent créer, mettre à jour, annoter et discuter des visualisations par assemblage de jetons. J’ai exploré dans les détails des tâches logiques et des actions liées à cette activité d’encodage visuel. Enfin, j’ai extraite des ces observations un

modèle qui décrit les actions, leurs objectifs et leurs relations. J'ai également étudié les séquences d'actions de construction et ai observé une diversité très élevée. Cette contribution est en réponse au [PROBLEM2.3](#) et est discutée dans Section 1.6.

- J'ai effectué une première exploration des détails sur les tâches et les actions logiques d'un phénomène précédemment considérés comme une "boîte noire": le processus d'encodage visuel. Je définis le premier modèle d'encodage visuelle dans un environnement constructif tangible. Ce modèle décrit les actions, leurs objectifs et les interactions qui en résultent dans la construction de représentations visuelles. Cette contribution est en réponse au [PROBLEM2.3](#) et est discutée dans Section 1.6.

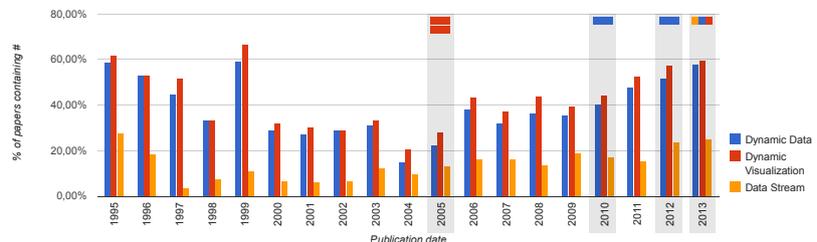
1.6 Outline

Cette thèse est divisé en trois parties. La première partie (I) contient deux chapitres présentant l'état de l'art du domaine, le premier concerne l'informatique et le second contient une courte histoire de l'utilisation de jetons pour faire des représentation visuel. La seconde partie (II) commence par le design et l'exploration de différents système de visualisation de donnée dynamique pour non experts, par le développement d'un nouvelle métaphore de visualisation basé sur les jetons: la sédimentation visuelle. La troisième partie (III) explore, la possibilité d'utilisé une approche basé sur les jetons comme un nouveau paradigme permettant d'autonomiser des utilisateurs non experts dans leur activité de création de leur propre visualisation d'informations.

Part I : *Tavaux précédents*

La première partie de cette dissertation définis le socle de cette recherche en fournissant un aperçu des travaux précédents. Ces travaux sont présentés en deux chapitres séparés. Le premier couvre la littérature qui concerne notre sujet dans le domaine de l'informatique. Le second propose un point de vue historique sur les systèmes de visualisation dynamique qui ont inspiré notre approche de conception.

Figure 1.6: Tracer du nombre d'articles renvoyé par le moteur de recherche de la librairie IEEE pour les requêtes "dynamic data", "dynamic visualization" et "data stream" dans la revue TVCG. Le pourcentage est obtenu par en comparaison du nombre d'article renvoyé avec la requêtes "ieee".



Chapitre 1.6 : *Tavaux précédent en Informatique.*

Ce chapitre synthetise une partie de la littérature en informatique qui concerne les domaines dans lesquels nous allons développer notre recherche: crowdsourcing, visualisation de données dy-

namique, and outils de création de visualization d'information.

Chapitre 1.6 : Une courte histoire des jetons utilisé dans des representation visuelles.



Figure 1.7: La visualisation d'information et d'autres domaines des systèmes de représentation monosémique [Bertin, 1977].

Ce chapitre prend un point de vue historique, explorant l'usage au travers des âges des jetons dans les représentations visuelles. Ce chapitre présente une collection non exhaustive d'artefact qui ont été conçus afin de manipuler des informations abstraites. Chacun de ces systèmes ont été choisis parce qu'ils partageaient un facteur commun: ils sont composés par des éléments visuels discrets, dits jetons.

Partie II : L'analyse de données dynamique pour les non experts.

La deuxième partie de cette thèse décrit le design et l'évaluation de différentes applications conçues et implémentées et déployées pour l'analyse collective de données dynamiques pour des personnes non-expertes en visualisation.

Chapitre 1.6 : *PolemicTweet: Engager un publique d'amateur dans un processus d'analyse de données dynamiques.*



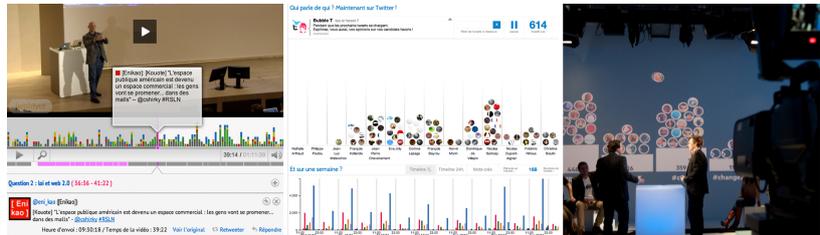
Figure 1.8: Les trois étapes de PolemicTweet: 1) diffusion des informations de connection & tags, 2) Client tweeter temps réels & et annotations structurées, 3) Lecteur vidéo permettant d'accéder à la vidéo par les tweets représenté en petits carré liens: <http://polemictweet.com/rsln/polemicaltimeLine.php>.

Ce chapitre présente le design et l'analyse de PolemicTweet, un système conçu pour crowdsourcer l'annotation et le marquage de vidéos avec des métadonnées structurées représentant des sentiments. Une évaluation de ce système déployé sur une longue période de temps m'a permis d'étudier: (1) si c'était possible d'engager des personnes durant des événements pour conduire des tâches d'analyse dynamique, (2) si ce type de tâches d'analyse peut être utile pour une utilisation future, et enfin (3) est-ce que les données produites sont utiles après l'expérience. Nos recherches confirment

qu'un protocole social bien conçu dédié au contexte et couplé avec une interface spécifique permet de collecter des informations utiles.

Chapitre 1.6 : *Bubble-T & TV, une exploration de l'espace de conception: Combiner un système de jetons avec des simulations physiques & des diagrammes de surfaces*

Figure 1.9: De gauche à droite: PolemicTweet metadata player (Section 1.6), Bubble-T webpage, Bubble-TV photography of the Tv show stage (Section 1.6).

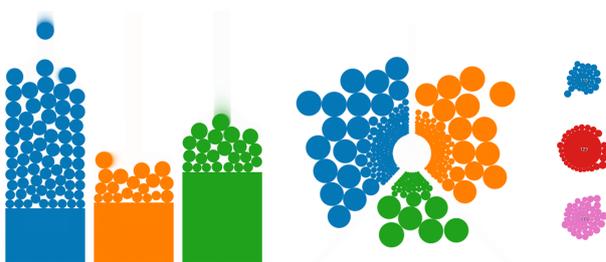


Ce chapitre contient une extension de la recherche menée au Section 1.6. Les considérations et observations faites lors des activités d'annotation et de marquage des vidéos par les utilisateurs non experts décrits dans le chapitre 4 nous a permis de concevoir deux nouveaux systèmes de visualisation basés sur la manipulation de jetons. Ces deux systèmes supportent des protocoles sociaux similaires mais permettent de faire interagir plus d'utilisateurs en même temps.

Bubble-T, est conçu afin de pouvoir représenter chaque instant de la mise à jours de données dynamiques. Il a été déployé lors d'événements bien plus long que ceux de PolemicTweet. Pour Bubble TV, le travail de conception a été orienté sur la possibilité de fournir un système d'annotation similaire à PolemicTweet, mais adapté à un événement avec une plus grande audience (l'audience d'une chaîne de télévision nationale).

Chapitre 1.6 : *Visual Sedimentation: un envoiage visuel générique pour les visualisation dynamique basé sur les jetons*

Figure 1.10: La métaphore de la sédimentation visuelle appliqué à une diagramme à bâtons (gauche), à un diagramme circulaire (centre), et un diagramme en bulle (droite).



Ce chapitre est un point de pivot entre le travail de recherche sur l'analyse social de données dynamique et la recherche sur la conception de représentations visuel par des non experts. A partir

de la recherche menée dans les chapitres précédents (Sections 1.6 and 1.6) nous avons identifié une barrière entre les visualisations discrètes basé sur des jetons et les visualisation continues. Dans ce chapitre nous adressons ce problème en introduisant une nouvelle métaphore qui permet de visualiser plusieurs échelles (discrètes et continues) de flux de données dynamiques. Cette métaphore est inspirée de la sédimentation telle qu'elle se passe dans le monde physique, cela aide la compréhension. Dans ce chapitre je démontre comment la métaphore de la sédimentation permet d'adresser les différents défis des visualisation de données dynamiques: assurer une transition douce et continue entre les données arrivantes et précédentes, ainsi qu'entre les representations par jetons et les représentations basées sur des surfaces. Ensuite l'espace de conception de la métaphore est décrit au travers de l'implémentation d'une boîte à outils qui facilite son exploration. Les possibilités générative de cet espace de design est important, afin de l'illustrer quelques exemples sont présentés.

Part III : *Démocratisation de la création de visualisation dynamique.*

La première partie de cette recherche a été dédiée à la création d'applications fonctionnelles pour différentes tâches d'analyse de données dynamiques, ainsi que le développement d'interfaces dédiées à la visualisation de données dynamiques par des publics de non expert de différentes tailles. Ces interfaces ont permis de démontrer que l'assemblage de jetons dans le temps peut être utile pour la conception de visualisation dynamique. Dans cette partie de la thèse, nous nous basons sur cette observation afin d'instruire: (1) un nouveau paradigme de conception d'outil de création de visualisation pour une population de non expert, et (2) la création d'une étude afin de mieux comprendre ce que cette approche peut nous apprendre comment des personnes non expert en infovis produisent des visualisations.

Chapitre 1.6 : *Définir le paradigme de visualisation constructive.*



Figure 1.11: Un enfant joue avec des jetons en apprenant le mathématique.

Cette phase de recherche commence par extraire les défis les plus significatifs pour démocratiser les outils de création de visualisation d'informations dynamiques. L'examen des possibilités pour adresser ces défis, nous a amené à définir un nouveau paradigme au travers duquel des personnes non expertes en infovis peuvent manipuler des données dynamiques en conjonction avec les variables visuels pour concevoir une association visuelle avec des données. Cela nous a amenée à: (1) porter un nouveau regard sur les théories de l'apprentissage, (2) observer les pratiques dans le domaine non académique où les personnes sont activement engagées dans la construction de leurs propres représentations visuelles, et (3) identifier à partir de ces pratiques les composants et processus de ce nouveau paradigme.

Chapitre 1.6 : *Etudier la construction visuelle ainsi que la construction de correspondance visuelle.*

Figure 1.12: La construction d'une visualisation avec des jetons: la main droite transporte des jetons, la main gauche pointe la correspondance avec des données.



Afin de mieux comprendre le processus d'association des données avec des variables visuelles (visual mapping) ainsi que l'implication du paradigme de *visualization constructive* présenté dans le chapitre précédent, j'ai conçu et conduit une étude exploratoire. Dans cette étude des personnes non expertes de la visualisation d'information sont amenées à créer, mettre à jour, et expliquer leurs propres visualisations d'information en utilisant des matériaux très simples comme des éléments de construction physique. Cela nous a permis d'apprendre que tous les participants, étaient capables de créer, mettre à jour dynamiquement, et expliquer leur propres visualisations. Sur la base de ces observations, je décris les actions des participants dans le contexte du développement de leurs représentations visuelles. À la lumière de ces observations j'en dérive les implications pour la conception d'outils de création de visualisation.

Chapitre 1.6 : *Conclusion and perspectives.*

Ce chapitre contient la conclusion et une discussion des travaux futurs. Ce chapitre résume le processus de recherche, déclare les contributions, et rapportent les limitations de ce travail.

A

Bibliography

James R. Beniger and Dorothy L. Robyn. Quantitative graphics in statistics: A brief history. *The American Statistician*, 32(1):1–11, 1978. URL <http://www.jstor.org/stable/2683467>.

Jacques Bertin. *Semiologie graphique*. Mouton; Paris: Gauthier-Villars, 1973. URL <http://books.google.fr/books?id=F4weAAAAMAAJ>.

Jacques Bertin. *La graphique et le traitement graphique de l'information*. Flammarion, 1977. URL http://www.persee.fr/web/revues/home/prescript/article/colan_0336-1500_1977_num_36_1_1162.

Alan F. Blackwell. The reification of metaphor as a design tool. *ACM TOCHI*, 13(4):490–530, 2006. URL <http://portal.acm.org/citation.cfm?doid=1188816.1188820>.

Stuart K Card, Jock D Mackinlay, and Ben Shneiderman. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999. URL <http://books.google.fr/books/about/?id=wdh2gqWfQmgC>.

J Carroll, W Kellogg, and R Mack. *Interface metaphors and user interface design*. Research reports // IBM, 1987. URL http://books.google.fr/books/about/?id=_xluYgEACAAJ.

John W Creswell. *Qualitative inquiry and research design: Choosing among five approaches*. Sage, 2012. URL <http://books.google.fr/books/about/?id=0JYEbDtKxq8C>.

Nigel Cross. Descriptive models of creative design: application to an example. *Design Studies*, 18(4):427–440, 1997. URL <http://oro.open.ac.uk/39443/8/Descriptive%20Models%20-%20DS.pdf>.

Dedre Gentner. Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2):155 – 170, 1983. ISSN 0364-0213. DOI: 10.1016/S0364-0213(83)80009-3. URL <http://www.sciencedirect.com/science/article/pii/S0364021383800093>.

Sarah Goodwin, Jason Dykes, Sara Jones, Iain Dillingham, Graham Dove, Alison Duffy, Alexander Kachkaev, Aidan Slingsby, and Jo Wood. Creative user-centered visualization design for energy analysts and modelers. *Visualization and Computer Graphics, IEEE*

Transactions on, 19(12):2516–2525, 2013. URL <http://openaccess.city.ac.uk/2618/>.

Saul Greenberg, Sheelagh Carpendale, Nicolai Marquardt, and Bill Buxton. *Sketching user experiences: The workbook*. Access Online via Elsevier, 2011. URL <http://sketchbook.cpsc.ucalgary.ca/>.

Christopher G Healey. Perception in visualization. *IEEE Transactions on Visualization and Computer Graphics*, 10:2008, 2007. URL <http://www.csc.ncsu.edu/faculty/healey/PP/>.

Jeffrey Heer and Ben Shneiderman. Interactive dynamics for visual analysis. *Queue*, 10(2):30:30–30:55, 2012. ISSN 1542-7730. DOI: 10.1145/2133416.2146416. URL <http://doi.acm.org/10.1145/2133416.2146416>.

Scarlett R Herring, Chia-Chen Chang, Jesse Krantzler, and Brian P Bailey. Getting inspired!: understanding how and why examples are used in creative design practice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 87–96. ACM, 2009. URL <http://dl.acm.org/citation.cfm?id=1518717>.

Thomas J Howard, Stephen J Culley, and Elies Dekoninck. Describing the creative design process by the integration of engineering design and cognitive psychology literature. *Design studies*, 29(2):160–180, 2008. URL <http://www.sciencedirect.com/science/article/pii/S0142694X08000173>.

Chris Johnson, Robert Moorhead, Tamara Munzner, Hanspeter Pfister, Penny Rheingans, and Terry S. Yoo. Nih/nsf visualization research challenges report. *LOS ALAMITOS, CA: IEEE COMPUTING SOCIETY*, 2006. URL <http://www.ncbi.nlm.nih.gov/pubmed/16548457>.

Paul Laseau. *Graphic thinking for architects & designers*. John Wiley & Sons, 2001. URL <http://books.google.fr/books/?id=KRCyqDjYSLAC>.

RC MacCormac. Froebel’s kindergarten gifts and the early work of Frank Lloyd Wright. *Environment and Planning B*, 1(1):29–50, 1974. URL <http://www.envplan.com/abstract.cgi?id=b010029>.

Wendy E Mackay and Anne-Laure Fayard. Hci, natural science and design: a framework for triangulation across disciplines. In *Proceedings of the 2nd conference on Designing interactive systems: processes, practices, methods, and techniques*, pages 223–234. ACM, 1997. URL <https://www.lri.fr/~mackay/pdf/FILES/DIS97.Triangulate.pdf>.

E Mcgrath. Methodology matters: Doing research in the behavioral and social sciences. In *Readings in Human-Computer Interaction: Toward the Year 2000 (2nd ed.* Citeseer, 1995. URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.361.1234>.

Robert Moorhead, Chris Johnson, Tamara Munzner, Hanspeter Pfister, Penny Rheingans, and Terry S. Yoo. Visualization research challenges: A report summary. *Computing in Science and Engineering*, 8(4):66–73, 2006. URL <http://www.csee.umbc.edu/~rheingan/pubs/cga06.pdf>.

Tamara Munzner. A nested model for visualization design and validation. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):921–928, 2009. URL <https://www.cs.ubc.ca/labs/imager/tr/2009/NestedModel/NestedModel.pdf>.

O. Neurath. *Modern Man in the Making*. Knopf, 1939. URL <http://books.google.ca/books?id=PxZmAAAAMAAJ>.

Alex F. Osborn. *Applied Imagination, Principles and Procedures of Creative Thinking*. Charles Scribner's Sons, 1953. URL http://www.amazon.com/Imagination-Principles-Procedures-Creative-Thinking/dp/B000H5J5WW/ref=pd_bbs_sr_1/102-8461719-9389763?ie=UTF8&s=books&qid=1181614703&sr=1-1.

Jean Piaget. *Six études de psychologie*. Ed. Denoël, 1989. URL <http://www.fondationjeanpiaget.ch/fjp/site/presentation/index.php?PRESMODE=1&DOCID=1195>.

Stuart Pugh. *Total design: integrated methods for successful product engineering*. Addison-Wesley Wokingham, 1991. URL <http://books.google.fr/books/?id=RKIQQAAMAAJ>.

Philip J Runkel and Joseph Edward MacGrath. *Research on human behavior*. Holt, Rinehart & Winston New York, 1972. URL <http://books.google.fr/books?id=wAZ9AAAAMAAJ>.

Andrew Sears and Julie A Jacko. *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*. CRC press, 2007. URL <http://books.google.fr/books/?id=b3rg6wmjc5QC>.

Michael Sedlmair, Miriah Meyer, and Tamara Munzner. Design study methodology: Reflections from the trenches and the stacks. *Visualization and Computer Graphics, IEEE Transactions on*, 18(12):2431–2440, 2012. URL <http://www.cs.ubc.ca/labs/imager/tr/2012/dsm/>.

Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pages 336–343. IEEE, 1996. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=545307.

Herbert Alexander Simon. *The sciences of the artificial*, volume 136. MIT press, 1969. URL <http://books.google.fr/?id=k5Sr0nFw7psC>.

Robert Spence and A Press. *Information visualization*. {Addison Wesley}, 2000. URL <http://books.google.fr/books/about/?id=gYQoAQAMAAJ>.

Tableau. Tableau Public. , 2014. URL <http://www.tableausoftware.com/public>. small Last access: 23 March 2014.

Anne Treisman. Preattentive processing in vision. *Computer vision, graphics, and image processing*, 31(2):156–177, 1985. URL <http://dl.acm.org/citation.cfm?id=5091>.

Edward R Tufte and PR Graves-Morris. *The visual display of quantitative information*, volume 2. Graphics press Cheshire, CT, 1983. URL http://www.edwardtufte.com/tufte/books_vdqi.

Andrew Vande Moere and Helen Purchase. On the role of design in information visualization. *Information Visualization*, 10(4):356–371, 2011. URL <http://infoscape.org/publications/ivs12.pdf>.

F.B. Viégas, M. Wattenberg, F. van Ham, J. Kriss, and M. McKeon. Manyeyes: a site for visualization at internet scale. *IEEE TVCG*, 13(6):1121–1128, 2007. URL <http://www.bewitched.com/manyeyes.html>.

Colin Ware. *Information visualization: perception for design*. Elsevier, 2004. URL <http://books.google.fr/books?id=qFmS95vf6H8C>.

Leland Wilkinson, D Wills, D Rope, A Norton, and R Dubbs. *The grammar of graphics*. Springer, 2006. URL <http://www.springer.com/statistics/computational+statistics/book/978-0-387-24544-7>.

Wesley Willett, Jeffrey Heer, Joseph Hellerstein, and Maneesh Agrawala. CommentSpace: structured support for collaborative visual analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3131–3140. ACM, 2011. URL <http://vis.stanford.edu/papers/commentSpace>.