

# Méthode de sélection et de configuration des représentations graphiques des corpus d'informations

Cyprien PIERRE

February 3, 2025

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Démarche</b>	<b>3</b>
<b>3</b>	<b>Sémantique</b>	<b>4</b>
<b>4</b>	<b>Eléments généraux</b>	<b>4</b>
4.1	Polices . . . . .	4
4.2	Tables . . . . .	4
<b>5</b>	<b>Graphiques</b>	<b>7</b>
5.1	Séries temporelles . . . . .	7
5.2	Classement . . . . .	7
<b>6</b>	<b>Méthode de sélection</b>	<b>7</b>
<b>7</b>	<b>Etudes futures</b>	<b>7</b>

## 1 Introduction

La visualisation de donnée est une discipline ancienne à la littérature riche et précise. Cependant, il n'existe, à ce jour pas de synthèse opérationnelle ni de méthode harmonisée dans la conception des représentations graphiques des informations. Ce rapport vise à fournir un point de départ cohérent à de futurs travaux en représentation visuelle d'informations. Il est basé sur

une revue de littérature rigoureuse et les recommandations qu'il porte sont issues des lectures réalisées.

L'origine de ces travaux prend place dans un besoin grandissant de solution de visualization d'information dans le secteur de la construction. Les technologies de Big Data, BIM, VDC et les disciplines de la science des données ont fortement pénétré ce secteur sur l'ensemble du cycle de vie des ouvrages. [asiauniversitytaichungtaiwanResearchApplicationFunctiontechnologyaesthetics2022] L'utilisation renforcée de l'IoT, l'émergence des BIS et BOS et l'intégration des solutions d'intelligences artificielles renforcent d'autant plus la nécessité pour ce secteur de se doter de moyens robuste en exploration et en communication des informations.

Le secteur de la construction est historiquement peu digitalisé en comparaison avec les autres industries. Pour adresser ce retard, il est important de publier des études opérationnelle claires et industrialisables dont chaque acteur de l'industrie de la construction puisse bénéficier immédiatement. Ce rapport s'inscrit dans cette démarche.

Il existe de multiples manières d'interagir avec un graphique. [frankelavsky2022, 1]

Ce rapport aborde :

- L'exploration visuelle
- La lecture des tables
- L'écoute des descriptions
- La sonification des données
- L'interactivité avec les commandes de clavier

Ces éléments impliquent la mise en oeuvre de solution programmatique d'accessibilité pouvant intégrer une logique globale de préparation de graphiques numériques.

Ce rapport n'aborde pas :

- L'exploration en réalité mixte
- Les retours haptiques
- L'emploi d'écran à relief (braille)

Ces éléments nécessitant l'emploi de matériels spécifiques.

## 2 Démarche

La conduite de cette revue s’est déroulé de la manière suivante : Nous avons commencé par dresser un panorama des types de graphiques existant en nous appuyant sur les travaux de From Data to Viz [**yanholtzDataViz2018**], The Graphic Continuum [**jonathanschwabishGraphicContinuum2014**], Lexique Visuel [**alansmithLexiqueVisuel**], Insights for ArcGis [**lindabealeInsightsArcGIS2017**]. Cette première action nous a permis de pré-catégoriser les graphiques selon leurs utilisation.

Ensuite, nous avons étudié les livres de divers auteurs et en avons tirés un ensemble de règles générales quant à la préparation des graphiques. Cela concerne les aspects visuels (couleurs, polices. . . ), des conseils sur la pertinence dans le choix des types graphiques suivant les objectifs attendus, quelques éléments d’accessibilités et des retours d’expériences. Sur cette base, nous avons affiné la catégorisation préétablies et commencé à regrouper les recommandations.

Pour compléter les dispositions propres à l’accessibilité, nous avons étudié le rapport Centering accessibility in data visualization [**schwabishCenteringAccessibilityData2022**]. Nous en avons tiré un aperçu global des approches possibles pour améliorer l’accessibilité des graphiques ainsi que de nombreux conseils et bonnes pratiques.

En parallele, nous avons étudié diverses chartes de publications issues de divers médias et institutions. Ces lectures nous ont permis de nourrir les recommandations en matière de conception visuelle des graphiques ainsi que de dresser une liste d’éléments standardisables et d’autres pouvant être laisser libre de personnalisation avec quelques conseils.

Pour finir, nous avons étudié les articles scientifiques publiés sur ces sujets avec une profondeur de recherche à 5 ans. Ce choix de profondeur résulte d’une évolution conséquente depuis le début de la décennie des sciences de l’information. Certains papiers étudiés sont plus anciens mais ont été très régulièrement cités par d’autres publications et nous avons choisis de les considérer. De cette étape, nous avons affiné les recommandations en identifiant les résultats d’expérimentations.

### 3 Sémantique

## 4 Eléments généraux

### 4.1 Polices

Le choix d'une police de texte a de multiples impacts sur la perception des graphiques. Il convient de sélectionner une police accessible et de préférence sans sérif pour un usage informatique (Arial, Calibri, Source Sans Pro, Verdana...).[andreaskrauseBestPracticesData2024] Les polices avec sérif peuvent être considérées pour la production d'éléments imprimés. Il n'y a pas de consensus clair sur l'impact des sérifs sur l'accessibilité des polices.[stephenfew8TableDesign2012] Il n'y a pas non plus de consensus clair sur l'efficacité des polices conçues pour adresser des problématiques d'accessibilités telles que la dyslexie. Il convient d'utiliser une police utilisant une hauteur fixe pour les chiffres.[stephenfew8TableDesign2012]

Il est conseillé de restreindre l'utilisation de l'italique car les textes affichés de la sorte sont plus difficiles à lire. Il est également conseillé de limiter l'utilisation de la graisse et du soulignement à des cas spécifiques pour ne pas surcharger les présentations.

La taille de la police joue un rôle majeur dans l'accessibilité du texte. Il est recommandé d'utiliser une hauteur de police de 12 points.[andreaskrauseBestPracticesData2024] Le nombre de tailles et de type de police doit être limité en nombre.[andreaskrauseBestPracticesData2024]

Certaines polices peuvent être utilisées pour projeter des icônes (NerdFont, StateFace...).[jonathanschwabish10Qualitative2021] L'intérêt des icônes est discuté plus tard dans ce rapport.

Il est important de prévoir le chargement de toutes polices non standard (eg. Source Sans Pro, NerdFont...) dans l'interface utilisateur si celles-ci sont utilisées puisqu'elles ne sont probablement pas installées dans le système d'exploitation de l'utilisateur. Prévoir leur chargement vise à assurer la bonne expérience des utilisateurs.

### 4.2 Tables

Une table ou tableur permet d'exposer des données brutes organisées en lignes ou en colonnes.

Sauf mention contraire, les recommandations sur les tables sont issues du livre *Show me the numbers* de Stephen Few.[stephenfew8TableDesign2012] L'auteur y rentre très en détails sur chaque point de paramétrage. Il y indique notamment les orientations en matière de construction de tableur lorsqu'il s'agit du choix premier d'affichage de données. Ce rapport s'intéresse

à la conception de graphique. Dans ce cadre, les tableurs sont des éléments complémentaires pouvant être affichés par l'utilisateur pour explorer plus précisément les données préalablement affichées.

Pour composer une table de donnée lisible, il est recommandé de :

- Séparer les entrées avec un espace vide,
  - Lorsque les données sont présentées en colonnes, l'espace entre deux colonnes doit être plus grand qu'entre deux lignes,
  - Lorsque les données sont présentées en ligne, l'espace entre deux ligne doit être plus grand qu'entre deux colonnes,
  - Insérer une ligne vide toutes les 5 lignes pour faciliter le balayage visuel **[ISO]**
- Utiliser une ligne horizontale pour séparer les entêtes de colonnes des données,
- Utiliser une ligne horizontale pour séparer les catégories lorsque les données sont triées par catégories,
  - Dans ce cas de figure, ne pas répéter le nom de la catégorie sur toutes les lignes,
  - Les noms des catégories doivent être dans la première colonne, les sous-catégorie dans la seconde colonne, etc. S'il y a plusieurs subdivisions, les noms des catégories doivent être apposées sur la même ligne,
  - Rappeler le nom de la catégorie en cas de changement de page,
  - Maintenir la structure du tableur sur toutes les catégories,
- Rappeler les titres des colones en cas de changement de page,
- Les catégories doivent être ordonnées suivant un ordre logique (eg. chronologique, alphabetique, par classement, etc.)
- Valeurs sommaisées
  - Utiliser une ligne verticale pour séparer les valeurs placées en colonne à droite de toutes les valeurs
  - Utiliser une ligne horizontale si ces valeurs sont placées en une ligne en bas du tableau,

- Si ces valeurs sont le message important de votre tableur, il convient de les affichés immédiatement à droite des colonnes de catégories ou immédiatement en dessous des entêtes de colonnes, suivant la nature du sommaire.
- Ne pas effectuer de rotation sur un tableur, son orientation doit respecter celle du texte du document,
- Les produits d'un calcul doivent être affichés dans la colonne immédiatement à droite de la colonne source de données
- Uniformiser les alignements
  - Les titres des colonnes suivent les alignements des données
  - Aligner les chiffres à droite,
    - \* Homogénéiser les décimales (généralement 2 ou 3 décimales suffisent suivant le contexte),
    - \* Indiquer les valeurs négatives avec le symbole moins (-), ici la notion de valeur symbolique est importante. Certains choisissent d'identifier les valeurs négatives entre parenthèses, cependant il ne s'agit pas d'une représentation naturelle répandue pour une telle identification.
    - \* Séparer les digit d'un nombre par un espace tous les 3 caractères. Dans le cas de grands nombres, il convient de les arrondir à la précision utile (dixaine, centaine...). Par défaut, la précision affichée doit correspondre à celle de la source d'information.
    - \* Si une valeur numérique réfère à une information de catégorie elle doit être traitée comme un texte.
  - Aligner les textes à gauche,
  - Centrer les dates et utiliser une convention stricte d'écriture de ces données telles que YYYY-MM-DD.[ISO] La composition de la date doit être indiqué à l'utilisateur. Le choix du format doit respecter le niveau de précision associée à la mesure. Implicitement, le choix d'un formatage de plus haut niveau que la précision de la mesure induit une agrégation des valeurs.
  - Centrer les données dont la largeur de caractère est fixe

Si des valeurs spécifiques doivent être mises en avant, il est possible d'utiliser l'une ou l'autre de ces solutions :

- mettre le texte en gras,
- remplir la cellule d'une couleur.

Il est recommandé de limiter cette opération à un nombre réduit de valeur. Si cela n'est pas possible, il convient de sélectionner un autre mode de visualisation.

Des prescriptions spécifiques à la préparation de tableurs pour certains graphiques sont apportés le cas échéant dans la suite de ce rapport.

## **5 Graphiques**

### **5.1 Séries temporelles**

La table de donnée sous-jacente doit être organisée par colonne de valeur temporelle croissante.

### **5.2 Classement**

La table de données sous-jacente doit être organisée par ligne de classement croissante (du premier au dernier en partant de la première ligne).

## **6 Méthode de sélection**

## **7 Etudes futures**

Des études complémentaires pourraient être menées sur les sujets suivants :

- La conception de bibliothèques de graphiques accessibles fonctionnant aussi bien dans un usage web que dans les PDF,