

Étude des ouvertures aux échecs

DataViz - Journal d'avancement

Maxime ANTOINE, Salim CHABCHOUB, Mustapha EL ABRIDI,
Jessica NDIAYE, Dorian TONNIS



Visualisation : <https://viz.maxime-antoine.eu/>

Organisation de l'équipe

Pour optimiser le développement, nous avons divisé l'équipe en deux :

- Équipe prétraitement des données : responsable du parsing des fichiers PGN de Lichess, du filtrage des données (Elo, cadence) et de la génération des fichiers JSON optimisés.
- Équipe visualisation et web : Responsable du développement des composants D3.js, de l'intégration de l'échiquier et de la mise en page.

Chronologie du projet

Phase 1 : Brainstorming : 05/11/2025 - 08/12/2025

Pendant cette période, nous avons chacun proposé des idées pour le projet et cherché des jeux de données. Nous nous sommes ensuite mis d'accord sur le thème des échecs et plus particulièrement des ouvertures. Nous avons ensuite réfléchi à plusieurs idées de visualisation avant de réaliser le document de cadrage.

Exemple de questions envisagées par l'équipe lors du brainstorming :

- 1) Quelle est la meilleure ouverture ?
 - 2) Quelle est la méthode la plus utilisée pour mat ?
 - 3) Quels sont les corps les plus gros ?
 - 4) Quelles sont les zones les plus / moins dangereuses du plateau ?
 - 5) Quelles pièces ont le plus de chance de survivre ?
 - 6) Quels points ont le plus de chance d'être promus ?
 - 7) Quelles pièces ont le plus de chance de mourir ?
Est-ce que tout ça dépend de l'elo ?
 - 8) A quel point la cadence de la partie influence le nombre d'erreurs ?
 - 9) Si on fait une erreur, à quel point c'est ratappable ? Est-ce que l'elo ou l'avancement dans la partie ont une influence ?
 - 10) Est-ce que occuper certaines cases lors de l'ouverture augmente la proba de victoires ?
 - 11) Est-ce qu'il y a des zones plus ou moins dangereuses pour le roi en fin de partie ?
 - 12) Combien de temps dure une partie ?
 - 13) Quelles pièces prennent le plus de pièces ?
- dans de la
à l'heure ?

Premières esquisses de visualisations :

Données :

Partie :

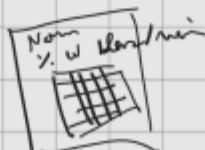
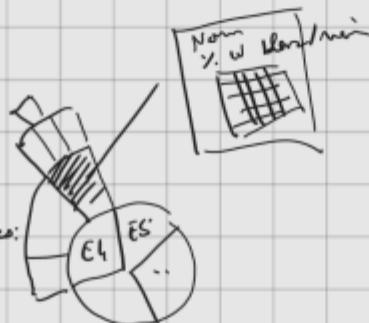
- type de partie
- lieu
- date et heure
- résultat
- elo des joueurs
- gain d'elo
- joueur bâti
- nom de l'ouverture
- raison de fin de partie

Coups :

- liste des coups (PGN)
- évaluation (6 % des parties)
- temps

Brainstorming :

Analyses d'ouvertures:



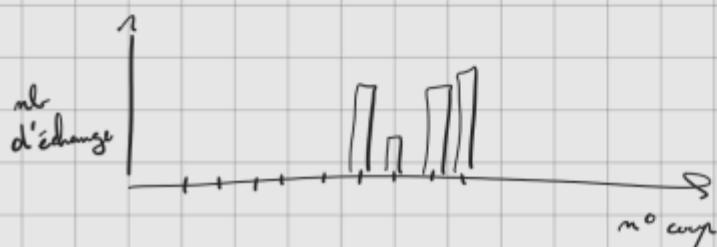
Couleur segment : gagnant

Vision dynamique :

quand on clique ça affiche
la couche suivante

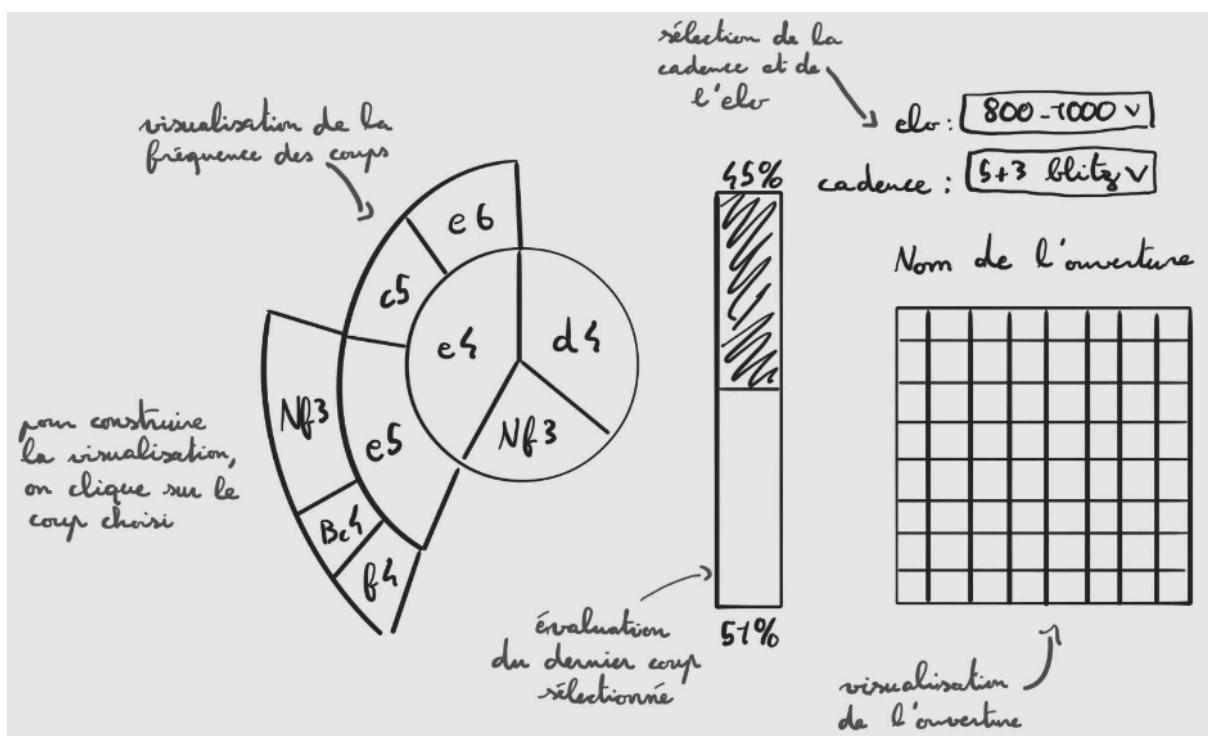
Filtre par elo

Heatmap des coups les plus
joués / où ai les piéces sont
pièces

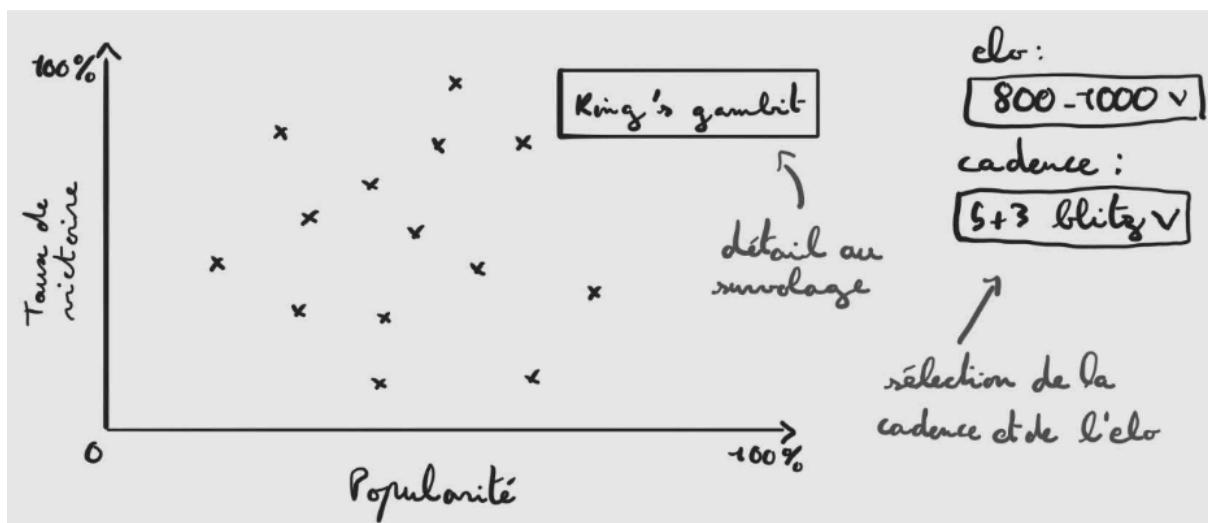


Analysé du nombre de
pièces par avancement de la
partie

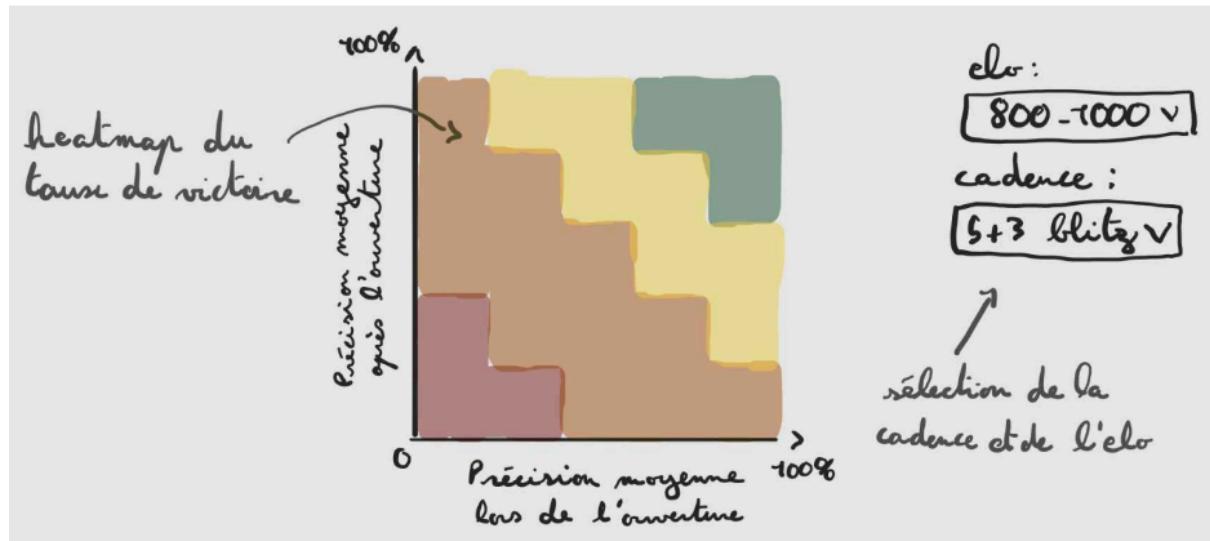
Esquisses finales après brainstorming :



Visualisation n°1 : Visualisation des ouvertures les plus jouées et de leur efficacité par elo et par cadence



Visualisation n°2 : Visualisation de la popularité et du taux de victoire des ouvertures.



Visualisation n°3 : Visualisation du taux de victoire en fonction de la précision lors de l'ouverture et au cours du reste de la partie.

Phase 2 : Traitement des données et fondations et premières visualisations : 09/12/2025 - 04/01/2026

Nous avons commencé le projet en nous mettant d'accord sur un format de données traitées afin que l'équipe prétraitement et l'équipe visualisation puissent travailler en parallèle.

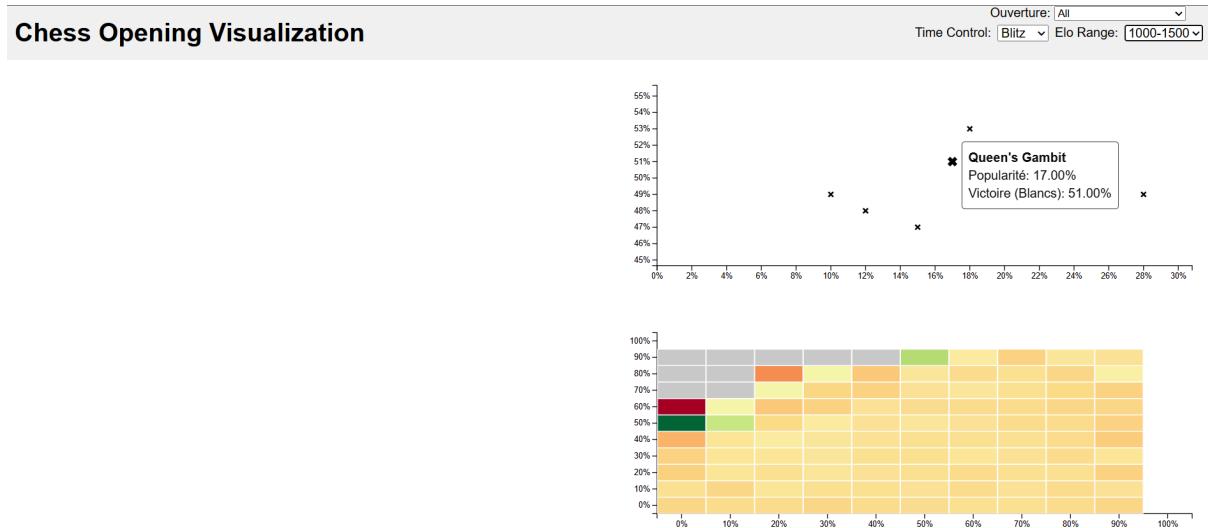
Format JSON convenu :

<code>opening_cadences: { "blitz": {}, "rapid": {}, "bullet": {} "0-500": {}, "500-1000": {}, "1000-1500": {}, "1500-2000": {}, "2000+": {} "blt": {} name: "Polish ...", white_min: 0,45, black_min: 0,5, next_moves: [...] } }</code>	<code>"accuracy": { "blitz": {}, "rapid": {}, "bullet": {} "0-500": {}, "500-1000": {}, "1000-1500": {}, "1500-2000": {}, "2000+": [] accuracy_after_opening opening_accuracy min_rate ...] }</code>	<code>"popularity": { "blitz": {}, "rapid": {}, "bullet": {} "0-500": {}, "500-1000": {}, "1000-1500": {}, "1500-2000": {}, "2000+": [{"name": "Polish ...", popularity: 0,2, min_rate: 0,3 }, ...] }</code>
---	--	--

Ensuite, nous avons choisi de débuter par les visualisations de popularité et de précision. Ces dernières étant plus simples à mettre en œuvre, car elles reposent sur des structures de données linéaires.

Une fois les deux premières visualisations réalisées, nous avons réorganisé le code pour factoriser les comportements similaires dans une classe abstraite. Nous avons également commencé à réfléchir à la propagation des informations entre les graphes en cas d'interactions (changement de filtre, clic sur une visualisation, survole d'un graph, etc.).

Première version des visualisations 2 et 3



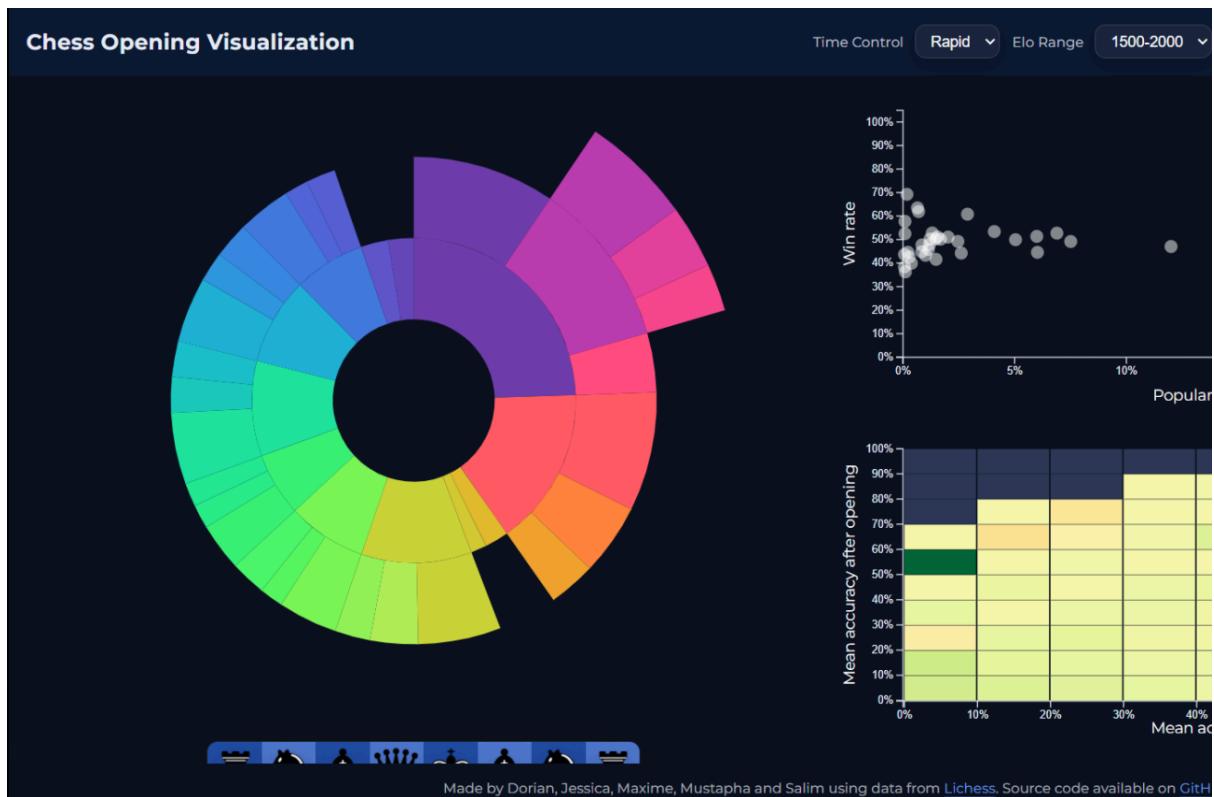
Phase 3 : Interactivité et dernière visualisation : 05/01/2025 - 13/01/2026

Les deux équipes avançant toujours en parallèle, nous avons ajouté la dernière visualisation, l'explorateur d'ouvertures.

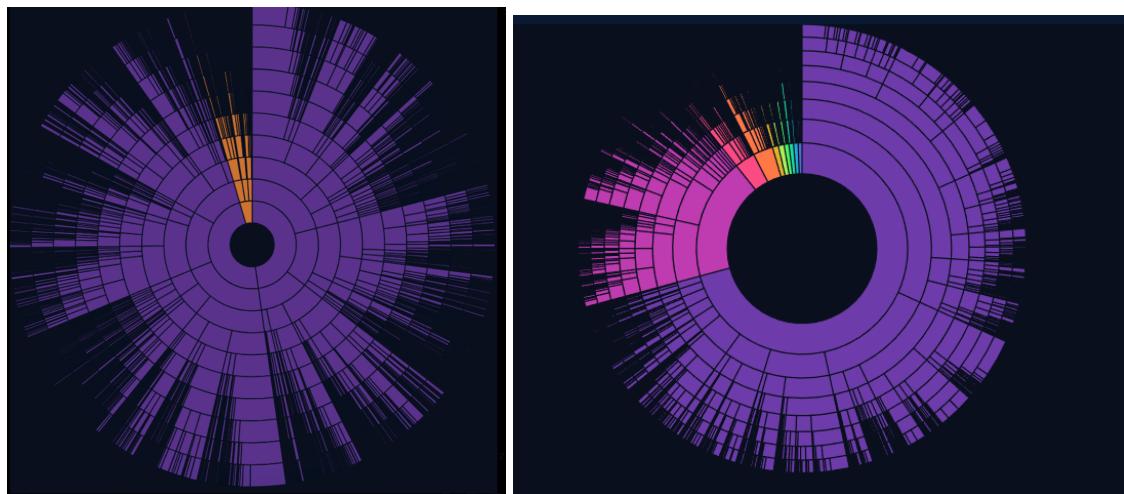
Nous avons ajouté l'échiquier pour visualiser les ouvertures et le diagramme sunburst pour les sélectionner. Cette visualisation a été plus difficile à réaliser, car la structure des données est en arbre ce qui a compliqué sa manipulation et a nécessité de manipuler des fichiers beaucoup plus volumineux.

Nous avons modifié la disposition des composants sur l'écran afin de mieux occuper l'espace disponible. Nous avons aussi ajouté des interactions entre toutes les visualisations. Désormais, sélectionner une ouverture depuis le filtre, l'explorateur d'ouverture, l'échiquier ou la visualisation de la popularité des ouvertures modifie les filtres globaux et met à jour l'ensemble des visualisations.

Première version de la visualisation 1 avec des données placeholder

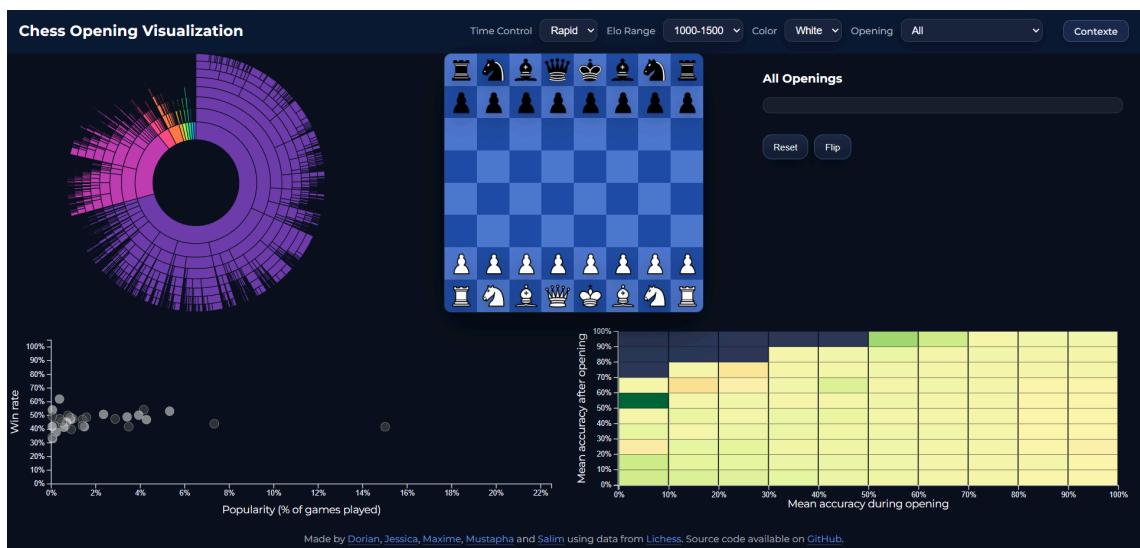


Ajout et ajustement des données lichess:



A gauche l'explorateur d'ouverture ne prenait en compte que 3 coups à chaque couche, mais nous nous sommes rendu compte qu'il était plus judicieux d'ajouter plus de possibilités sur les premières couches, donc nous avons retravaillé les données pour avoir 10 coups sur les deux premières couches et ainsi obtenir la visualisation à droite.

Version terminée pour la soutenance :



Phase 4 : Dernières améliorations : 14/01/2026 - 19/01/2026

Suites aux conseils donnés lors de la soutenance, nous avons mis à jour dynamiquement l'axe Y de la visualisation de la popularité des ouvertures, ajouté une échelle de couleur sur la dernière visualisation et retiré les données non représentatives des visualisation 2 et 3.

Version finale :

