Big Data Chess Project – Planning

# Problem

User Story: To create an application which analysis your chess games, and tells you in which openings you struggle the most or are the strongest.  
  
What we planned so far is having an analysis only on the given player. But if this turns out to be less complicated than we thought it would be, we could extend the analysis.

# Relevant Data

## Chess games:

<https://www.reddit.com/r/chess/comments/9ifkaq/how_i_downloaded_all_my_chesscom_games_using/>  
  
Chess.com and Lichess API implemented in another github: <https://github.com/evamvid/chesscom-games/blob/master/chesscom-games.py>

Lichess lets you download all the games of a given month:   
<https://database.lichess.org/>

They are around 30 gb big.

## Chess Openings (For Clustering)

chessTempo: <https://old.chesstempo.com/chess-openings.html>  
has all chess openings, but I did not find an option to download

Github eco.json : <https://github.com/hayatbiralem/eco.json>  
another Git Project , which calls itself Encyclopedia of Chess Openings. Have to look more into detail into it, but it seems to have all chess openings in json formats.

# Existing Applications

Opening Tree: Free Opensourece software which allows loading in all games from different chess sites, and lets you chose which time formats. Lets you discover the moves that you play the most just in general lets you explore your openings. But does nto show quality of opening only game win percentage at any given position.   
all on git hub as well.

<https://www.openingtree.com> github: <https://github.com/openingtree/openingtree>

Chess com Insights:

Has the insight on percentage of games were you are the person which deviates from the book move

Ein Bild, das Text, Screenshot, Software enthält.

Automatisch generierte Beschreibung

Interesting insight on how well you know the opening theory

# Format of saving lichess games:

PGN:

[Event "Rated Classical game"]

[Site "https://lichess.org/j102bwii"]

[White "Chessyshev"]

[Black "varianteimpasse"]

[Result "1-0"]

[UTCDate "2013.03.31"]

[UTCTime "22:03:44"]

[WhiteElo "1319"]

[BlackElo "1347"]

[WhiteRatingDiff "+14"]

[BlackRatingDiff "-16"]

[ECO "C00"]

[Opening "Rat Defense: Small Center Defense"]

[TimeControl "300+8"]

[Termination "Normal"]

1. e4 e6 2. d4 d6 3. Nf3 Be7 4. Nc3 h6 5. h3 a6 6. Bc4 Nd7 7. O-O b5 8. Bb3 c6 9. d5 cxd5 10. exd5 e5 11. Re1 Bb7 12. a3 Ngf6 13. h4 O-O 14. g3 Ng4 15. Nd2 h5 16. f3 Ngf6 17. Nde4 g6 18. Nxf6+ Bxf6 19. f4 Bg7 20. Ne4 f6 21. Nxd6 Qc7 22. Ne4 Rfe8 23. d6+ Kh8 24. dxc7 exf4 25. Qxd7 Kh7 26. Nxf6+ Kh8 27. Rxe8+ Rxe8 28. Qxe8+ Bf8 29. Qxf8# 1-0

**Suggestion how to save move order: ??**

1. e4 e6 2. d4 d6 3. Nf3 Be7 4. Nc3 h6 5. h3 a6 6. Bc4 Nd7 7. O-O b5 8. Bb3 c6 9. d5 cxd5 10. exd5 e5 11. Re1 Bb7 12. a3 Ngf6 13. h4 O-O 14. g3 Ng4 15. Nd2 h5