Project 3 - Magic: The Gathering

Online Tournament Metagame Analysis  
**Data Engineering Track**

# Team

* Davin Frankosky

# Github

<https://github.com/Koskadelli/Magic-Decklist-Analysis-Tool>

# Overview

Magic: The Gathering is a collectible card game with an estimated 50 million players worldwide. Played both in paper and online, players leverage decks and strategic decision making against one another to claim victory. The game is played competitively in tournaments, where players showcase their deck-building prowess and strategic acumen, leading to a complex and intricate metagame that evolves as new card sets are released. Analyzing the tournament metagame provides insights into the prevailing deck archetypes and strategies that dominate competitive play.

This project seeks to capture the metagame via web-scraping tournament winning decks which have been published online, and using the ETL process to ingest them into a database for further analysis and visualization to identify trends and help make improved decisions on deck choice and card purchases.

# Dataset and Transformation Process

1. Use Python (and specifically Beautiful Soup) to scrape card data from <https://www.mtgo.com/decklists>, (specifically format leagues) for all of January 2024. These are all decks that went 5-0 and will well represent the metagame of specific tournament formats.
2. Clean and re-organize the scraped data to build individual dataframes in Python using Pandas (including information such as format, date of tournament, total number of copies of each card, card name). Structure these in such a way that I can load them into a database by being conscious of the need for unique IDs, etc.
3. Utilize Python-based library MTG SDK (<https://github.com/MagicTheGathering/mtg-sdk-python>), to pull in additional information for cards.
   * This is a new library not covered in class to meet that project criteria.
4. Push this info to a Postgres SQL database for further use.

# Optional Visualizations

1. Flask powered or DB -> JSON -> HTML/JS
2. Select a card and see usage for each tournament of the month by how many copies were played per 5-0 deck on average.
3. Graph the above against changes in card price over time to see if usage correlates to price. (interactive ‘enter a card name’ or dropdown)
   1. Online Pricing information (additional datasource will use: https://www.goatbots.com/download-prices)
4. Colors and card types by volume played. (interactive dropdown of format, pie charts of color and type breakdowns)
5. Prototypes:



