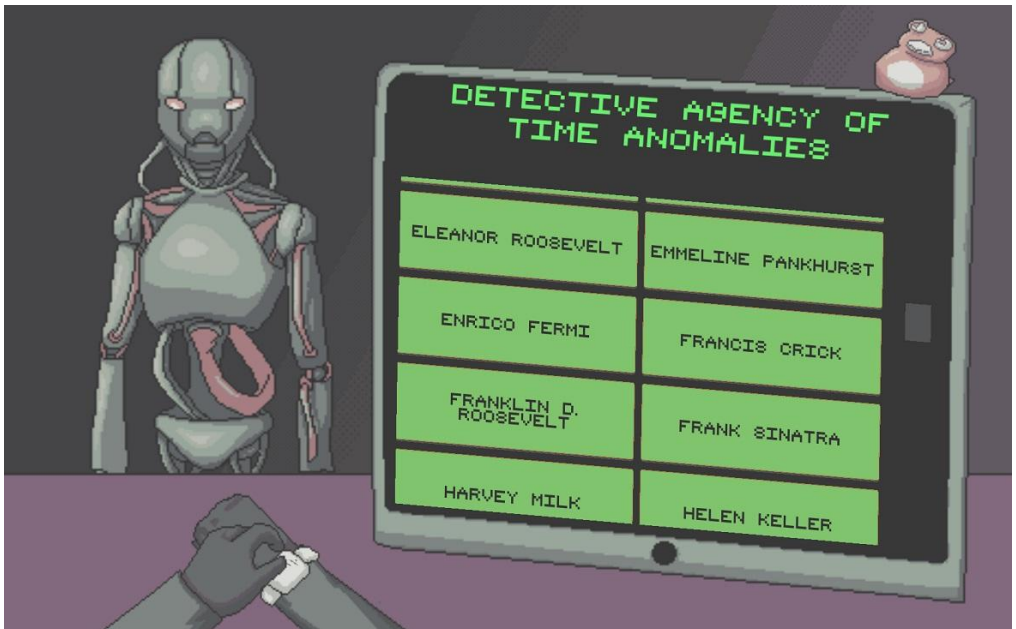


Real-world Data as a Seed

By Antonios Liapis

<http://antoniosliapis.com> | @SentientDesigns

Living in the Information Age, we are surrounded by news, facts, propaganda, and advertisements. We are bombarded with information from TV screens in bars or from wall projections on the subway, and we can look up any type of obscure information on our phones within seconds (thus ruining trivia night). There are community efforts to fact-check and compile information in sites such as Wikipedia, PolitiFact or the Europeana collections. There are also far less noble efforts at using consumers' interaction data to target them with messages of all kinds. While games exist in all types of devices, we do not often think of games as sources of information --- least of all factual. However, games can take advantage of all this information available in repositories, websites and social media to create new ways of engaging players as well as disseminating information during gameplay.





Gabriella Barros, Mike Green, Julian Togelius and I designed a game which takes advantage of information on open data repositories and transforms it into an adventure game. The game is called "DATA Agent" and the player is an agent of the Detective Agency of Time Anomalies (DATA) tasked with solving a bizarre mystery. An assassin has traveled back in time and killed a famous person, masquerading as another famous person somehow related to the victim. Since the assassin does not know everything about the person they impersonate, the DATA Agent must find in a lineup of suspects which one does not have all their facts right. The correct facts about the suspects can be found by talking to other people in different cities. Finding the suspects themselves is no easy task either: the agent must talk to other people, read books and break into dark and locked places.

Most of the game mechanics in *DATA Agent* require the player to interact with real-world data, transformed into game elements such as locations, non-player characters, items, books and facts. The centerpiece is the murder victim, which is chosen by a designer before the game is generated. Around the victim, the generator starts by finding possible suspects among people with Wikipedia articles that share as many common attributes with the victim as possible (e.g. the same birth date or the same thesis advisor) but also have different attributes with other suspects. Once suspects are found, one of them is randomly chosen to be the culprit (the time-traveling doppelganger) and one of their facts is changed. Each suspect is linked back to the victim through a chain of entities in the Wikipedia knowledge ontology (essentially, finding the degrees of separation in Wikipedia:

https://en.wikipedia.org/wiki/Wikipedia:Six_degrees_of_Wikipedia); those links are transformed into in-game characters or books, placed in locations around the world based on their origin (e.g. their birthplace) or places found in the chain. Dialog with characters is generated based on templates to point to the next clue (unlocking a new character, item and/or location) but can also provide some



information about the character (such as their date of birth or subject). Finally, some puzzle elements are added by ``locking'' some locations and ensuring that ``keys'' can be found by the player (torches to unlock dark places, crowbars to unlock chained places).



As expected, adventures generated by the DATA Agent algorithms do not always present a challenge --- or make sense. Suspects are usually connected to the victim, for example the suspects for the murder of Albert Einstein are all physicists. On the other hand, some of the falsehoods used to pinpoint the culprit can be obvious from common knowledge without playing the game. Some of the chains leading the player from the victim to the suspects also take bizarre



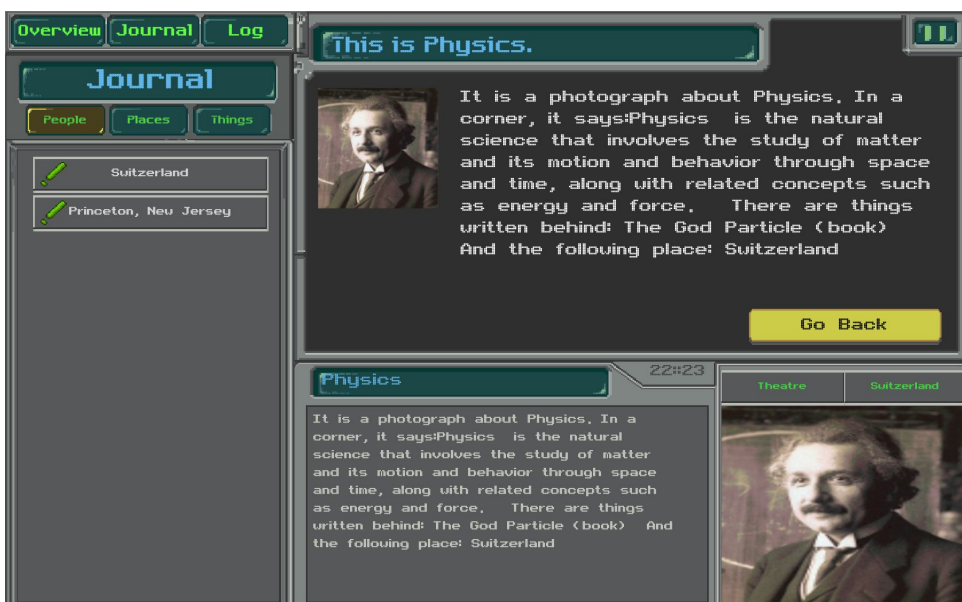
paths, for instance talking to Caligula and Khrushchev to find where the suspects of Frank Sinatra's murder live. Errors in the actual Wikipedia knowledge-base (DBpedia) can also lead to absurd results such as a non-player character named Argentina in the case of Louis Armstrong's murder (Argentina is of type ``person" in the DBpedia ontology, leading to this error). You can find out more bizarre outcomes, and play 99 generated adventure games by downloading *DATA Agent* from <https://champchampchamp.itch.io/data-agent>.



DATA Agent is far from perfect, but its aspirations are worth examining. Using real-world knowledge through open data repositories allows games to be more relevant and closer to the real world, recent news, or trending web searches. Trying to show the ``degrees of separation" in Wikipedia articles through a playful environment and a murder mystery narrative



(however absurd that is) requires both the generator and the player to rationalize why these links exist. *DATA Agent* transforms information of real-world people for the game, changing their time/place of death (for the victim) or blaming them for a murder they did not commit (the culprit); it explicitly states that history has been changed by a murder and a time-traveling doppelganger masquerading as a real person. However, in different scenarios (beyond murder mysteries, most likely) the real-world information could be kept intact within the game to allow players to explore real-world information, even for learning purposes. Using other sources of contemporary or opinion-laden snippets such as trending Twitter topics or activist websites can result in games rich in critique and conflicting viewpoints (but possibly sparse on facts, and thus unfit for learning). Games that highlight and expose data can even be used to debug or fact-check the original repositories, closing the feedback loop by correcting the knowledge-base that game assets were built from.



In short, *DATA Agent* is only one instance of how a semantically rich and narrative-heavy game built on real-world data can be used for entertainment, data visualization or critique.

Relevant readings:

Michael Cerny Green, Gabriella A. B. Barros, Antonios Liapis and Julian Togelius: "DATA Agent" in Proceedings of the 13th Conference on the Foundations of Digital Games, 2018.

Gabriella A. B. Barros, Michael Cerny Green, Antonios Liapis and Julian Togelius: "Data-driven Design: A Case for Maximalist Game Design," in Proceedings of the International Conference of Computational Creativity, 2018.

Website:

<https://champchampchamp.itch.io/data-agent>

