Building a data marketplace

Introduction

According to the World Economic Forum[a], the Economist[b], and others data has been characterised as the oil of our modern economy while politicians started noting that people should get a cheque for their data from companies like Facebook, Amazon, Google.[c] Facebook, Amazon, Google and other corporations in the list of the most valuable firms in the world have access to personal information of billions of people in an international level and a big fraction of their revenue comes from using this information.[d] Additionally, many companies try to augment datasets they have collected from their customers with external data, and for that, they transact with data marketplaces.

The popularity of data marketplaces can be easily justified by the continuously increasing need for data for training systems based on artificial intelligence.[e] For example, data from multiple electricity smart metres can be used to provide feedback to households and assist in reducing their electricity bills or even detect malfunctioning appliances. Data from Internet of Things (IoT) devices can speed up the development of personalised models that improve users' quality of experience. By collecting multiple types of information, online service providers can offer more sophisticated services that provide a better quality of experience to the users while also boosting their revenue. Unfortunately, users who generate this information do not get financial rewards. This is a consequence of an existing "understanding" between people and corporations that people receive unpaid access to online services in exchange for unpaid access to their personal data.

In this project, you will work in teams to develop your own data marketplace. Detailed requirements for the functionality of your marketplace and submission deadlines are provided below.

[a] reports.weforum.org/rethinking-personal-data/executive-summary
[b] www.economist.com/leaders/2017/05/06/ the-worlds-most-valuable-resource-is-no-longer-oil-but-data
[c] https://www.cnbc.com/2019/10/17/andrew-yang-facebook-amazon-google-should-pay-for-users-data.html
[d] worldin2019.economist.com/killerrobots
[e] Posner, Eric A., and E. Glen Weyl. Radical markets: Uprooting capitalism and democracy for a just society. Princeton University Press, 2018.

Short description

Team registration deadline: Sunday 20 February @ 23:59
Progress report deadline: Wednesday 2 March @ 23:59
Project submission deadline: Saturday 12 March @ 23:59
Total credits: 60% | Credits for team registration: 1% | Credits for progress report: 3% | Credits for providing additional functionalities: 9%
Submission: Via Brightspace. Different assignments will be created for each deadline. Use the following naming convention for each of your submission `{student_ids}__{registration | progress | final}.zip`. For example, the final report of a team with four students with ids 1234, 5678, 9012, and 3456 should be named `1234_5678_9012_3456__final.zip`
Send an email to `dimitris.chatzopoulos@ucd.ie` for any questions.

Detailed description

In this project you will work in teams of 3 or 4 and you will implement a data marketplace. Single person teams will be accepted under special conditions. The minimum functionalities that should be supported by your marketplace are a catalogue, a shopping cart, a checkout, and a (fake) purchasing mechanism (e.g., PayPal).

(i) **Catalogue.** The catalogue should include pages for browsing list of datasets and a page for displaying details of an individual dataset. The dataset list page should present all the available datasets to the customer. For each dataset, there should be a name, and a price per datapoint. You should be able to click on a dataset to go to its individual dataset page. The individual dataset page should include, as a minimum, a name, a description, a price per datapoint, the number of the available datapoints, and the minimum value and the maximum value. You should also have a purchase button. *You are welcome to add more information.*

(ii) **Shopping cart.** The shopping cart should display a list of the datasets from which you intend to buy a sample of datapoints. It should allow you to change the number of datapoints or remove datasets from the list. It should provide the price of each dataset part (e.g., if a customer wants to buy 100 datapoints from a dataset that has 10000 datapoints and the price per datapoint is 0.01$, the written price should be 1$) and give the total price to be paid. There should be a checkout button to take you to the checkout and purchasing part of the system. If a customer selects to buy n datapoints from a dataset that has N datapoints, each datapoint in the dataset will be selected with $1/N$ probability. *You are welcome to add more datapoint selection approaches (e.g., by selecting equally spread in the dataset datapoints)*.

(iii) **Checkout.** The checkout should capture key customer information and request payment details. You can pretend that every user enters valid details for now. Completion of payment should create an order that is made visible to the marketplace owner. The fake payment system should generate some form of payment id that should be associated with the order. The order state should be modifiable to indicate that the order is new, has been cancelled or has been delivered.

(iv) **Customer.** Customers must sign up to purchase data so you don't need to implement guest payment (i.e. if you don't have and account or are not logged in, you cannot buy data). A customer should be able to log on to the system and see a history of their previous purchases. When viewing the available datasets, a customer can select how many of the datapoints (s)he is interested in buying.

(v) **Marketplace owner.** An administrator login should be provided that allows the marketplace owner to create new products, review orders, and change the state of the order. The marketplace owner should be able to change the dataset details (but not the name). It is essential that, when a dataset price per datapoint is changed, the historical prices on existing orders does not change, but that they reflect the price actually paid. The marketplace owner should be able to hide, but not delete existing datasets (all datasets must be kept so that they can be matched to items in orders).

	Requirement ID & Description	Credit
general	G_0 : View the datasets currently on offer.	2%
	G_1 : View details about a particular dataset.	2%
	G_2 : Add a set of datapoints to a shopping cart.	2%
	G_3 : Change the number of the selected datapoints in the shopping cart.	3%
	G_4 : Remove a selected dataset from the shopping cart.	2%
	G_5 : Create a customer account.	2%
	G_6 : Log in to their account.	2%
	G_7 : Something extra	3%
Customer	C_0 : Go to the checkout to purchase orders in their shopping cart.	2%
	C_1 : Use a (fake) payment portal to pay for their goods.	2%
	C_2 : View their order history.	3%
	C_3 : Something extra	3%
Owner	O_0 : Login to the marketplace.	1%
	O_1 : Add datasets to the marketplace.	3%
	O_2 : Hide datasets to make them no longer available.	3%
	O_3 : View all orders.	3%
	O_4 : Change the state of orders.	1%
	O_5 : Edit dataset details.	3%
	O_6 : Something extra	3%
Documentation	D_0 : Team registration	1%
	D_1 : Progress report	3%
	D_2 : Video Demo	5%
	D_3 : Wireframe design of the project	2%
	D_4 : Final report	4%

Table 1: Table of requirements.

Documentation guidelines

(i) **Team registration.** You need to submit an HTML page that will contain: *(i)* team's name, *(ii)* the name, the email and the student ID of each team member, and *(iii)* a photo of each of the team member or a photo with all the team members. Create a zip file with the HTML file and supplementary material and submit it via Brightspace. Please note that the deadline for this is Sunday 20 February @ 23:59.

(ii) **Progress report.** Update the HTML page you submitted for the team registration by adding the following information: *(i)* the team member who is responsible for each requirement. More than one team members can work on one requirement but one has to be responsible for its delivery. *(ii)* Mention some potential extras you are thinking to implement in the project in order to get the extra credits (this is not a commitment). *(iii)* Add a table with ten 30-minute slots the whole team is available between Friday 4 March and Tuesday 8 March to have a Zoom meeting with Dimitris Chatzopoulos and discuss the progress of your project and get feedback. Submit the zip file through Brightspace. Please note that the deadline for this is Wednesday 2 March @ 23:59.

(iii) **Video Demo.** Demonstrate your marketplace while recording your screen. You can use Zoom to do that but you are free to you any other tool you want. The video format should be mp4.

(iv) **Wireframe design of the project.** A wireframe design of your project will help you with the development of your project as well. So, do not consider it as an additional task that gives you extra credit but use this design to decide how you will develop the project.

(v) **Final report.** Update the HTML page you submitted for the intermediate report to produce the final report. The final report should include *(i)* a table with all the requirements you think you have completed and what additional requirements you added to get the extra credit. This table should include a column that will contain the member of the team who was responsible for each of the requirements and any other members that contributed. Additionally, the final report should contain *(ii)* the wireframe design, *(iii)* instructions on how to install and run your project, and *(iv)* reflections on what you have learnt through the project and what you would do differently.