

Εργασία στο Μάθημα “Τεχνολογίες Υπηρεσιών Λογισμικού”

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Εθνικό Μετσόβιο Πολυτεχνείο

Εαρινό εξάμηνο 2021-2022

Διδάσκοντες: Β. Βεσκούκης, Π. Τσανάκας

Περιγραφή

Η εργασία περιλαμβάνει την πλήρη ανάπτυξη μιας μικρής δικτυακής εφαρμογής λογισμικού που θα παρέχεται ως υπηρεσία (SaaS), οι προδιαγραφές της οποίας δίνονται από τους διδάσκοντες. Ο κύκλος της ανάπτυξης δεν περιλαμβάνει διερεύνηση απαιτήσεων, ανάλυση επιχειρησιακών διαδικασιών του πελάτη κ.λπ., αλλά επιχειρεί να προσομοιώσει συνθήκες όπου, στην πορεία της ανάπτυξης μιας εφαρμογής, εξειδικεύονται προδιαγραφές οι οποίες είναι σε ένα βαθμό γνωστές κατά την έναρξη της ανάπτυξης. Η εστίαση επικεντρώνεται στην αρχιτεκτονική, τις τεχνολογίες και τα εργαλεία ανάπτυξης εφαρμογών που παρέχονται ως υπηρεσίες καθώς και στην εφαρμογή ευέλικτων (agile) μεθοδολογιών ανάπτυξης, επίσης με τα αντίστοιχα σύγχρονα εργαλεία.

Το “EnergyLive2022” είναι μια δικτυακή εφαρμογή - υπηρεσία που προσφέρει τη δυνατότητα οι χρήστες να παρακολουθούν τιμές μεγεθών της αγοράς ηλεκτρικής ενέργειας στην Ευρώπη μέσω του φυλλομετρητή τους. Τα δεδομένα, τα οποία διατίθενται ως ανοικτά δεδομένα, θα συλλέγονται αυτόματα από την εφαρμογή, με σύνδεση μέσω ftp ή API και θα εμφανίζονται στην ιστοσελίδα της εφαρμογής, στην οποία συνδέονται οι χρήστες για να λάβουν την υπηρεσία. Με τον τρόπο αυτό, οι χρήστες δεν έχουν την επιμέλεια να συλλέξουν και να ομογενοποιήσουν δομικά και εννοιολογικά μεγάλο όγκο δεδομένων, υπηρεσία για την οποία δέχονται να πληρώσουν. Οι ελάχιστες υποστηριζόμενες λειτουργίες είναι οι ακόλουθες:

- Επιλογή μεγέθους της αγοράς ενέργειας (φορτίο, παραγωγή, εισαγωγές-εξαγωγές), χώρας και, κατά περίπτωση, άλλων παραμέτρων, καθώς και ημερομηνίας έναρξης της απεικόνισης
- Απεικόνιση των τιμών του μεγέθους με μορφή διαγράμματος από μια καθορισμένη ημερομηνία έναρξης, μέχρι την τρέχουσα στιγμή, με αυτόματη ενημέρωση όταν προστίθενται νέα δεδομένα
- Κατέβασμα των δεδομένων που περιέχονται στο διάγραμμα σε μορφή csv, json ή σε μορφή εικόνας
- Εμφάνιση των υπολειπόμενων ημερών συνδρομής και δυνατότητα ανανέωσης
- Sign in μέσω της υπηρεσίας "Sign in with Google"

Προαιρετικά (υποχρεωτικά για πολυπληθείς ομάδες), θα υλοποιηθούν επιπλέον λειτουργίες:

- Απεικόνιση περισσότερων του ενός διαγραμμάτων στην οθόνη
- Απεικόνιση σε χάρτη της Ευρώπης (θα σας διατεθεί αρχείο με χώρες)

Οι διδάσκοντες, σε λόγο υποθετικού πελάτη, διαθέτουν μια πρώτη περιγραφή του τρόπου που φαντάζονται την αλληλεπίδραση των χρηστών με την εφαρμογή, ο οποίος δεν είναι πλήρης ούτε τελικός. Τα περιεχόμενα της περιγραφής θα εξειδικευτούν ή θα χρειαστεί να ερμηνευτούν από κάθε ομάδα στην πορεία της εργασίας, η οποία θα αναπτυχθεί με agile μεθοδολογία. Θα διατεθεί αρχείο

Visual Paradigm με την προδιαγραφή του επιθυμητού τρόπου αλληλεπίδρασης του χρήστη με την εφαρμογή, ο οποίος θα χρησιμοποιηθεί ως αναφορά, δηλαδή δεν είναι υποχρεωτικό να υλοποιηθεί ακριβώς με αυτό τον τρόπο.

Πηγές δεδομένων

Τα δεδομένα που θα απεικονίζονται διατίθενται ως ανοικτά δεδομένα από τον οργανισμό entso-e (European Network of Transmission System Operators for Electricity: www.entsoe.eu) μέσω sftp, σύμφωνα με τις οδηγίες που υπάρχουν εδώ: https://transparency.entsoe.eu/content/static_content/Static%20content/knowledge%20base/SFTP-Transparency_Docs.html. Από τα δεδομένα αυτά θα χειριστούμε 3 επιλεγμένα datasets όπως φαίνεται στον ακόλουθο πίνακα.

Πίνακας 1. Επιλεγμένα datasets

Dataset	Περιγραφή δεδομένων	Αρχείο στο ftp
Actual Total Load	Πραγματικό φορτίο ηλεκτρικής ενέργειας	YYYY_MM_ActualTotalLoad_6.1.A.csv
Aggregated Generation per Type	Συνολική παραγωγή ηλεκτρικής ενέργειας ανά τύπο	YYYY_MM_AggregatedGenerationPerType_16.1.B_C.csv
Physical flows	Διασυνοριακές ροές ηλεκτρικής ενέργειας	YYYY_MM_PhysicalFlows_12.1.G.csv

Κάθε τέτοιο σύνολο περιέχει κάποιες στήλες (πεδία) που χαρακτηρίζουν τα δεδομένα (χώρα, τύπος παραγωγής, κ.ά.), καθώς και αυτό καθεαυτό το "φορτίο" της χρονοσειράς των τιμών του εκάστοτε μεγέθους (data load), όπως θα εξηγηθεί αναλυτικά στο μάθημα. Τα πεδία χαρακτηρισμού των δεδομένων (χώρα, τύπος παραγωγής κ.λπ.) αναφέρονται και ως "διαστάσεις". Κάθε χρονοσειρά αποτελείται από χρονοσημάνσεις (timestamps) και τιμές δεδομένων. Ο αριθμός των τιμών της χρονοσειράς που αντιστοιχούν σε μία ημέρα, καθορίζεται από τη χρονική ανάλυση της χρονοσειράς και είναι 24, 48 ή 96, για δεδομένα ανά ώρα, μισάωρο, ή τέταρτο, αντίστοιχα.

Τα δεδομένα διατίθενται από τον entsoe σε αρχεία csv (tab delimited), ένα αρχείο για κάθε μήνα, όπως φαίνεται στη δεξιά στήλη του πίνακα 1. Κάθε μέρα του τρέχοντος μήνα προστίθενται στο αρχείο νέες εγγραφές. Σημειώνεται ότι είναι πιθανό να υπάρχουν κενά (πχ κάποιες χώρες οι οποίες για κάποια ημέρα δεν έχουν δεδομένα) τα οποία μπορεί να συμπληρωθούν στο μέλλον. Επίσης, σε επόμενη έκδοση του αρχείου csv οι τιμές σε κάποια μη-κενά δεδομένα μπορεί να διορθωθούν. Αμφότερες οι συμπεριφορές αυτές (συμπλήρωση, διόρθωση) επεξηγούνται στην ενότητα "Γενική δομή και κύκλος ενημέρωσης δεδομένων" που ακολουθεί.

Για τους σκοπούς της εργασίας, θα σας διατεθούν τα παραπάνω αρχεία csv με δομή ακριβώς ίδια με αυτή του entsoe, όμως με μια χρήσιμη για την εργασία διαφορά: τα αρχεία θα διατίθενται με ωριαία αντί ημερήσια ενημέρωση, δηλαδή το όνομά τους ακολουθεί τη λογική π.χ. YYYY_MM_DD_HH_ActualTotalLoad_6.1.A.csv. Με τον τρόπο αυτό μπορεί να φαίνεται ευκολότερα κατά τη λειτουργία της εφαρμογής σας η "ζωντανή" αυτόματη ανανέωση των διαγραμμάτων που απαιτεί η εργασία. Αυτό δεν απέχει πολύ από την πραγματικότητα, καθώς κάποια δεδομένα διατίθενται μέσω API και ενημερώνονται ανά ώρα ή και συχνότερα.

Στον πίνακα 2 που ακολουθεί φαίνεται πώς ενημερώνεται ανά ώρα το αρχείο που περιέχει το dataset "Actual Total Load". Προσέξτε ότι το αρχείο "μεγαλώνει" κάθε φορά κατά μία ώρα, και περιέχει τα δεδομένα από την αρχή του μήνα και όχι μόνο αυτά που έχουν προστεθεί την τελευταία ώρα.

Πίνακας 2. Διάθεση αρχείων για την εργασία

Αρχείο entsoe	Αρχείο εργασίας	Ενημέρωση από	Ενημέρωση έως
2022_03_ ActualTotalLoad_6.1.A.csv	2022_03_01_01_ ActualTotalLoad_6.1.A.csv	1.3.2022 00:00	1.3.2022 01:00
	2022_03_01_02_ ActualTotalLoad_6.1.A.csv	1.3.2022 00:00	1.3.2022 02:00
	2022_03_01_03_ ActualTotalLoad_6.1.A.csv	1.3.2022 00:00	1.3.2022 03:00
	...	1.3.2022 00:00	...
	2022_03_02_00_ ActualTotalLoad_6.1.A.csv	1.3.2022 00:00	2.3.2022 00:00
	...	1.3.2022 00:00	...
	2022_03_15_05_ ActualTotalLoad_6.1.A.csv	1.3.2022 00:00	15.3.2022 05:00
	...	1.3.2022 00:00	...

Θα σας διατεθεί το σύνολο των αρχείων των τριών προαναφερθέντων datasets από την αρχή του χρόνου, δηλαδή για κάθε dataset 31*24 αρχεία για τον Ιανουάριο, 28*24 αρχεία για τον Φεβρουάριο και 8*24 αρχεία μέχρι περίπου και την 8.3.2022 (σύνολο περίπου 1600 αρχεία για κάθε ένα από τα 3 dataset, με μέγεθος αρκετών GB).

Κάθε εβδομάδα, θα αφιερώνονται 10-15' από το μάθημα για συζήτηση εφ' όλης της ύλης της εργασίας.

Ομάδες, εργαλεία

Η εργασία θα γίνει από ομάδες των 4 ατόμων, οι οποίες καλούνται να υλοποιήσουν **το σύνολο των ελάχιστων υποχρεωτικών λειτουργιών** του EnergyLive2022 που αναφέρονται παραπάνω, ως δικτυακή εφαρμογή η οποία θα γίνει deploy σε παραγωγικό περιβάλλον που θα ανακοινωθεί. Η αρχιτεκτονική που θα υλοποιήσετε θα είναι η αρχιτεκτονική **microservices**.

Ομάδες των 5 ατόμων θα υλοποιήσουν μία από τις "επιπλέον λειτουργίες" της επιλογής τους, ενώ ομάδες των 6 ατόμων θα υλοποιήσουν και τις δύο επιπλέον λειτουργίες. Ομάδες με λιγότερα από 4 άτομα θα ασχοληθούν μόνο με τα δύο πρώτα datasets και για όλες τις υποχρεωτικές λειτουργίες.

Τα εργαλεία που θα χρησιμοποιηθούν έχουν ως ακολούθως:

- Σχεδίαση, αρχιτεκτονική: Visual Paradigm CE
- Υλοποίηση: WebStorm, node, express, Bootstrap, highcharts ή άλλο αντίστοιχο JS περιβάλλον, καθώς και Python. DBMS της επιλογής σας.
- Διαχείριση και σχολιασμός πηγαίου κώδικα: github

- Διαχείριση έργου Agile: YouTrack
- Deployment: (θα ανακοινωθεί)

Με την οριστικοποίηση των ομάδων θα διατεθούν από τους διδάσκοντες τα αντίστοιχα YouTrack project και github repositories.

Παραδοτέα, βαρύτητες

Τα παραδοτέα της εργασίας έχουν ως εξής:

1. **Αρχείο vnrp με την αρχιτεκτονική** της υλοποίησής σας σε μορφή διαγραμμάτων UML component, deployment και sequence, καθώς και σχεδίαση της υποδομής δεδομένων (ER ή NoSQL json τεκμηρίωση). Σε καθένα από τα components που θα περιέχεται στα διαγράμματα, θα πρέπει να υπάρχει σύνδεσμος στο αντίστοιχο αρχείο πηγαίου κώδικα στο github. (25%)
2. **Github repository με τον πηγαίο κώδικα**, συμπεριλαμβανομένων deployment scripts (45%)
3. **Η εφαρμογή σε λειτουργία**, στην υποδομή cloud που θα ανακοινωθεί (10%)
4. **Stress tests** με το εργαλείο jMeter (10%)
5. Ενημερωμένος δικτυακός τόπος της υπηρεσίας **YouTrack**, με την πορεία διαχείρισης του έργου με 3-4 sprints, καθώς και **Σχολιασμός κώδικα** στο GitHub (20%)

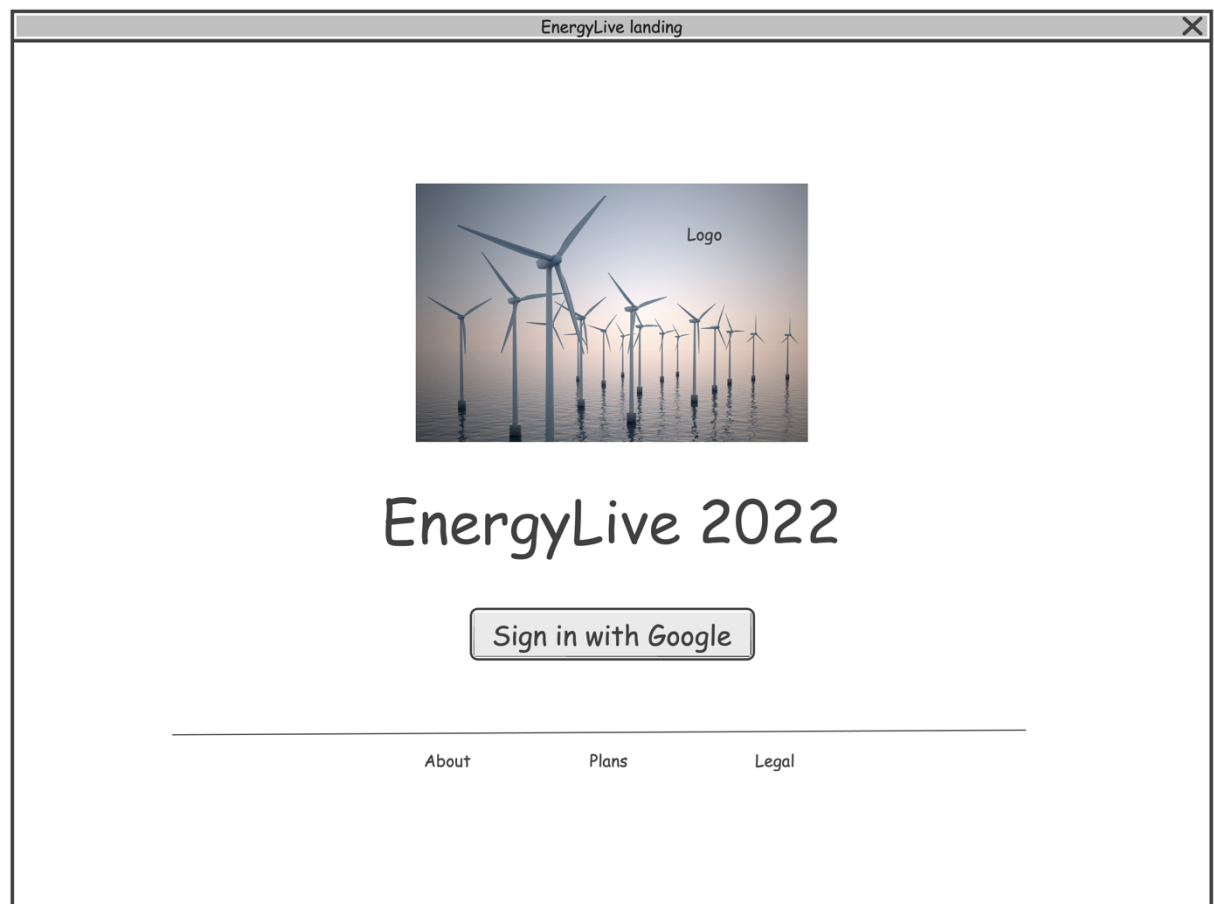
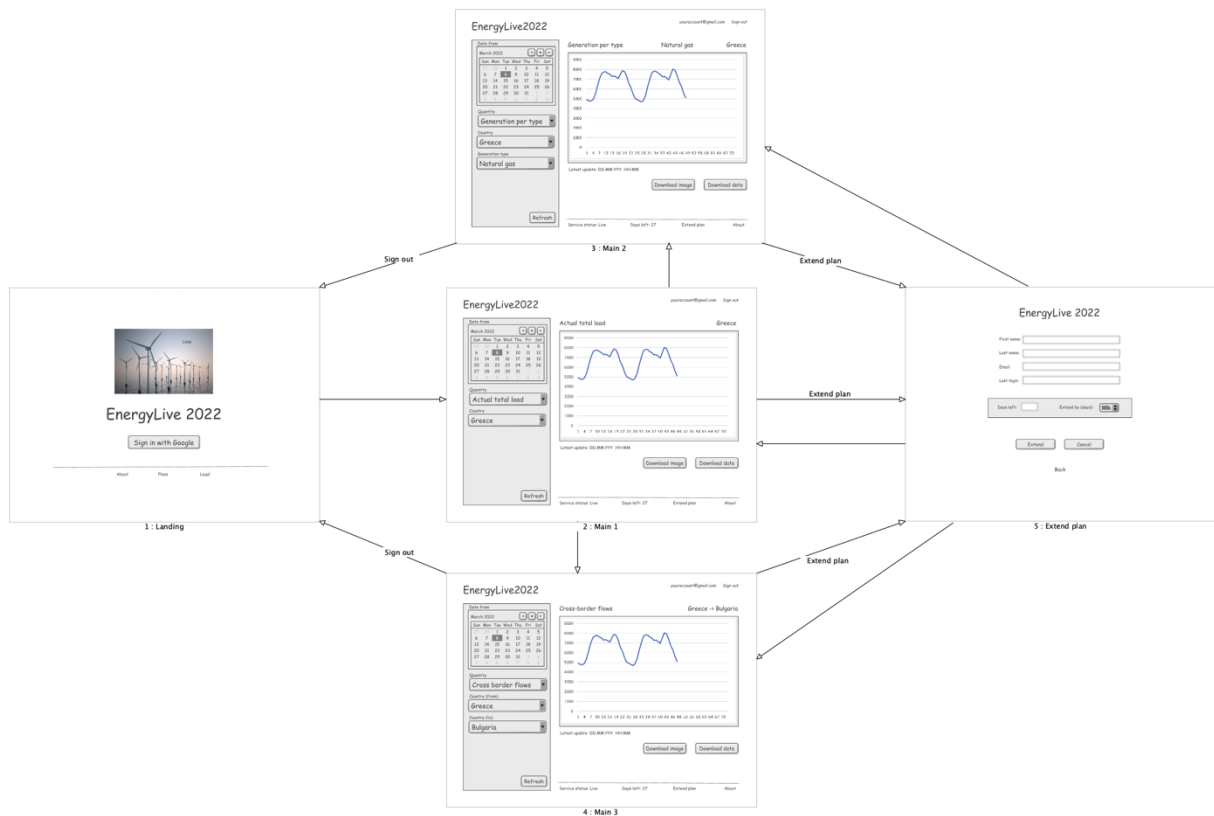
Χρονοδιάγραμμα, ενδιάμεσα παραδοτέα

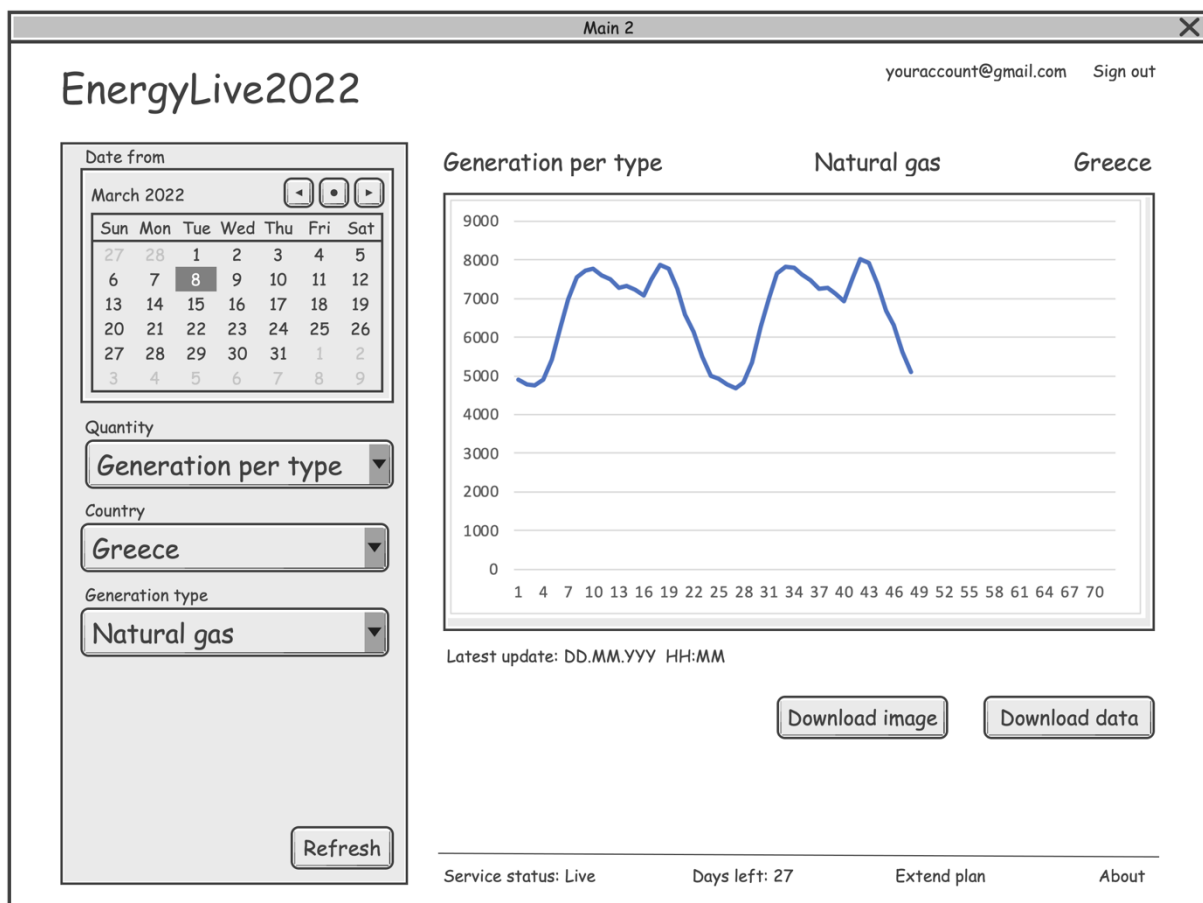
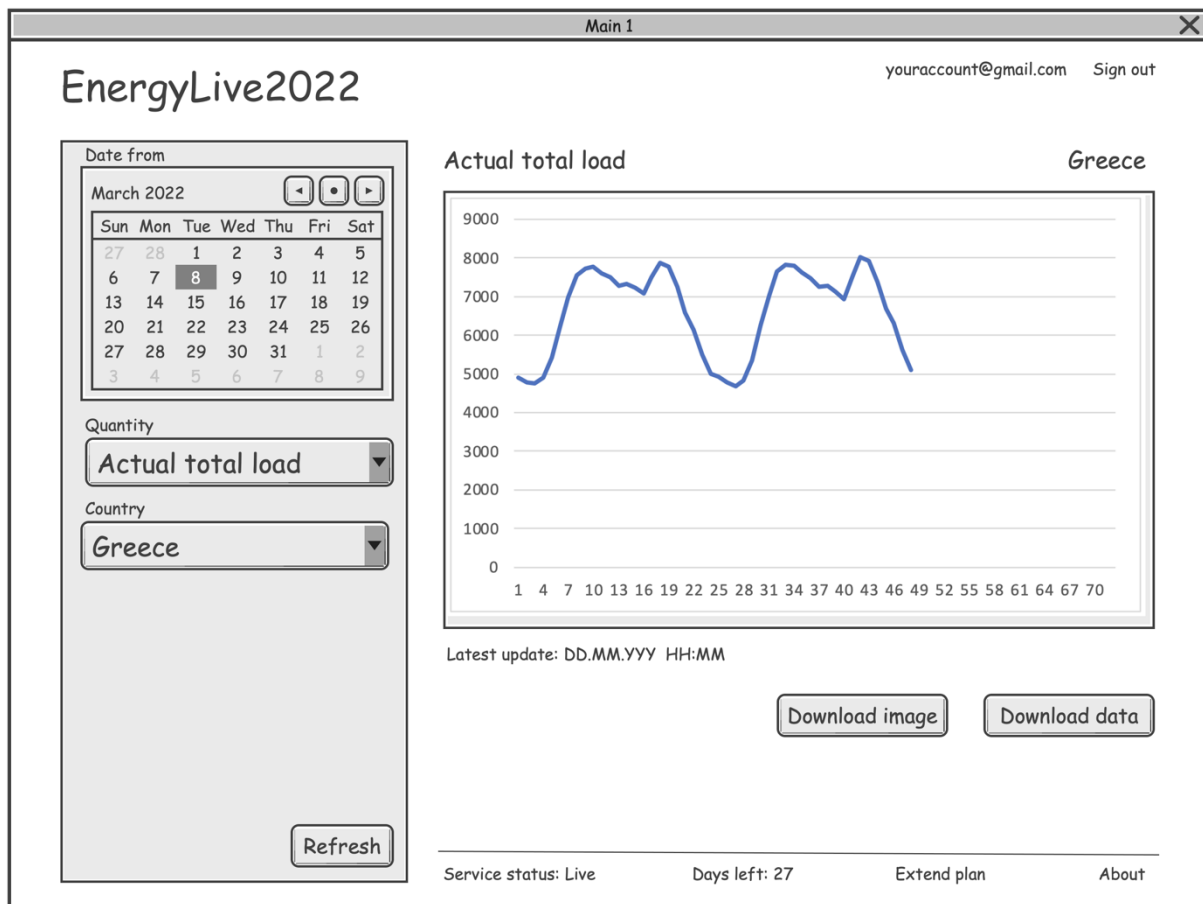
Σύμφωνα με τη λογική του agile software development δεν υπάρχουν ενδιάμεσα παραδοτέα με την έννοια π.χ. της ολοκληρωμένης εκδοχής των διαγραμμάτων ή του κώδικα. Ως τέτοια θα ληφθούν **στιγμιότυπα** του github repository κάθε ομάδας), του YouTrack έργου, καθώς και υποβολές του αρχείου vnrp μέσω του helios, σε συγκεκριμένες χρονικές στιγμές ως ακολούθως:

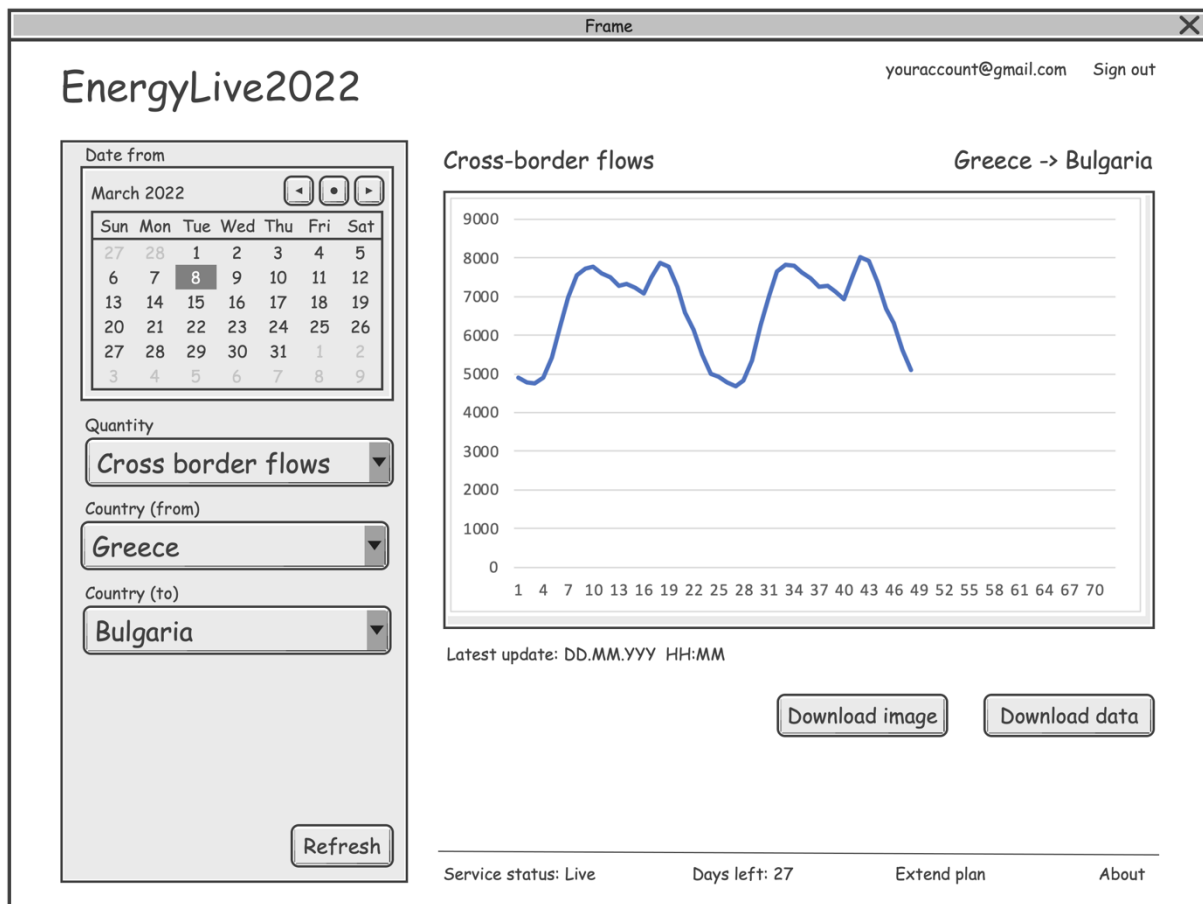
- **Εβδομάδα 25-29.4.2021** (αμέσως μετά από τις διακοπές του Πάσχα): Κάθε ομάδα θα παρουσιάσει στον υποθετικό πελάτη την πρόοδό της με τρόπο που θα ανακοινωθεί και επίσης θα υποβάλει το αρχείο vnrp με τη μέχρι το σημείο εκείνο αρχιτεκτονική και σχεδίαση της εφαρμογής.
- **Την τελευταία εβδομάδα του εξαμήνου**: Κάθε ομάδα θα παρουσιάσει στον υποθετικό πελάτη την πρόοδό της με τρόπο που θα ανακοινωθεί και επίσης θα υποβάλει το αρχείο vnrp με τη μέχρι το σημείο εκείνο αρχιτεκτονική και σχεδίαση της εφαρμογής.
- **Τελική παράδοση**: μετά την εξέταση του μαθήματος με υποβολή του τελικού vnrp και με 10 λεπτες προφορικές παρουσιάσεις

Απαιτήσεις του πελάτη που δίνονται αρχικά

Ακολουθούν οι αρχικές εκδοχές του υποθετικού πελάτη αναφορικά με την επιθυμητή συμπεριφορά της εφαρμογής.







Γενική δομή και κύκλος ενημέρωσης δεδομένων

Ακολουθούν σχήματα από τις διαφάνειες διαλέξεων, με τη δομή και τις ιδιαιτερότητες ενημέρωσης των δεδομένων, όπως καταγράφονται από τα πηγαία αρχεία.

