Παραδοτέο Π5.3: Σχεδιασμός, υλοποίηση και συνεχής λειτουργία του ΚΔΔΔ PANACEA

Πακέτο Εργασίας	ΠΕ 5
Νο Παραδοτέου	Π5.3
Επικεφαλής δικαιούχος	Εθνικό Αστεροσκοπείο Αθηνών
Συγγραφείς	Βασίλης Αμοιρίδης, Αθανάσιος Γεωργίου, Δημήτρης Βαλλιανάτος, Αννα Γιαλιτάκη, Μανώλης Προεστάκης
Συμμετέχοντες	
Εσωτερικοί κριτές	
Τύπος Παραδοτέου	
Επίπεδο Διάχυσης	Εσωτερικό PANACEA
Εκτιμούμενη ημερομηνία Παράδοσης	05/2022
Πραγματική ημερομηνία Παράδοσης	05/2022
Έκδοση	2022-05-01– 1
Σχόλια	

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ ΟΙΚΟΝΟΜΙΑΣ & ΑΝΑΠΤΥΞΗΣ
ΕΠΑνΕΚ 2014-2020
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΤΗΤΑ
ΕΣΠΑ 2014-2020

Περιεχόμενα

Εισαγωγή στο Κέντρο Διαχείρισης και Διάθεσης Δεδομένων (ΚΔΔΔ) 3

Αρχιτεκτονική ΚΔΔΔ 4

Υπηρεσία Διαχείρισης Δεδομένων (ΥΔΔ) 7

Εισαγωγή στο Κέντρο Διαχείρισης και Διάθεσης Δεδομένων (ΚΔΔΔ)

Η Ερευνητική Υποδομή (ΕΥ) ΠΑΝΑΚΕΙΑ αποσκοπεί στην ομογενοποίηση και την συγκέντρωση όλων των ατμοσφαιρικών παρατηρήσεων, κάτω από μία ενιαία εθνική στέγη. Ως προς την εκπλήρωση αυτού του σκοπού, σχεδιάστηκε και υλοποιήθηκε το Κέντρο Διαχείρισης και Διάθεσης Δεδομένων (ΚΔΔΔ). Το ΚΔΔΔ προσφέρει δυνατότητες συλλογής δεδομένων από επίγεια όργανα παθητικής ή ενεργητικής τηλεπισκόπισης, λειτουργίες αυτόματης επεξεργασίας τους αλλά και εργαλεία που υποστηρίζουν την διάχυση των δεδομένων.

Το ΚΔΔΔ σχεδιάστηκε κατάλληλα ώστε να μπορεί να εξυπηρετήσει έναν μεγάλο αριθμό εφαρμογών και εξ' αυτού αποτελείται από ένα σύνολο διακριτών υπηρεσιών και εργαλείων, οι οποίες μπορούν να χρησιμοποιηθούν είτε μεμονωμένα, είτε συνδυαστικά. Χρησιμοποιώντας τα εργαλεία αυτά, οι αναπτύξεις χρησιμοποιήθηκαν για την διαχείριση ή/και την επεξεργασία των δεδομένων από 4 επιστημονικές καμπάνιες. Ο σχεδιασμός όμως επιτρέπει την περαιτέρω ανάπτυξη ώστε να καλυφθούν και άλλες ανάγκες, όπως η συλλογή και διάθεση δεδομένων σε υπηρεσιακή κλίμακα.

Η έκθεση αυτή έχει δομηθεί έτσι ώστε, αρχικά, να αναλυθεί η γενική αρχιτεκτονική του ΚΔΔΔ, ώστε ο αναγνώστης να μπορεί να κατανοήσει πως τα επιμέρους στοιχεία μπορούν να λειτουργήσουν συνδυαστικά για την κάλυψη επιστημονικών αναγκών. Στις επόμενες ενότητες, κάθε ένα από τα εργαλεία αναλύεται σε βάθος ώστε να γίνεται πλήρως κατανοητός ο σχεδιασμός του και η χρήση του. Τέλος, υπάρχει ενότητα που περιγράφει πως το ΚΔΔΔ χρησιμοποιήθηκε σε διάφορες επιστημονικές εφαρμογές και τα αποτελέσματα των πειραμάτων αυτών.

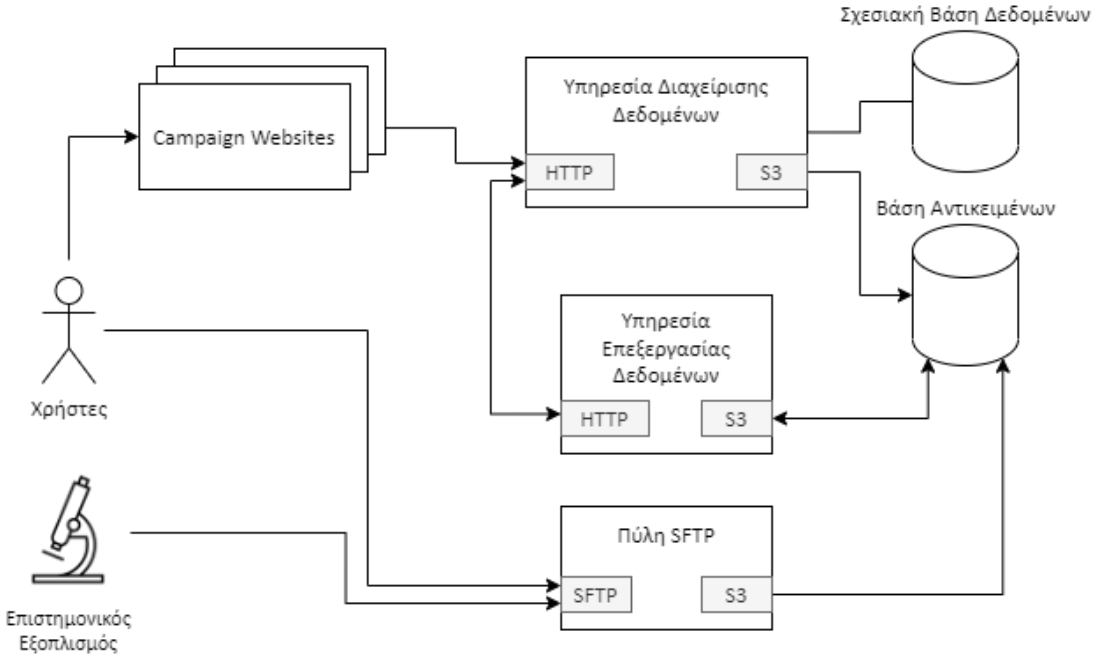
ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ
ΟΙΚΟΝΟΜΙΑΣ & ΑΝΑΠΤΥΞΗΣ
ΕΠΑνΕΚ 2014-2020
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ
ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΤΗΤΑ
ΕΣΠΑ 2014-2020

Αρχιτεκτονική ΚΔΔΔ

Το ΚΔΔΔ αποτελεί έναν κεντρικό κόμβο για την διαχείριση δεδομένων ατμοσφαιρικής τηλεπισκόπισης και σχεδιάστηκε ώστε να έχει την δυνατότητα να υποστηρίξει μεγάλο εύρος εφαρμογών. Κατά τον σχεδιασμό, ταυτοποιήθηκαν οι πυλώνες γύρω από τους οποίους χτίστηκε το ΚΔΔΔ. Συγκεκριμένα, το ΚΔΔΔ θα πρέπει να:

- Υποστηρίζει την αποθήκευση δεδομένων από διαφορετικά είδη οργάνων και για διάφορες φυσικές παραμέτρους
- Επιτρέπει την επεξεργασία των δεδομένων σε σχεδόν πραγματικό χρόνο και να φιλοξενεί τα προϊόντα της επεξεργασίας
- Βοηθά στην διάχυση των δεδομένων μέσω ιστοσελίδων και SFTP αποθετηρίων
- Βοηθά στην οργάνωση επιστημονικών πειραμάτων (campaigns) με την προβολή γραφημάτων (quicklooks) των μετρήσεων αλλά και από αποτελέσματα αριθμητικών ατμοσφαιρικών μοντέλων

Οι παραπάνω επιστημονικές ανάγκες απαιτούν ο σχεδιασμός του ΚΔΔΔ να είναι ευέλικτος και επεκτάσιμος, ώστε να μπορεί να χρησιμοποιηθεί για διαφορετικές επιστημονικές εκστρατείες.



Εικόνα 1: Αρχιτεκτονικό Διάγραμμα του ΚΔΔΔ. Στο σχήμα φαίνονται οι υπηρεσίες από τις οποίες αποτελείται το κέντρο, καθώς και τα πρωτόκολλα που χρησιμοποιούνται για την επικοινωνία ανάμεσα τους.

Η αρχιτεκτονική που υλοποιήθηκε φαίνεται σχηματικά στην Εικόνα 1. Η κεντρική υπηρεσία που είναι υπεύθυνη για την διαχείριση των δεδομένων και την επίβλεψη των άλλων υπηρεσιών είναι η «Υπηρεσία Διαχείρισης Δεδομένων» (ΥΔΔ). Η ΥΔΔ συνδέεται απευθείας με δύο βάσεις δεδομένων. Η πρώτη, η σχεσιακή βάση δεδομένων, είναι υπεύθυνη για την αποθήκευση πληροφοριών για τα επιστημονικά όργανα, τους χρήστες της υποδομής, τα πειράματα που πραγματοποιούνται, καθώς και άλλα σχετικά μεταδεδομένα. Η δεύτερη βάση δεδομένων, η «Βάση Αντικειμένων», είναι υπεύθυνη για την αποθήκευση των αρχείων με τις μετρήσεις που ανεβαίνουν στο κέντρο από τα επιστημονικά όργανα.

Η βάση αντικειμένων είναι ιδανική για την αποθήκευση ανομοιόμορφης πληροφορίας, όπως είναι τα επιστημονικά δεδομένα που προέρχονται από διαφορετικά όργανα. Στην ουσία, σε κάθε αρχείο ανατίθεται ένας μοναδικός κωδικός (key), μέσω του οποίου μπορεί κάποια άλλη υπηρεσία να ανακτήσει τα περιεχόμενα από την βάση. Η βάση δεδομένων που χρησιμοποιήθηκε ακολουθεί το πρότυπο S3 και μπορεί εύκολα να αντικατασταθεί από υπηρεσία αποθήκευσης δεδομένων στο cloud (π.χ. Amazon Object Storage, Azure Blob Storage), εάν απαιτείται από τον όγκο των δεδομένων.

Τα δεδομένα που φτάνουν στην ΥΔΔ και αποθηκεύονται στην βάση αντικειμένων είναι αδιάφανα προς αυτή, δηλαδή δεν υπάρχει μηχανισμός για την ανάλυση των περιεχομένων τους. Ο σχεδιασμός αυτός είναι σκόπιμος, καθώς επιτρέπει τον χειρισμό κάθε είδους μέτρησης, ανεξάρτητα από την μορφή των αρχείων. Είναι όμως επιθυμητό να επεξεργάζονται αυτά τα αρχεία αυτόματα, ώστε να παράγεται «μεταπληροφορία» για τα περιεχόμενα (για παράδειγμα, ποια είναι η περίοδος μέτρησης) αλλά και παράγωγα αρχεία όπως γραφήματα (quicklooks) ή δεδομένα μεγαλύτερου επιπέδου. Τέτοια επεξεργασία δεν μπορεί παρά να σχεδιάζεται ξεχωριστά για κάθε είδος μετρήσεων ή για κάθε φυσική παράμετρο, επομένως προκύπτει η ανάγκη από το ΚΔΔΔ να υποστηρίξει διαφορετικές διεργασίες επεξεργασίας. Η ανάγκη αυτή καλύπτεται από την Υπηρεσία Επεξεργασίας Δεδομένων (ΥΕΔ), η οποία έχει πρόσβαση στα δεδομένα και, σε συγκεκριμένες περιπτώσεις (triggers) μπορεί να εκτελέσει ρουτίνες επεξεργασίας. Οι ρουτίνες αυτές είναι προγράμματα (scripts), που μπορούν να συγγραφούν από το επιστημονικό προσωπικό που διαχειρίζεται το όργανο.

Η επιστημονική κοινότητα χρησιμοποιεί τακτικά τα πρωτόκολλα μεταφοράς δεδομένων FTP και SFTP. Για συμβατότητα του κέντρο με αυτή την παράδοση, αλλά και διατήρηση της ευελιξίας που προσφέρει η βάση αντικειμένων, η Πύλη SFTP προσφέρει πρόσβαση στα αποθηκευμένα δεδομένα μέσω του πρωτοκόλλου SFTP. Χρησιμοποιώντας οποιοδήποτε συμβατό λογισμικό SFTP client, οι χρήστες μπορούν να συνδεθούν στο Κέντρο και να μεταφορτώσουν δεδομένα, χωρίς να γνωρίζουν πως λειτουργεί η υποδομή ή να πρέπει να ανανεώσουν ροές εργασιών με νέο λογισμικό.

Μέσω της ΥΕΔ δίνεται η δυνατότητα για διασύνδεση του κέντρου με άλλες εθνικές και διεθνείς υποδομές, καθώς οι ρουτίνες επεξεργασίας μπορούν να μεταφορτώσουν δεδομένα από και προς το κέντρο. Χρησιμοποιώντας την ΥΕΔ έγινε η διασύνδεση του ΚΔΔΔ με το Single Calculus Chain[1], το σύστημα αυτόματης επεξεργασία δεδομένων Lidar του Ευρωπαϊκού Δικτύου EARLINET.

Εάν απαιτείται από την επιστημονική εκστρατεία ή το πείραμα, υπάρχει η δυνατότητα ανάπτυξης ιστοσελίδας, μέσω της οποίας οι συντελεστές και το κοινό μπορούν να έχουν πρόσβαση σε σχετικές πληροφορίες και τα δεδομένα που προέκυψαν. Χρησιμοποιώντας ως βάση το ΚΔΔΔ, είναι εφικτό να παραχθεί μία ιστοσελίδα με τον ελάχιστο δυνατό κόπο, καθώς

[1] https://scc.imaa.cnr.it/

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ ΥΠΟΥΡΓΕΙΟ ΟΙΚΟΝΟΜΙΑΣ & ΑΝΑΠΤΥΞΗΣ
ΕΠΑνΕΚ 2014-2020 ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΤΗΤΑ
ΕΣΠΑ 2014-2020

οι λειτουργίες έχουν υλοποιηθεί και απλώς αλλάζει ο «σχεδιασμός» της σελίδας για να ταιριάζει με το πείραμα. Αυτή η υπηρεσία χρησιμοποιήθηκε για τα πειράματα ASKOS[2] και CALISTHO[3].

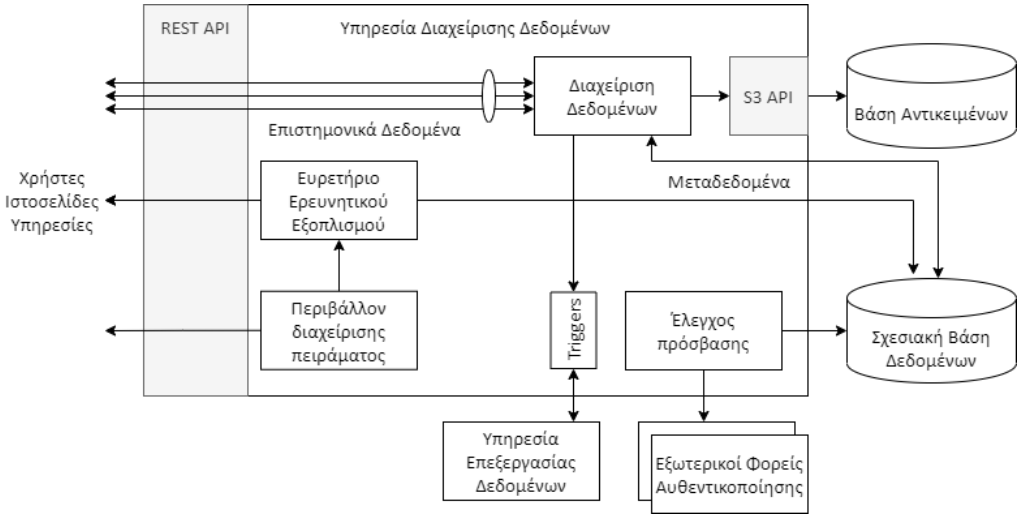
Στην ανάπτυξη του ΚΔΔΔ χρησιμοποιήθηκε αποκλειστικά λογισμικό ανοιχτού κώδικα. Η επικοινωνία με το Κέντρο και μεταξύ των υπηρεσιών του γίνεται μέσω ανοιχτών πρωτοκόλλων. Τα επιμέρους στοιχεία της υποδομής (βάση αντικειμένων, βάση δεδομένων, κτλ.) μπορούν να αντικατασταθούν με άλλες υλοποιήσεις εφόσον αυτό κριθεί απαραίτητο για την κάλυψη των αναγκών του κέντρου. Οι διακριτές υπηρεσίες που αποτελούν το ΚΔΔΔ επικοινωνούν μέσω δικτύου, επομένως δεν υπάρχει η απαίτηση να φιλοξενούνται στην ίδια φυσική υποδομή. Αυτό προσφέρει την ικανότητα για δια-ιδρυματική υποστήριξη του Κέντρου, εάν αυτό απαιτηθεί από τον όγκο των δεδομένων που καλείται να διαχειριστεί.

[2] https://askos.space.noa.gr/
[3] https://calishto.panacea-ri.gr/

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ
ΟΙΚΟΝΟΜΙΑΣ & ΑΝΑΠΤΥΞΗΣ
ΕΠΑνΕΚ 2014-2020
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ
ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΤΗΤΑ
ΕΣΠΑ
2014-2020

Υπηρεσία Διαχείρισης Δεδομένων (ΥΔΔ)

Η Υπηρεσία Διαχείρισης Δεδομένων (ΥΔΔ) είναι η καρδιά του Κέντρου, καθώς διαχειρίζεται την ροή δεδομένων από και προς το Κέντρο, την οργάνωση των διεργασιών επεξεργασίας και διατηρεί το ευρετήριο των διαθέσιμων δεδομένων. Η ΥΔΔ είναι επίσης υπεύθυνη για πιο πεζές αλλά εξίσου σημαντικές λειτουργίες, όπως η αυθεντικοποίηση των χρηστών και άλλων συστημάτων, ο έλεγχος πρόσβασης, την συντήρηση των βάσεων δεδομένων, κ.α. Η εσωτερική αρχιτεκτονική της υπηρεσίας φαίνεται στην Εικόνα 2.



Εικόνα 2: Αρχιτεκτονική της Υπηρεσίας Διαχείρισης Δεδομένων.

Κεντρική στην ΥΔΔ είναι έννοια του πειράματος, κάτω από την οποία οργανώνονται τα επιστημονικά όργανα, οι μετρήσεις και τα δεδομένα τους, καθώς και άλλες οντότητες όπως οι χρήστες και οι διεργασίες αυτόματης επεξεργασίας. Ουσιαστικά, το Κέντρο έχει αναπτυχθεί με την λογική «multi-tenancy», υπό την οποία μπορούν οι υπηρεσίες του να διατίθενται σε διαφορετικούς χρήστες ξεχωριστά. Ωστόσο, το γεγονός ότι υπό την ίδια στέγη στεγάζονται δεδομένα από πολλά πειράματα δεν παραμένει ανεκμετάλλευτο, καθώς επιτρέπεται ο υπηρεσιακός διαμοιρασμός δεδομένων ανάμεσα στα πειράματα. Παράδειγμα «πειραμάτων» για την πλατφόρμα είναι οι επιστημονικές εκστρατείες των ερευνητικών κέντρων.

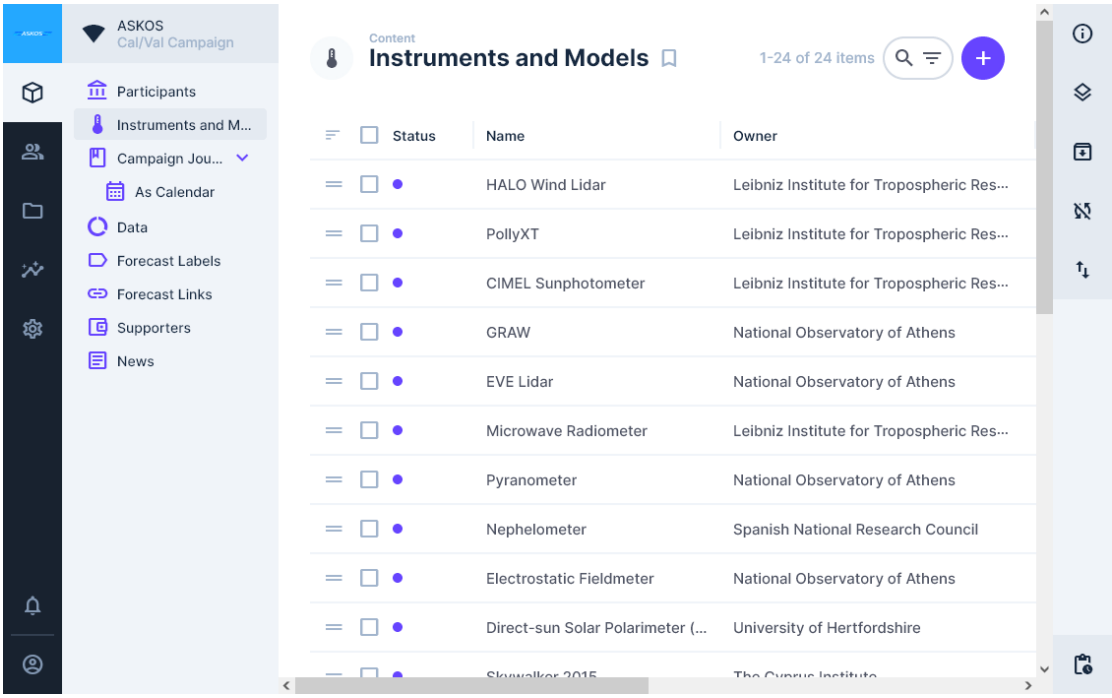
Η αλληλεπίδραση των χρηστών και των άλλων υπηρεσιών με την ΥΔΔ γίνεται μέσω του REST API. Το API αυτό διέπετε από ένα σύστημα ελέγχου πρόσβασης ώστε να έχει ο κάθε χρήστης πρόσβαση στα δεδομένα για τα οποία είναι υπεύθυνος και να μπορεί να διαχειριστεί τα πειράματα στα οποία συμμετέχει. Το κέντρο έχει την δυνατότητα να επεκταθεί με εξωτερικούς φορείς αυθεντικοποίησης (μέσω του πρωτοκόλλου OAUTH2) αλλά αυτή η λειτουργία δεν έχει ενεργοποιηθεί. Εκτός από χρήστες, το Κέντρο μπορεί να διαχειριστεί την αυθεντικοποίηση προγραμμάτων, αποκαλούμενα «bots». Αυτά είναι προγράμματα που μπορούν να χρησιμοποιηθούν από την επιστημονική κοινότητα για την αυτοματοποίηση κάποιας αλληλεπίδρασης με την πλατφόρμα, όπως για παράδειγμα την μεταφόρτωση μετρήσεων από κάποιο σύστημα.

Κατά την μεταφόρτωση δεδομένων από όργανα προς το Κέντρο, η ΥΔΔ είναι υπεύθυνη για να διατηρεί το ευρετήριο δεδομένων ενημερωμένο και για να δίνει εντολή στην ΥΕΔ να εκτελέσει τις απαραίτητες ρουτίνες επεξεργασίας. Στην συνέχεια, το φυσικό αρχείο αποθηκεύε-

ται στην Βάση Αντικειμένων με ένα μοναδικό κλειδί που το χαρακτηρίζει. Στο κλειδί περιέχεται οι αριθμοί του πειράματος και του οργάνου, καθώς και ένας μοναδικός κωδικός για το αρχείο. Η ΥΔΔ υποστηρίζει την χρήση πολλαπλών βάσεων αντικειμένων είτε για λόγους ασφάλειας των δεδομένων (π.χ. αποθήκευση σε δύο data centers) είτε για λόγους επέκτασης χωρητικότητας.

Για κάθε μέτρηση (ή σχετικό αρχείο), είναι δυνατή η αποθήκευση «μεταδεδομένων» που περιγράφουν τα περιεχόμενα. Αυτές οι πληροφορίες μπορούν να προέρχονται κατευθείαν από την πηγή των δεδομένων μαζί με το αρχείο, είτε να παράγονται μέσω κάποιας αυτόματης ροής επεξεργασίας που αναλύει τα περιεχόμενα του αρχείου. Για παράδειγμα, τέτοιες πληροφορίες μπορεί να είναι η φυσική παράμετρος, η περίοδος μέτρησης, η κατάσταση του οργάνου, και άλλα. Σε δεύτερο χρόνο, δίνεται η δυνατότητα αναζήτησης στα δεδομένα με βάση κάποια από αυτές τις τιμές. Δεν υπάρχει περιορισμός στον αριθμό και την μορφή αυτών των μεταδεδομένων, ως προς την καλύτερη εξυπηρέτηση διαφορετικών αναγκών.

Το ΚΔΔΔ προσφέρει στα πειράματα που το απαιτούν ένα περιβάλλον διαχείρισης μέσω ιστοσελίδας, μέσω του οποίου μπορούν οι συμμετέχοντες να διαχειρίζονται τα δεδομένα του πειράματος αλλά και άλλες πληροφορίες, όπως τα όργανα και οι φορείς που συμμετέχουν, τα νέα και το πρόγραμμα του πειράματος. Στην Εικόνα 3 παρατίθεται φωτογραφία του περιβάλλοντος διαχείρισης του πειράματος ASKOS. Το περιβάλλον διαχείρισης βασίστηκε στο λογισμικό ανοιχτού κώδικα Directus[4].



Εικόνα 3: Περιβάλλον διαχείρισης του πειράματος ASKOS.

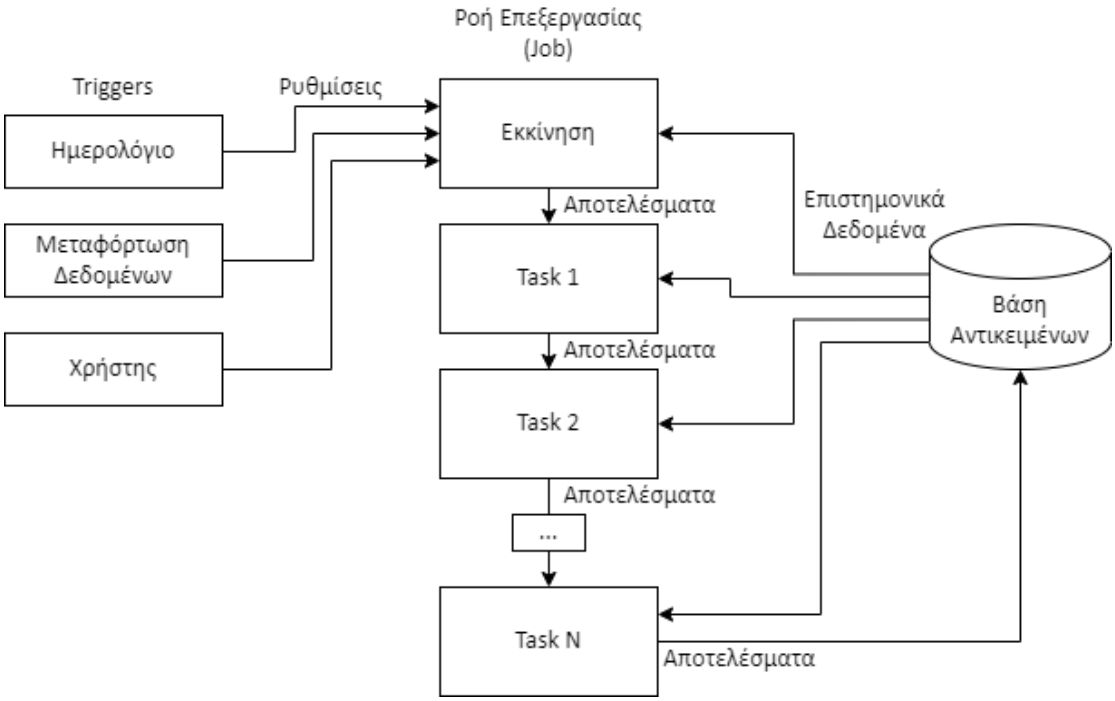
Η ΥΔΔ έχει υλοποιηθεί χρησιμοποιώντας Node.JS και μπορεί να φιλοξενηθεί σε οποιοδήποτε περιβάλλον Linux. Ως σχεσιακή βάση δεδομένων υποστηρίζεται η PostgreSQL ή συμβατές της.

[4] https://directus.io/

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ
ΟΙΚΟΝΟΜΙΑΣ & ΑΝΑΠΤΥΞΗΣ
ΕΠΑνΕΚ 2014-2020
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ
ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΤΗΤΑ
ΕΣΠΑ
2014-2020

Υπηρεσία Επεξεργασίας Δεδομένων

Για την κάλυψη των αναγκών αυτόματης επεξεργασίας δεδομένων του Κέντρου, αναπτύχθηκε σχετική υπηρεσία για την εκτέλεση ρουτίνων επεξεργασίας. Η Υπηρεσία Επεξεργασία Δεδομένων (ΥΕΔ) διαχειρίζεται την εκτέλεση ροών επεξεργασίας που ονομάζονται jobs. Οι ροές αυτές είναι ακολουθίες από παραμετροποίησιμα προγράμματα Python (ονομασμένα Tasks), τα οποία εκτελούνται αυτόματα ή χειροκίνητα. Στην Εικόνα 4 φαίνεται ένα σχεδιάγραμμα που περιγράφει την λειτουργία της υπηρεσίας.



Εικόνα 4: Σχεδιάγραμμα λειτουργίας της Υπηρεσίας Επεξεργασίας Δεδομένων.

Οι ροές επεξεργασίας αποτελούνται από προγράμματα Python, τα οποία μπορούν να παραμετροποιηθούν από ένα σύνολο ρυθμίσεων. Οι ρυθμίσεις αυτές περιλαμβάνουν τα κλειδιά των νέων δεδομένων προς επεξεργασία αλλά μπορούν να τροποποιούν το πρόγραμμα και με άλλους τρόπους, ώστε να αλλάζει η μορφή επεξεργασίας ή το είδος του αποτελέσματος. Μέσω της παραμετροποίησης, μπορεί να χρησιμοποιηθεί η ίδια ροή επεξεργασίας για διαφορετικά όργανα, χρησιμοποιώντας διαφορετικό σύνολο ρυθμίσεων. Κάθε ροή αποτελείται από επί μέρους προγράμματα που ονομάζονται Tasks, τα οποία είναι υπεύθυνα για μία εργασία το κάθε ένα. Επιτρέπεται ο διαμοιρασμός των Tasks ανάμεσα σε ροές, όπου αυτό είναι θεμιτό.

Οι ροές επεξεργασίας έχουν πρόσβαση στα δεδομένα από την βάση αντικειμένων, στην οποία αποθηκεύονται και τα αποτελέσματα της επεξεργασίας. Επιτρέπεται επίσης στις ροές να επικοινωνήσουν με την ΥΔΔ ώστε να έχουν πρόσβαση στο ευρετήριο δεδομένων. Οι ροές εκτελούνται σε διαφορετικό φυσικό υπολογιστή από τα υπόλοιπα στοιχεία της υποδομής, συμπεριλαμβανομένου της ΥΕΔ, ώστε να μην υπάρχει περίπτωση κάποιο σφάλμα στον κώδικα επεξεργασίας να επηρεάσει τις υπηρεσίες του Κέντρου.

Για την εκκίνηση των ροών επεξεργασίας υποστηρίζεται μία συλλογή από «triggers». Εκτός από την βασική εκκίνηση που μπορεί να κάνει ένας χρήστης χειροκίνητα, δίνοντας στο σύστημα τις κατάλληλες ρυθμίσεις για την ροή, μπορεί να γίνει εκκίνηση με βάση ένα ημερολόγιο (π.χ. κάθε μέρα) ή κατά την μεταφόρτωση νέων δεδομένων στην πλατφόρμα.

Κάθε εκτέλεση των ροών αποθηκεύεται στο σύστημα, μαζί με οποιαδήποτε πληροφορία παράγουν τα προγράμματα. Το ιστορικό αυτό διατηρείται ώστε να γίνεται ευκολότερη η διερεύνηση σφαλμάτων που τυχών συμβούν, αλλά και για να μειωθεί η τριβή που είναι φυσική όταν πρέπει να αναπτυχθεί λογισμικό που τρέχει σε κάποιο απομακρυσμένο σύστημα. Οι συμμετέχοντες στα πειράματα έχουν πρόσβαση σε όλες τις πληροφορίες για τις ροές του πειράματος και μπορούν να ενεργοποιούν ή να απενεργοποιούν κάποια ροή σε πραγματικό χρόνο.

Αν και οι ροές είναι προγράμματα Python, επιτρέπεται η εκτέλεση προγραμμάτων που έχουν συγγραφεί με άλλες γλώσσες προγραμματισμού. Στο πείραμα ASKOS, η ΥΕΔ εκτελούσε καθημερινά μία ροή Python, η οποία με την σειρά της εκτελούσε ένα αριθμητικό μοντέλο της ατμόσφαιρας, γραμμένο στην γλώσσα FORTRAN. Έπειτα, η ροή φρόντιζε για την μεταφόρτωση των αποτελεσμάτων του μοντέλου στην πλατφόρμα, από όπου γινόντουσαν διαθέσιμα στην ιστοσελίδα του πειράματος. Για να επιτευχθεί η μέγιστη δυνατή συμβατότητα με τα ήδη υπάρχοντα προγράμματα επεξεργασίας, οι ροές μπορεί να αποτελούνται από προγράμματα μικρού μεγέθους που απλώς καλούν άλλα, ήδη υπάρχοντα προγράμματα.

Η ΥΕΔ χρησιμοποιεί το σύστημα ανοιχτού κώδικα Dagster[5] για την διαχείριση της εκτέλεσης των ροών.

[5] https://www.dagster.io/

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ
ΟΙΚΟΝΟΜΙΑΣ & ΑΝΑΠΤΥΞΗΣ
ΕΠΑνΕΚ 2014-2020
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ
ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΤΗΤΑ
ΕΣΠΑ
2014-2020

Εφαρμογές του ΚΔΔΔ

Για την σωστή αξιολόγηση της ανάπτυξης του ΚΔΔΔ, το σύστημα εφαρμόστηκε σε 3 επιστημονικά πειράματα. Μέσα από την εφαρμογή των υπηρεσιών του ΚΔΔΔ σε πραγματικές επιστημονικές εκστρατείες συλλέχθηκαν σχόλια για την αποτελεσματικότητα του και έγιναν τροποποιήσεις στον σχεδιασμό ώστε να υποστηρίζονται καλύτερα οι επιστημονικοί στόχοι. Στις υπό-ενότητες που ακολουθούν παρατίθενται πληροφορίες για κάθε ένα από τα πειράματα που χρησιμοποιήσαν πόρους του ΚΔΔΔ.

ASKOS

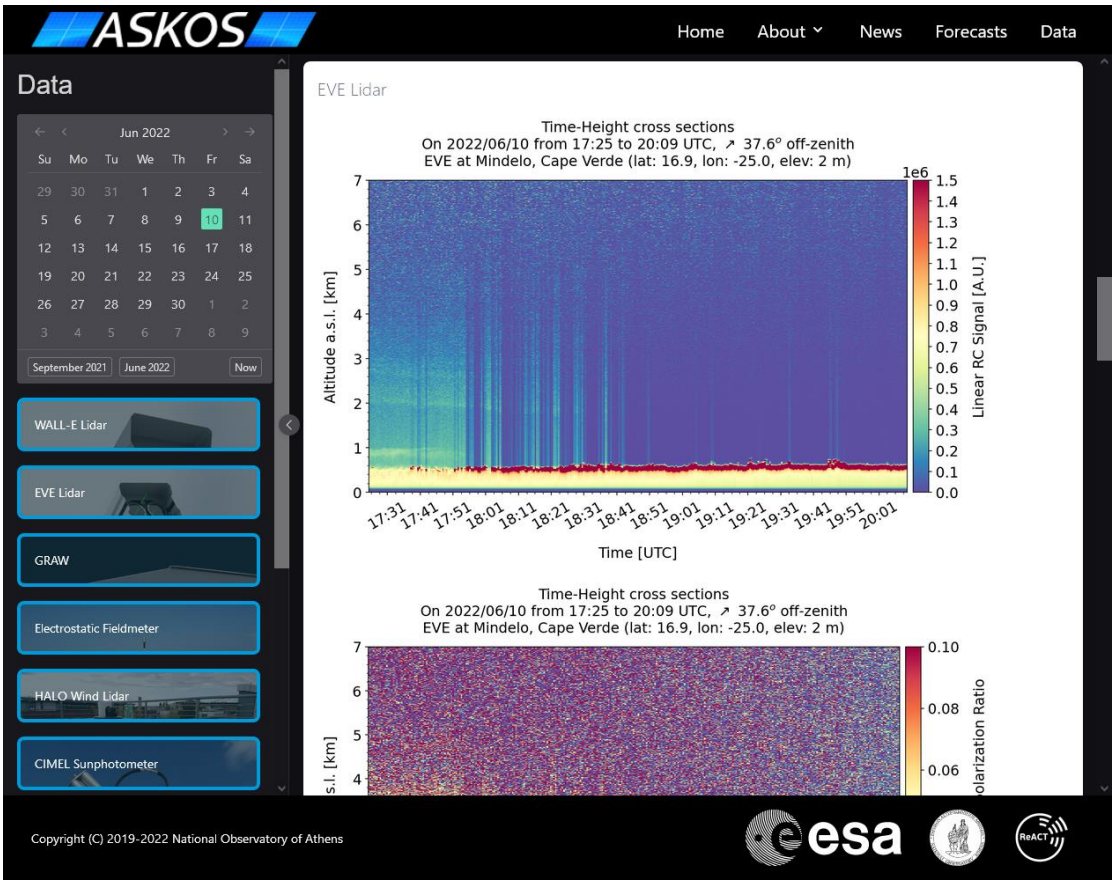
Το πείραμα ASKOS είναι μία δράση του Εθνικού Αστεροσκοπείου Αθηνών (ΕΑΑ) σε συνεργασία με την Ευρωπαϊκή Διαστημική Υπηρεσία (ESA) και 22 άλλους φορείς, με σκοπό την αξιολόγηση των προϊόντων της δορυφορικής αποστολής AEOLUS. Το πείραμα έβαλε χώρα στο Πράσινο Ακρωτήρι, που λόγω της ιδιαίτερης θέσης του θεωρείτε ιδανικό για την παρατήρηση της μεταφοράς αερολυμάτων από τις ερήμους της βόρειας Αφρικής προς των Ατλαντικό Ωκεανό. Οι πρώτες δύο φάσεις του πειράματος έχουν ολοκληρωθεί, η πρώτη το καλοκαίρι του 2021 και η δεύτερη το καλοκαίρι του 2022 και αναμένεται η πραγματοποίηση της τελευταίας φάσης τον Σεπτέμβριο του 2022.

Το ΚΔΔΔ υποστήριξε την διεθνή δράση του ΕΑΑ με την προσφορά υπηρεσιών για τις ανάγκες διαχείρισης δεδομένων του πειράματος. Μέσω των υπηρεσιών του ΚΔΔΔ έγινε η αποθήκευση των δεδομένων από 12 διαφορετικά επιστημονικά όργανα, για τα οποία υπήρξαν ρουτίνας αυτόματης επεξεργασίας για την παραγωγή γραφημάτων σε σχεδόν πραγματικό χρόνο. Το ΚΔΔΔ ήταν επίσης υπεύθυνο για την αυτόματη λήψη δεδομένων από διεθνείς υποδομές που συμμετείχαν στο πείραμα μέσω διασυνδέσεων που αναπτύχθηκαν. Πιο συγκεκριμένα, αναπτύχθηκαν ροές διασύνδεσης του ΚΔΔΔ με τις διεθνείς υποδομές ACTRIS/Cloudnet[6], PollyNET[7] και AERONET[8]. Τέλος, το ΚΔΔΔ ήταν υπεύθυνο για την εκτέλεση αριθμητικών μοντέλων της ατμόσφαιρας σε πόρους του ΕΑΑ και για την εμφάνιση των αποτελεσμάτων στην ιστοσελίδα του πειράματος.

Εικόνα 5: Φωτογραφία της ιστοσελίδας του πειράματος ASKOS στις 28/07/2022. Η σελίδα είναι διαθέσιμη στο https://askos.space.noa.gr.

[6] https://cloudnet.fmi.fi/
[7] https://polly.tropos.de/
[8] https://aeronet.gsfc.nasa.gov/

Η ιστοσελίδα του πειράματος αναπτύχθηκε χρησιμοποιώντας το λογισμικό του ΚΔΔΔ και προσφέρει πρόσβαση σε διαγράμματα των μετρήσεων και των προβλέψεων από ατμοσφαιρικά μοντέλα. Ο χρήστης μπορεί μέσω ενός εύχρηστου περιβάλλοντος να επιλέξει την ημερομηνία μέτρησης (ή πρόβλεψης) και να περιηγηθεί στα διαθέσιμα δεδομένα. Ως προς την ευκολία των ερευνητών που εργάζονται στο πεδίο κατά την διάρκεια του πειράματος, η ιστοσελίδα είναι διαθέσιμη και σε κινητές συσκευές.



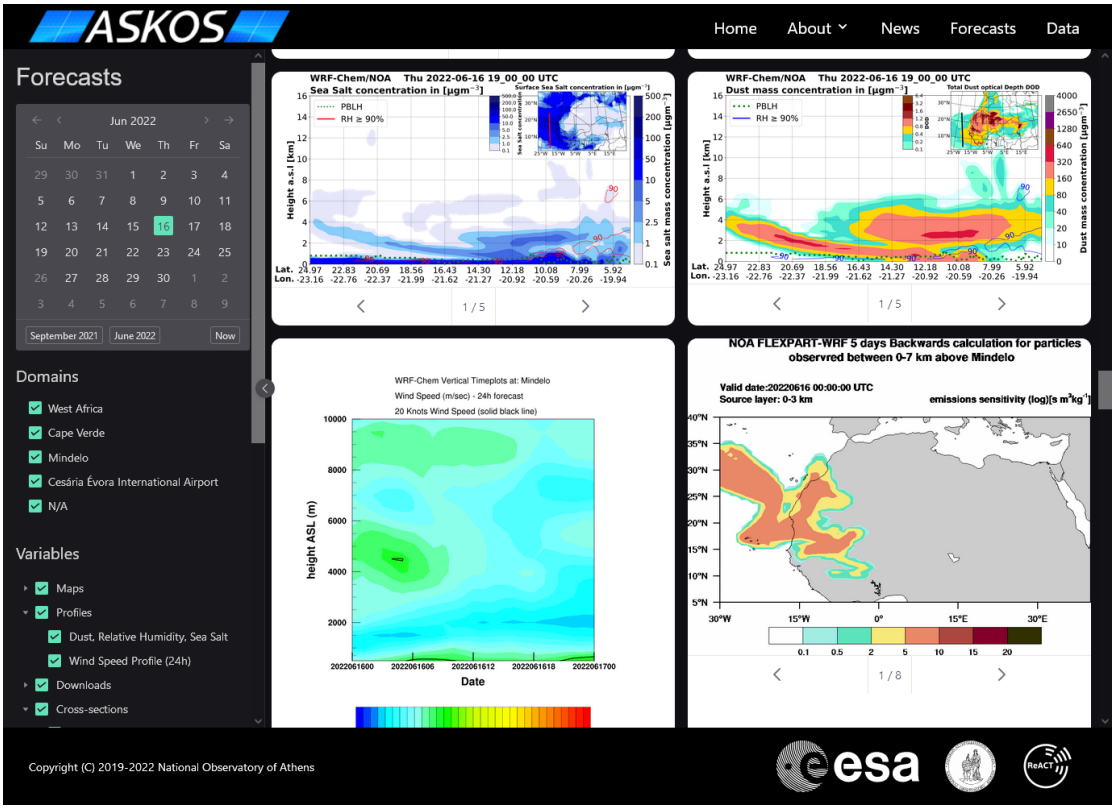
Εικόνα 6: Φωτογραφία του περιβάλλοντος περιήγησης δεδομένων του πειράματος ASKOS.

Για την υποστήριξη του πειράματος αναπτύχθηκαν πολλαπλές ροές επεξεργασίας δεδομένων. Αυτές χωρίζονται σε τρεις διακριτές κατηγορίες:

- Δημιουργία διαγραμμάτων για τα δεδομένα που μεταφέρονται στο ΚΔΔΔ κατά την διάρκεια του πειράματος
- Διασύνδεση του ΚΔΔΔ με άλλους φορείς για την μεταφόρτωση δεδομένων σε πραγματικό χρόνο, συγκεκριμένα:
    - Μεταφόρτωση δεδομένων μετρήσεων και διαγραμμάτων από Cloud Radar από την διεθνή ACTRIS/Cloudnet
    - Μεταφόρτωση δεδομένων Lidar από την υποδομή PollyNET
    - Μεταφόρτωση γραφημάτων από την υποδομή AERONET
- Διαχείριση αριθμητικών μοντέλων της ατμόσφαιρας και των αποτελεσμάτων τους.

Το ΚΔΔΔ έδωσε την δυνατότητα στην ερευνητική ομάδα να έχει πρόσβαση σε διαγράμματα και προβλέψεις σε σχεδόν πραγματικό χρόνο, διευκολύνοντας το επιστημονικό έργο στο πεδίο και την λήψη αποφάσεων.

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ
ΟΙΚΟΝΟΜΙΑΣ & ΑΝΑΠΤΥΞΗΣ
ΕΠΑνΕΚ 2014-2020
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ
ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΤΗΤΑ
ΕΣΠΑ 2014-2020



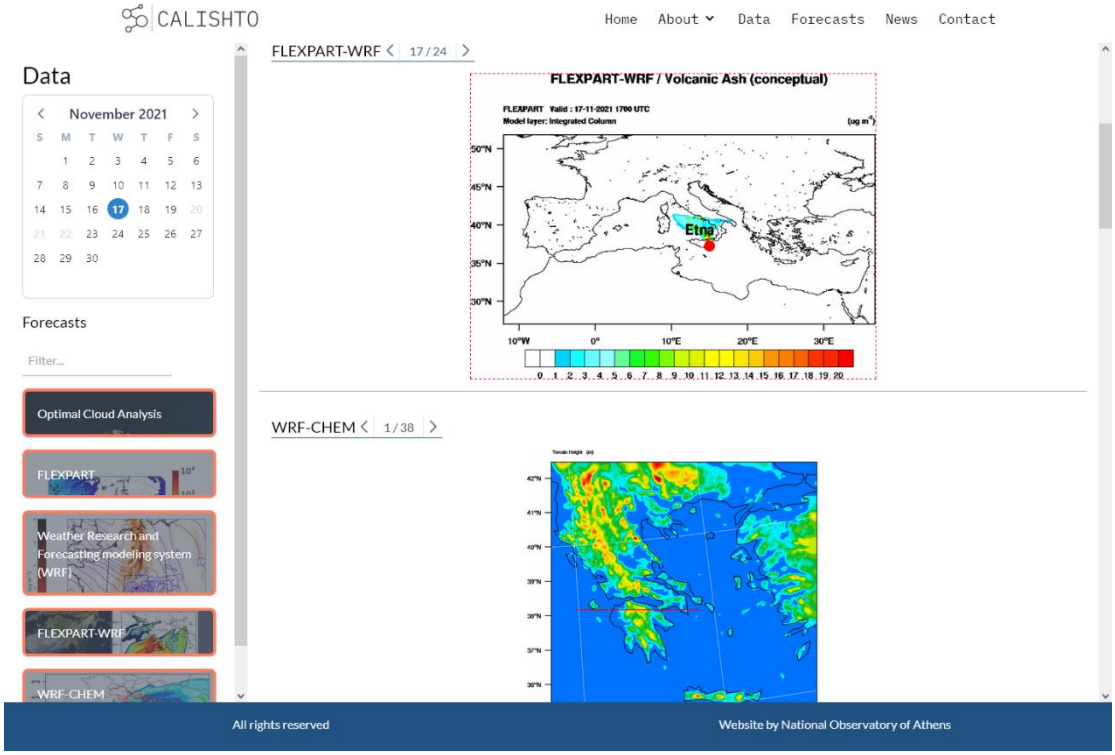
Εικόνα 7: Φωτογραφία περιβάλλοντος περιήγησης προβλέψεων του πειράματος ASKOS.

CALISHTO

Η CALISHTO είναι μία πειραματική εκστρατεία που έλαβε μέρος το φθινόπωρο του 2021 στο Παρατηρητήριο Ατμοσφαιρικών Αερολυμάτων και Κλιματικής Αλλαγής του Χελμού (HAC). Το πείραμα διοργανώθηκε από το Εθνικό Κέντρο Έρευνας Φυσικών Επιστημών «Δημόκριτος», σε συνεργασία με 10 άλλους φορείς και πανεπιστήμια. Το ΚΔΔΔ υποστήριξε το πείραμα με την κατασκευή ιστοσελίδας για την διοργάνωση του πειράματος και την διάχυση προβλέψεων από αριθμητικά μοντέλα.

Εικόνα 8: Φωτογραφία της ιστοσελίδας του πειράματος CALISHTO. Η ιστοσελίδα είναι διαθέσιμη στο https://calishto.panacea-ri.gr/.

Μέσω της ιστοσελίδας είναι δυνατή η πρόσβαση σε διαγράμματα προβλέψεων από αριθμητικά μοντέλα που χρησιμοποιήθηκαν για την διάρκεια του πειράματος. Υπήρχε η δυνατότητα για εμφάνιση διαγραμμάτων από μετρήσεις αλλά δεν χρειάστηκε για την επίτευξη των επιστημονικών στόχων του πειράματος και δεν χρησιμοποιήθηκε.



Εικόνα 9: Φωτογραφία του περιβάλλοντος περιήγησης προβλέψεων του πειράματος CALISHTO.

Για την μεταφόρτωση των διαγραμμάτων για τις ατμοσφαιρικές προβλέψεις, υλοποιήθηκε ροή διασύνδεσης του ΚΔΔΔ με υποδομή της Ακαδημίας Αθηνών, από την οποία προσφέρθηκαν πόροι για την εκτέλεση των ατμοσφαιρικών μοντέλων.

EARLINET COVID19 Campaign

Τον Μάιο του 2020 οργανώθηκε επιστημονική εκστρατεία από το δίκτυο EARLINET για την ποσοτικοποίηση της μείωσης των αερολυμάτων στην ατμόσφαιρα λόγω των απαγορεύσεων κυκλοφορίας που πολλά κράτη είχαν επιβάλει λόγω του ιού SARS-COV-19[9]. Το Παρατηρητήριο Κλιματικής Αλλαγής PANGEA των Αντικυθήρων, ως μέλος του δικτύου EARLINET, συμμετείχε στο πείραμα με την πραγματοποίηση μετρήσεων 24/7 και την μεταφορά τους σε πραγματικό χρόνο στο σύστημα Single Calculus Chain για την επεξεργασία τους.

Η μεταφορά των μετρήσεων πραγματοποιήθηκε μέσω του ΚΔΔΔ, δημιουργώντας μία ροή μεταφοράς δεδομένων από τα Αντικύθηρα στο Κέντρο, και μία ροή διασύνδεσης του Κέντρου με το σύστημα Single Calculus Chain. Η ροή διασύνδεσης υλοποιήθηκε πρώτα σαν αυτόνομο πρόγραμμα Python, το οποίο είναι διαθέσιμο στην επιστημονική κοινότητα[10], και έπειτα με-

[9] https://www.earlinet.org/index.php?id=covid-19
[10] https://github.com/NOA-ReACT/PollyXT-SCC-Pipelines

τατράπηκε σε ροή για το ΚΔΔΔ. Αυτή η τακτική επιτρέπει και σε άλλους οργανισμούς να χρησιμοποιούν τις αναπτύξεις που έγιναν για το ΚΔΔΔ χωρίς να χρησιμοποιούν αναγκαστικά και τους πόρους της υποδομή.