

Πιλοτικές Εφαρμογές

Προδιαγραφές, Σχεδιασμός, Ανάπτυξη και Αξιολόγηση

Γιώργος Αλεξίου

Επιστημονικός Συνεργάτης

Ινστιτούτο Πληροφοριακών Συστημάτων



ΑΟΗΝΑ

Έρευνα & Καινοτομία
Τεχνολογίες Πληροφορίας

Τα προαναφερθέντα έργα υλοποιήθηκαν στο πλαίσιο της Πρόσκλησης με κωδικό 031 με τίτλο «ΔΡΑΣΗ ΣΤΡΑΤΗΓΙΚΗΣ ΑΝΑΠΤΥΞΗΣ ΕΡΕΥΝΗΤΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΩΝ ΦΟΡΕΩΝ», που χρηματοδοτείται από το Επιχειρησιακό Πρόγραμμα «Ανταγωνιστικότητα, Επιχειρηματικότητα και Καινοτομία», ΕΣΠΑ 2014-2020 ΕΣΠΑ 2014-2020, με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης (Ευρωπαϊκό Ταμείο Περιφερειακής Ανάπτυξης). Αναλυτικά: "Υπολογιστικές Επιστήμες και Τεχνολογίες Δεδομένων, Περιεχομένου και Αλληλεπίδρασης" (MIS 5002437) με 3 υποέργα: Υποέργο 1 «Γλωσσικές Τεχνολογίες Ανάλυσης Περιεχομένου και Αλληλεπίδρασης», Υποέργο 2 «Τεχνολογίες Ανάλυσης Περιεχομένου στον Πολιτισμό», Υποέργο 3 «ΜΕΔΑ: Μεγάλα Δεδομένα – Προκλήσεις, Μέθοδοι και Τεχνικές Αποδοτικής Διαχείρισης» και "Καινοτόμος εφαρμογή του βιομηχανικού διαδικτύου των πραγμάτων σε ευφυή περιβάλλοντα" (MIS 5002434) με Υποέργο 1: "Καινοτόμος εφαρμογή του βιομηχανικού διαδικτύου των πραγμάτων σε ευφυή περιβάλλοντα".



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Ταμείο
Περιφερειακής Ανάπτυξης



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ
ΑΝΑΠΤΥΞΗΣ ΚΑΙ ΕΠΙΧΕΙΡΗΣΕΩΝ
ΕΘΝΙΚΗ ΓΡΑΜΜΑΤΕΙΑ ΔΙΑΧΕΙΡΙΣΗΣ ΠΡΟΓΡΑΜΜΑΤΩΝ
ΕΘΝΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ ΕΠΙΧΕΙΡΗΣΙΑΚΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ
ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΤΗΤΑ
ΚΑΙΝΟΤΟΜΙΑ



ανάπτυξη - εργασία - αλληλεγγύη

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Στόχοι και Εργασίες

Ανάλυση Απαιτήσεων



Σχεδιασμός Πλοτικών Εφαρμογών




Αξιολόγηση Πλοτικών Εφαρμογών



Ανάπτυξη Πλοτικών Εφαρμογών





Ανάλυση Απαιτήσεων και Σχεδιασμός Πιλοτικών Εφαρμογών

1. Σχεδιασμός πιλοτικών εφαρμογών με βάση τις μεθόδους και τα εργαλεία που προέκυψαν από τις προηγούμενες ενότητες
 - Αποτύπωση των λειτουργικών απαιτήσεων
 - Απάντηση βασικών ερωτημάτων μέσω της ανάλυσης των δεδομένων για κάθε πιλοτική εφαρμογή





Μέθοδοι Προετοιμασίας, Επιμέλειας και Εξερεύνησης Δεδομένων

2. Σχεδιασμός και υλοποίηση μεθόδων προετοιμασίας, επιμέλειας και εξερεύνησης των δεδομένων.
 - καθαρισμός (π.χ., απαλοιφή διπλοτύπων),
 - μοντελοποίηση (π.χ. να αναπαρασταθούν σε κάποιο κατάλληλο μοντέλο δεδομένων),
 - ανωνυμοποίηση (για την απαλοιφή ευαίσθητων στοιχείων)
 - Σημασιολογικός εμπλουτισμός
 - εξερεύνηση των χαρακτηριστικών τους, (π.χ. η ανίχνευση της δομής, του είδους συσχετίσεων μέσα στο σύνολο δεδομένων, η στατιστική ανάλυση τιμών, κ.α)





Ανάπτυξη Πιλοτικών Εφαρμογών

1. Ανάπτυξη των πιλοτικών που θα ενσωματώνουν τις τεχνολογίες που έχει αναπτύξει το ΙΠΣΥ αλλά και διαθέσιμων με τις οποίες εξοικειώθηκε στο προσωπικό του στις προηγούμενες ενότητες.
 - Αποτελεσματική υλοποίηση των βασικών αλγορίθμων και της τεχνογνωσίας του ΙΠΣΥ.
 - Ευκολία χρήσης.
 - Ανταπόκριση σε πραγματικές ανάγκες.





Αξιολόγηση Πιλοτικών Εφαρμογών

2. Αποτίμηση των δυνατοτήτων των εφαρμογών αλλά και προσαρμογή της υλοποίησή τους στους προαναφερθέντες στόχους.
 - Εκτεταμένη πειραματική αποτίμηση των μεθόδων, σύμφωνα με τις βέλτιστες επιστημονικές πρακτικές
 - Αξιολόγηση από χρήστες.



Πιλοτικές Εφαρμογές





PHONY



Αυτόματη δημιουργία συνόλων δεδομένων για την
ανίχνευση ψευδών ειδήσεων στα μέσα κοινωνικής
δικτύωσης





PHONY – Προδιαγραφές

Προδιαγραφές

- Δημιουργία συνόλων δεδομένων για εκπαίδευση μοντέλων ανίχνευσης ψευδών ειδήσεων
- Δυνατότητα επιλογής τύπου, χρονικής περιόδου και πηγής των ειδήσεων
- Αυτόματη δημιουργία
- Δυνατότητα εξαγωγής όλων των ιδιοτήτων (γλωσσικές, σημασιολογικές, δικτυακές)
- Πλήρης αναπαράσταση της διαδρομής fake news στο δίκτυο

Δεδομένα από:

- Streaming tweets
- Ψευδείς ειδήσεις από fact-checking services



PHONY – Ανάπτυξη

Ευρετήριο Streaming Tweets

- διεύρυνση με περιεχόμενο URL
- ανεστραμμένο ευρετήριο: term, {tweet1..tweetn}

Συλλογή και Ανάλυση Ψευδών Ειδήσεων

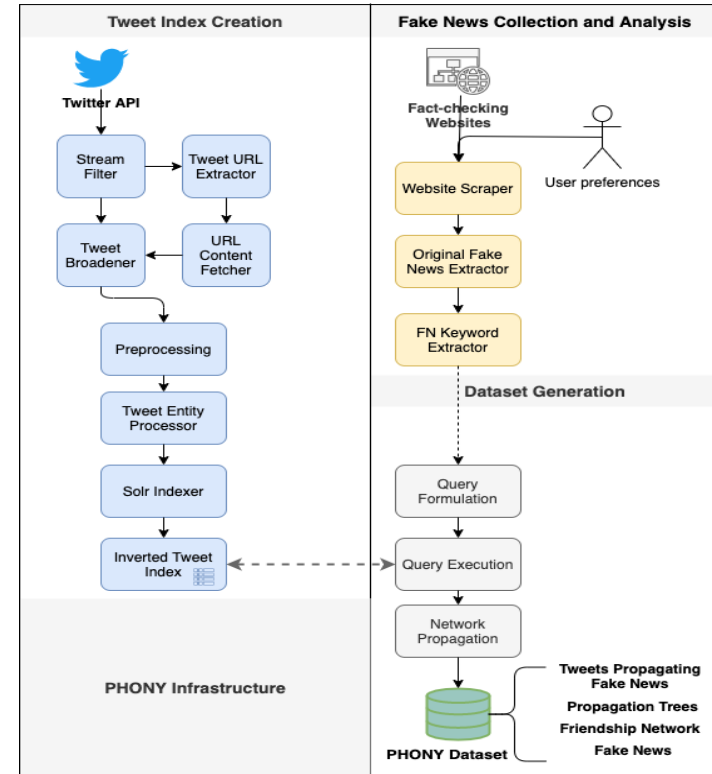
- scraping υπηρεσιών ελέγχου ειδήσεων
- εξαγωγή λέξεων κλειδίων

Δημιουργία Συνόλου Δεδομένων

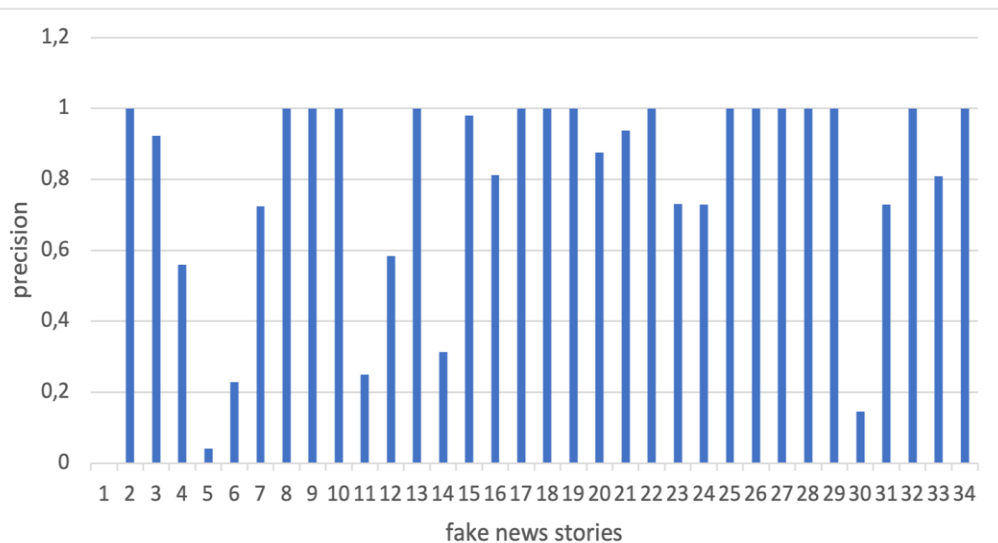
- δημιουργία και εκτέλεση query στο ευρετήριο
- εξαγωγή Δεδομένων διάδοσης

PHONY Datasets

- ψευδείς ειδήσεις
- tweets που διαδίδουν ψ.ε.
- δένδρα διάδοσης ψ.ε.
- δίκτυο φίλων που διαδίδει ψ.ε.



PHONY – Αξιολόγηση



- Ελληνικά tweets από Ιανουάριο 2016 έως Απρίλιο 2018
- ψευδείς ειδήσεις από EllinikaHoaxes από Ιανουάριο 2016 έως Απρίλιο 2018
- average precision 77,56%
- 19 fake news (54%) πέτυχαν precision >90%





BIP! Finder



BIP! Finder

Μηχανή αναζήτησης επιστημονικών δημοσιεύσεων
με βάση τη δημοφιλία και την επιρροή





BiP! Finder – Προδιαγραφές

Ανάπτυξη μηχανής αναζήτησης που ενσωματώνει μετρικές απήχησης (δημοφιλία – επιρροή): BiP! Finder

- Πλούσια δεδομένα διεπιστημονικών γράφων αναφορών (Crossref)
- Αναζήτηση βάσει επιρροής
- Αναζήτηση βάσει δημοφιλείας
- Δυνατότητα συγκρίσεων άρθρων με βάση την απήχηση κ.α.

Ανάπτυξη μηχανής αναζήτησης για άρθρα σχετικά με covid-19:
BiP4Covid19

- Επέκταση κριτηρίων κατάταξης χρήση επιπλέον #readers (mendeley) & #tweets που έγιναν πρόσφατα



BIP! Finder – Ανάπτυξη

BiP! Finder Interface



BIP! Finder

clear all Influence

Exceptional (Top 0.01%)
 Substantial (Top 1%)
 Average (All)

Popularity

Exceptional (Top 0.01%)
 Substantial (Top 1%)
 Average (All)

Start Year

Starting Publication Year

End Year

Ending Publication Year

Venue

Select Venues

text mining

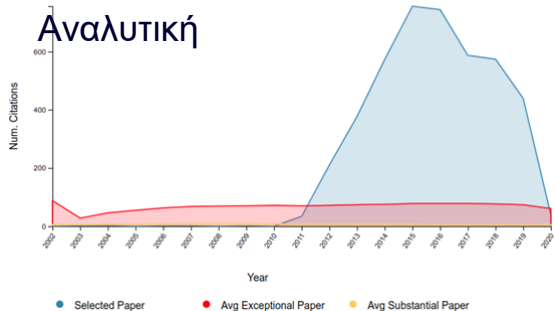
Order by: Popularity Influence Year Combine with keyword relevance: Yes No

6381 results (320 pages)

Click on entries for comparison

Title	Venue	Year	Impact
Systematic and integrative analysis of large gene lists using DAVID bioinformatics resourc (...)	context Nature Protocols	2008	🔥 📄 📌
The WEKA data mining software: an update	context Sigmdd Explorations	2009	🔥 📄 📌
The STRING database in 2017: quality-controlled protein-protein association networks, ma (...)	context Nucleic Acids Research	2017	🔥 📄 📌
Opinion Mining and Sentiment Analysis	context Foundations And Trends In (...)	2008	🔥 📄 📌

BiP! Finder Αναλυτική



BiP4Covid19 Interface

BIP! Finder for COVID-19

This version of **BiP! Finder** aims to ease the exploration of **COVID-19**-related literature by enabling **ranking articles based on various impact metrics**.

Last Update: 04 - 06 - 2021

Provided impact measures:

- Popularity:** Citation-based measure reflecting the current impact.
 - Influence:** Citation-based measure reflecting the total impact.
 - Reader Attention:** The current number of Mendeley readers.
 - Social Media Attention:** The number of recent tweets related to this article.
- *More details on these impact measures can be found [here](#).

Score interpretations:

- Exceptional score (in top 0.01%).
- Substantial score (in top 1%).
- Average score (in bottom 99%).
- ? Score not available.

Main data sources:

- COVID-19 dataset⁽¹⁾ (list of papers)
- LitCovid hub⁽²⁾ (list of papers)
- PMC & PubMed (citations)
- Mendeley (number of readers)
- COVID-19-TweetDs⁽³⁾ (tweets)

Search...

Rank by:

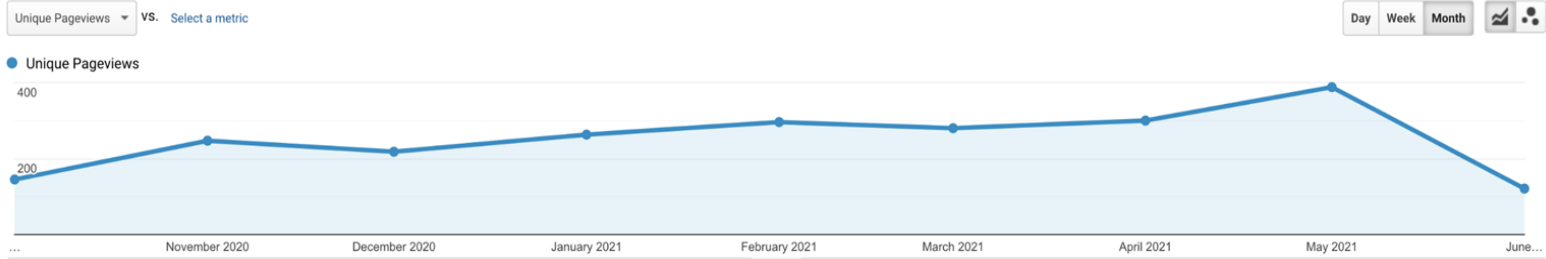
Popularity Influence Reader Attention Social Media Attention

Title	Venue	Year	Impact	Source
1 Safety and Efficacy of Imatinib for Hospitalized Adults with COVID-19: A structured summary of a study protocol for a randomised controlled trial	Trials	2020	🔥 📄 📌	LitCov and COVID-19
2 Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China	Lancet	2020	🔥 📄 📌	LitCov and COVID-19
3 Controlled, double-blind, randomized trial to assess the efficacy and safety of hydroxychloroquine and chloroquine in SARS-CoV-2 infection in healthcare personnel in the hospital setting: A structured summary of a study protocol for a randomised controlled trial	Trials	2020	🔥 📄 📌	LitCov and COVID-19
4 Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study	Lancet	2020	🔥 📄 📌	LitCov and COVID-19
5 Beyond the black stump: rapid reviews of health research issues affecting regional, rural and remote Australia	Med J Aust	2020	🔥 📄 📌	COVID-19
6 Epidemiological and clinical characteristics of coronavirus disease cases at a screening clinic during the early outbreak period: a single-center study	J Med Microbiol	2020	🔥 📄 📌	LitCov and COVID-19



BIP! Finder – Αξιολόγηση

BIP! Finder επισκεψιμότητα



Bip4Covid19

1^ο σε επισκέψεις covid-19 dataset στο zenodo

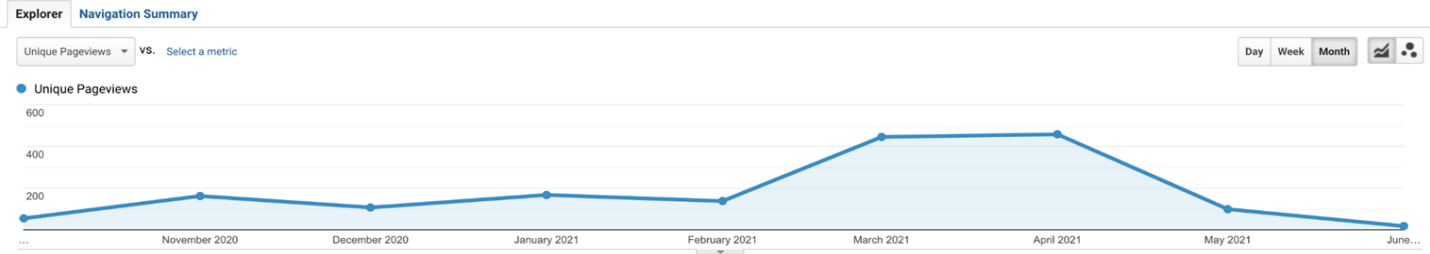
3^ο για covid-19 στο zenodo γενικά

113,970

views

15,044

downloads





Amnesia



Πλατφόρμα για την ανωνυμοποίησης ευαίσθητων
δεδομένων





Amnesia – Προδιαγραφές

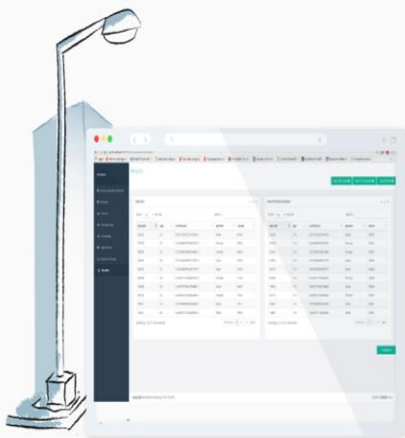
- Φόρτωση, επεξεργασία και αποθήκευση αρχείων ιεραρχιών
- k -ανωνυμοποίηση και k^m -ανωνυμοποίηση
- Εξερεύνηση όλων των συνδυασμών των ανωνυμοποιήσεων, ώστε να επιλέξει αυτή που του ταιριάζει
- Εμφάνιση των αρχικών και των ανωνυμοποιημένων δεδομένων, ώστε να παρατηρούμε τις αλλαγές που έχουν γίνει
- Αποθήκευση των ανωνυμοποιημένων δεδομένων, καθώς και των κανόνων ανωνυμοποίησης
- Εφαρμογή σε προσωπικό υπολογιστή, διαδικτυακή εφαρμογή και μέσω γραμμής εντολών



Amnesia – Ανάπτυξη

Λειτουργία του Amnesia

Ανωνυμοποίηση Δεδομένων σε 3 απλά βήματα



1 Εισαγωγή Δεδομένων

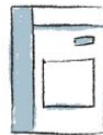
Το Amnesia δέχεται πολύπλοκα σχεσιακά δεδομένα αντικειμένων σε οριοθετημένα αρχεία κειμένου.

2 Επιλογή και Προεπισκόπηση δεδομένων για ανωνυμοποίηση

Οι οπτικές αναπαραστάσεις παραμέτρων και αποτελεσμάτων ανωνυμοποίησης επιτρέπουν σε μη ειδικούς χρήστες να προσαρμόσουν τη διαδικασία ανωνυμοποίησης στις ανάγκες τους.

3 Αποθήκευση των Ανωνυμοποιημένων Δεδομένων

Η διαδικασία ολοκληρώνεται χωρίς την έκθεση ευαίσθητων δεδομένων!



Υψηλή ευχρηστία και ευελιξία

- ✓ Εύκολη διεπαφή χρήσης
- ✓ Προσφέρει ειδικές ρυθμίσεις για τις διάφορες ανάγκες κάθε χρήστη
- ✓ Οπτικοποίηση των επιλογών ανωνυμοποίησης



Ανάλυση ευαίσθητων δεδομένων

- ✓ Εγγύεται μη-σύνδεση με τα αρχικά δεδομένα
- ✓ Προσφέρει k-anonymity & km-anonymity
- ✓ Επιτρέπει ελάχιστη μείωση της ποιότητας των πληροφοριών



Συμβατότητα με το GDPR

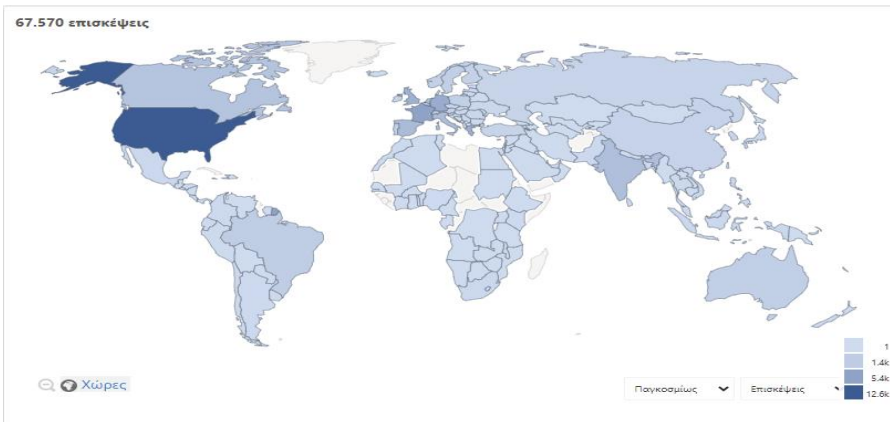
- ✓ Εγγύηση Ανωνυμίας
- ✓ Προσφέρει κάτι παρτάνω από την απλή ψευδο-ανωνυμοποίηση
- ✓ Τα ανώνυμα δεδομένα δεν περιορίζονται από το GDPR



Amnesia - Αξιολόγηση

Από το 2017-2021:

- ~67.000 χρήστες έχουν επισκεφτεί την σελίδα του Amnesia
- ~4.000 χρήστες έχουν κατεβάσει το Amnesia στον προσωπικό τους υπολογιστή
- ~40.000 χρήστες έχουν χρησιμοποιήσει την διαδικτυακή εφαρμογή του Amnesia



Πιλοτικές Εφαρμογές – Demos



BIP! Finder

Ηλίας Κανέλλος



Μανώλης Τερροβίτης



Δανάη Πλα Καρύδη





VIRTUAL EVENT

Ημέρες Δράσης

23-24-25
ΙΟΥΝΙΟΥ
2021

The logo features a central purple hexagon with the text 'Ημέρες Δράσης' and dates. It is surrounded by several teal hexagonal icons: a brain with gears, a lightbulb, an open book, a circuit board, a globe, and a graduation cap. The background is a dark blue grid with glowing blue dots and white lines forming a network pattern.

Thank you!