



ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ  
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ  
ΤΜΗΜΑΤΟΣ ΕΦΑΡΜΟΣΜΕΝΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ

---

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ  
με θέμα

*Σχεδίαση και Ανάπτυξη Διαδραστικής Διεπαφής Χρήστη για την  
Παρουσίαση Αποτελεσμάτων Εξόρυξης Γνώμης*

ΜΟΥΣΕΝΙΚΑΣ ΔΗΜΗΤΡΙΟΣ

ΘΕΣΣΑΛΟΝΙΚΗ, ΦΕΒΡΟΥΑΡΙΟΣ 2021

- Στόχος αυτοματοποιημένης ανάλυσης συναισθημάτων (sentiment analysis), ή εξόρυξης γνώμης (opinion mining).
- Οπτικοποίηση αναλύσεων συναισθημάτων και στοιχείων.
- Θεματικά πεδία οπτικοποίησης συναισθημάτων.
- Η επίλυση του προβλήματος είναι σημαντική:
  - για τις επιχειρήσεις και τους οργανισμούς.
  - για νέες ευκαιρίες και προκλήσεις

# Σκοπός – Στόχοι (1/2)

- Ανάλυση συναισθημάτων και αναπαράσταση των αποτελεσμάτων.
- Βασικοί στόχοι:
  - **Ευρεία βιβλιογραφική ανασκόπηση του θεματικού πεδίου ανάλυσης και οπτικοποίησης συναισθημάτων:**
    - *Παρουσίαση και επεξήγηση των πλέον χαρακτηριστικών τρόπων.*
    - *Παρουσίαση και λεπτομερής ανασκόπηση και κατηγοριοποίηση των τεχνικών οπτικοποίησης συναισθημάτων.*
    - *Καταγραφή εργαλείων και εφαρμογών οπτικοποίησης (εμπορικών και ανοιχτού κώδικα – open source).*

- **Μελέτη περίπτωσης (Case Study):**
  - *Ανάπτυξη διαδραστικής διεπαφής.*
  - *Χρήση συνόλου δεδομένων (data set) με κριτικές και αξιολογήσεις πελατών για προϊόντα και υπηρεσίες*
  - *Η διεπαφή είναι προσπελάσιμη και απεικονίζει τα γραφικά μέσω περιηγητή ιστοσελίδων (web browser).*
  - *Απεικόνιση πολλαπλών γραφικών και πληροφοριών.*
  - *Η εφαρμογή πολλαπλών σελίδων (multi page application).*

- Οφέλη παρούσας εργασίας:
    - *Ερευνητές οπτικοποίησης συναισθημάτων.*
    - *Ερευνητές άλλων επιστημονικών πεδίων αλλά και για επαγγελματίες.*
  - Περιοχές συνεισφοράς παρούσας εργασίας:
    - *Ευρεία βιβλιογραφική ανασκόπηση του θεματικού πεδίου ανάλυσης και οπτικοποίησης συναισθημάτων και έρευνας.*
    - *Επισκόπηση ενός ευρύτατου συνόλου τεχνικών οπτικοποίησης συναισθημάτων.*
    - *Λεπτομερής κατηγοριοποίηση και ανάλυση των χρησιμοποιούμενων τεχνικών οπτικοποίησης συναισθημάτων.*
    - *Πρακτικές υλοποίησης οπτικοποίησης ανάλυσης συναισθημάτων.*
    - *Προτάσεις για μελλοντικές ερευνητικές κατευθύνσεις.*
-

- Ανάλυση συναισθημάτων ή εξόρυξη γνώμης:
  - *Η ανάλυση δεδομένων κειμένου που μπορεί να καταδείξει την «πολικότητα» των συναισθημάτων που εμπεριέχονται.*
  - *Εναλλακτικά, θα μπορούσε κανείς να ορίσει την ανάλυση συναισθημάτων ως μια διαδικασία εντοπισμού και υπολογιστικής ανάλυσης απόψεων και κρίσεων.*
- Οπτικοποίηση συναισθημάτων (Sentiment Visualization):
  - *Οι τρόποι και οι τεχνικές, μέσα από τις οποίες μπορούν να απεικονιστούν και να αναλυθούν οι πληροφορίες ανάλυσης συναισθημάτων.*

# Τύποι Ανάλυσης Συναισθημάτων

- Λεπτομερής Ανάλυση Συναισθήματος (Fine-grained Sentiment Analysis)
- Ανάλυση Ανίχνευσης Συναισθήματος (Emotion Detection Sentiment Analysis)
- Μονοδιάστατη Ανάλυση Συναισθήματος (Aspect-based Sentiment Analysis)
- Ανάλυση Προθέσεων (Intent Analysis)

- Τεχνικές Ανάλυσης Συναισθήματος που Στηρίζονται σε Κανόνες (Rule-based)
- Τεχνικές Μηχανικής Μάθησης
- Υβριδικές Τεχνικές



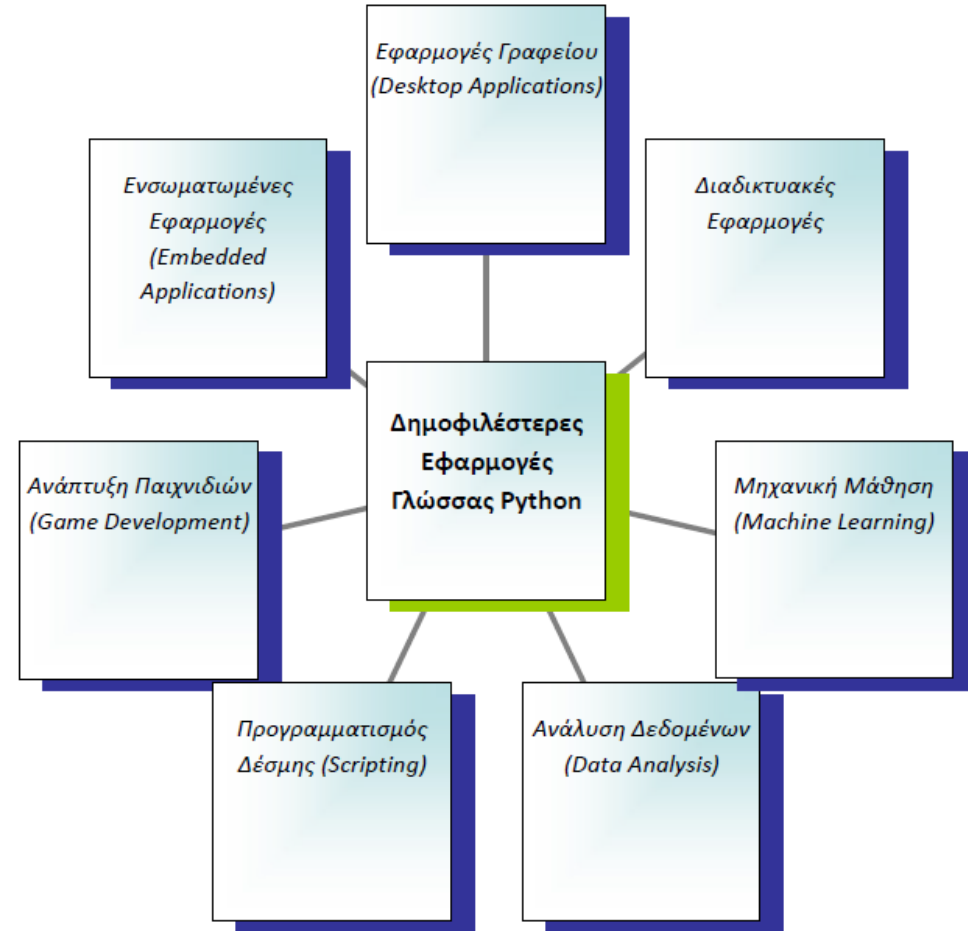
- Μελέτη ανταγωνισμού επιχειρήσεων
- Προσδιορισμός υποστηρικτών και επικριτών
- Παρακολούθηση συναισθημάτων κοινού σε βάθος χρόνου
- Ανάλυση αναφορών επωνυμίας επιχειρήσεων
- Διεξαγωγή ερευνών αγοράς

- Οπτικοποίηση συναισθημάτων ως υποσύνολο οπτικοποίησης κειμένου.
- Τεχνικές οπτικοποίησης συναισθημάτων από βιβλιογραφία:
  - *Τεχνικές με βάση τα χαρακτηριστικά των δεδομένων*
  - *Τεχνικές με βάση την πηγή προέλευσης και ιδιότητες των δεδομένων*
  - *Τεχνικές με βάση τις εκτελούμενες εργασίες*
  - *Τεχνικές με βάση τα χαρακτηριστικά οπτικοποίησης*

# Εργαλεία και Λογισμικό Οπτικοποίησης Ανάλυσης Συναισθημάτων

- Apache OpenNLP
  - Aylien
  - Bismart Folksonomy
  - GATE - Java NLP tools
  - Google Cloud NLP/ Google Fusion Tables
  - IBM Watson
  - Infogram
  - MeaningCloud
  - Microsoft Cognitive Services
  - MonkeyLearn
  - Open text
  - Opinion Crawl
  - Orange
  - QDA Miner
  - Qlik
  - Quick Search
  - RapidMiner
  - Semantria/ Lexanalytics
  - SAS Sentiment Analysis
  - Salience
  - Tableau
  - Texminer
  - Thematic
  - Trackur
  - Voyant Tools
-

# Εργαλεία Ανάπτυξης και Δημοφιλέστερες Εφαρμογές της Python



# Μελέτη Περίπτωσης: Ανάπτυξη Διαδραστικής Διεπαφής Παρουσίασης Αποτελεσμάτων Ανάλυσης Συναισθημάτων

- Περιλαμβάνει:
  - την προετοιμασία, ανάλυση και παρουσίαση ενός συνόλου δεδομένων κειμένου.
  - την παρουσίαση των αποτελεσμάτων ανάλυσης συναισθημάτων.
- Χρήση γλώσσας Python και βιβλιοθηκών γραφικών όπως είναι η Plotly και η Dash.
- Αξιοποίηση λειτουργιών από την χρήση τεχνικών οπτικοποίησης συναισθημάτων.

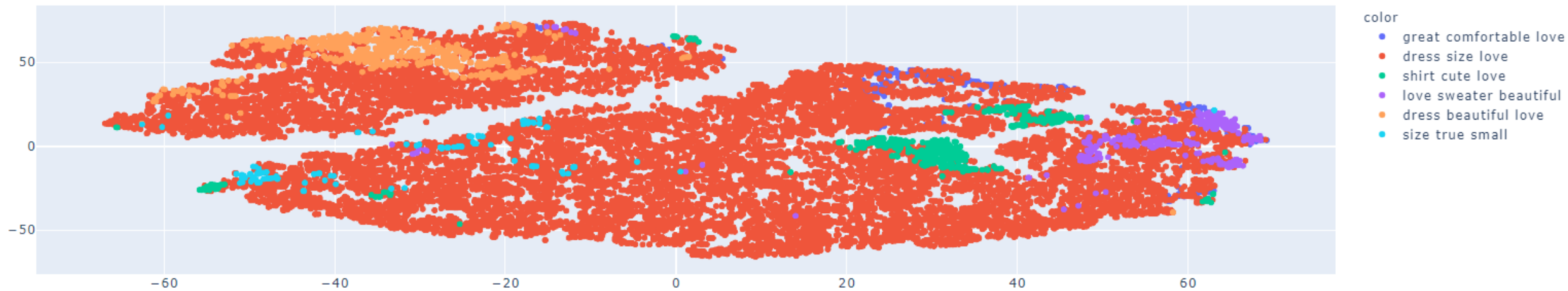
Οι λογικές ενότητες/σελίδες της web-εφαρμογής είναι οι ακόλουθες:

- Ενότητα Α – Data Sample Basics
    - *Εισαγωγική σελίδα, η οποία περιλαμβάνει βασικές πληροφορίες για το σύνολο δεδομένων (data set) και σχετικό δείγμα.*
  - Ενότητα Β - Visualizations
    - *Ομάδα ραβδογραμμάτων (bar charts), που οπτικοποιούν τη στατιστική ανάλυση των κριτικών και την κατανομή τους βάσει διαφορετικών κριτηρίων.*
  - Ενότητα Γ – Box Plots
    - *Γραφήματα τύπου box plot, που απεικονίζουν με ιδιαίτερα παραστατικό και διαδραστικό τρόπο τα συναισθήματα για τους επιμέρους τύπους ρούχων ή σκορ κριτικής, βάσει διαφορετικών κριτηρίων.*
  - Ενότητα Δ – Latent Analysis
    - *Γραφική απεικόνιση λανθάνουσας ανάλυσης (latent analysis), που παρουσιάζει τα έξι (6) δημοφιλέστερα θέματα συζήτησης (topics).*
-

# Αρχεία Εφαρμογής και Λειτουργικότητα

Αρχείο	Λειτουργικότητα
<code>index.py</code>	Συνδεδειγμένος κρίκος των επιμέρους σελίδων της εφαρμογής.
<code>app.py</code>	Διασφάλιση των κλήσεων των επιμέρους τμημάτων της εφαρμογής.
<code>E-Sales Reviews_v6.xlsx</code>	Σύνολο δεδομένων για ανάλυση.
<code>/assets/sentiment-analysis.png</code>	Εικόνα εικαστικής βελτίωση της εφαρμογής.
<code>/apps/__init__.py</code>	Απαραίτητο για την ορθή λειτουργία της εφαρμογής.
<code>/apps/sent_visual_v13a.py</code>	Απαιτούμενος Python κώδικας για την υλοποίηση της πρώτης σελίδας της εφαρμογής ( <i>Data Sample Basics</i> ).
<code>/apps/sent_visual_v13b.py</code>	Απαιτούμενος Python κώδικας για την υλοποίηση της δεύτερης σελίδας της εφαρμογής ( <i>Visualizations</i> ).
<code>/apps/sent_visual_v13c.py</code>	Απαιτούμενος Python κώδικας για την υλοποίηση της τρίτης σελίδας της εφαρμογής ( <i>Box Plots</i> ).
<code>/apps/sent_visual_v13d.py</code>	Απαιτούμενος Python κώδικας για την υλοποίηση της τέταρτης σελίδας της εφαρμογής ( <i>Latent Analysis</i> ).

t-SNE Clustering of 6 Latent Semantic Analysis Topics



*Scatter Chart with the six (6) most popular 3-word topics of the reviews are displayed based on latent analysis.*

CASE STUDY IN WHICH AN INTERACTIVE GRAPHICAL USER INTERFACE HAS BEEN DEVELOPED, THROUGH WHICH THE RESULTS OF SENTIMENT ANALYSIS ARE PRESENTED, ANALYZED AND EVALUATED. FOR THIS PURPOSE, A DATA SET HAS BEEN USED, WHICH INCLUDES CUSTOMER REVIEWS FOR PRODUCTS AND SERVICES PROVIDED BY A WOMEN'S CLOTHING COMPANY. THE USER IS NAVIGATED TO THE APPLICATION PAGES BY SELECTING PAGE FROM THE UPPER RIGHT PART NAMED "EXPLORE".



# Τεχνικές Οπτικοποίησης Ανάλυσης Συναισθημάτων Βιβλιογραφικής Έρευνας και Υλοποίησης (1/3)

- Τεχνική ανάλυσης και οπτικοποίησης συναισθημάτων επί κριτικών/αναφορών:
  - Ερευνητικά πρωτοτύπων/εργαλείων, όπως είναι τα *Affect Inspector* (Subasic & Huettner, 2001), *Pulse* (Gamon et al., 2005), *Opinion Observer* (Liu et al., 2005), *AMAZING* (Miao et al., 2009).
- Τεχνική οπτικοποίησης ανάλυσης συναισθημάτων πάνω σε ροές δεδομένων (streams):
  - Ερευνητικές εργασίες των Marcus και συν. (2011), Zhao και συν. (2012), Liu και συν. (2013), Kranjc και συν. (2015), Steed και συν. (2015) και Tsirakis και συν. (2016)
- Τεχνική ανάλυσης πολικότητας (polarity analysis):
  - Προτάθηκαν στα υλοποιηθέντα πρωτότυπα/εργαλεία *SATISFI* (Taskaya & Ahmad, 2003), *Vox Civitas* (Diakopoulos et al., 2010), καθώς επίσης και αυτές των Annett και Kondrak (2008), Puri και συν. (2014), Brooks και συν. (2014) και Gamon και συν. (2008).
- Τεχνική μονοδιάστατης (aspect based) ανάλυσης συναισθήματος, η οποία στη βιβλιογραφία συχνά έχει εφαρμοσθεί πάνω σε δεδομένα κριτικών πελατών:
  - Αντίστοιχες δουλειές στη βιβλιογραφία έχουν παρουσιαστεί από τους Yatani και συν. (2011), Alper και συν. (2011), Di Caro και Grella (2013), Gorg και συν. (2013), Wensel και Sood (2008), Cervantes και συν. (2015) και Mahmud και συν. (2016).

# Τεχνικές Οπτικοποίησης Ανάλυσης Συναισθημάτων Βιβλιογραφικής Έρευνας και Υλοποίησης (2/3)

- Τεχνικές που στηρίζονται σε εργασίες αναπαράστασης και αλληλεπίδρασης, οι οποίες υποστηρίζονται απευθείας από τεχνικές απεικόνισης. Τεχνικές ομαδοποίησης/κατηγοριοποίησης (clustering/classification), οι οποίες κάνουν χρήση ετικετών (tagging):
  - Προτάθηκαν από τους *Oelke και συν. (2009)*, *Brew και συν. (2011)*, *Kim και Lee (2014)* και *Sung και συν. (2016)*.
- Τεχνικές συγκεντρωτικής απεικόνισης (overview), οι οποίες επιτυγχάνουν να απεικονίσουν τη «μεγάλη εικόνα», οπτικοποιώντας μεγάλο αριθμό δεδομένων ή μέσω συγκεντρωτικών (aggregated) αναπαραστάσεων:
  - Προτάθηκαν στις δουλειές των *Wu και συν. (2013)*, *Liu και συν. (2016)*, *Miao και συν. (2009)* και *Scharl και συν. (2016)*.
- Τεχνικές πλοήγησης/εξερεύνησης (navigation/exploration), αφού έχει υλοποιηθεί εφαρμογή διαδραστικής πλοήγησης και εξερεύνησης των δεδομένων:
  - Μελέτες των *Duan και συν. (2012)*, *Brooks και συν. (2014)*, *Li και συν. (2016)*, *Lu και συν. (2016)*, *Hao και συν. (2013)*, *Torkildson και συν. (2014)*, και *Chen και συν. (2015)*.

- Τεχνικές με βάση τα χαρακτηριστικά οπτικοποίησης (visualization aspects) και τον τύπο των χρησιμοποιούμενων γραφικών:
  - ❑ Ειδικότερα, με βάση τις ιδιότητες των γραφικών (visual variables), χρησιμοποιείται η πολύ συνήθης πρακτική απεικόνισης της πολικότητας συναισθημάτων με χρωματική διαβάθμιση:
    - *Ερευνητικές εργασίες των Wensel και Sood (2008), Tsirakis και συν. (2016), Wu και συν. (2012; 2013), Calderon και συν. (2014), Chen και συν. (2014), Hoque και Carenini (2014), Hoque και Carenini (2016), Yu και συν. (2016), Kang και Ren (2011), Zhao και συν. (2014a), El-Assady και συν. (2016), Almutairi (2013) και Wang και συν. (2015).*
  - ❑ Επιπρόσθετα, το μέγεθος ή το εύρος των σχηματικών αναπαραστάσεων χρησιμοποιούνται ευρύτατα ως απεικονιστικές τεχνικές για να γίνει εύκολα αντιληπτή η κατανόηση του μεγέθους ή της ποσότητας των απεικονιζόμενων μεταβλητών, κυρίως μέσα από ραβδογράμματα ή άλλα αντίστοιχα γραφήματα:
    - *Σχετική βιβλιογραφία Alper et al., 2011; Mohammad et al., 2016; Marcus et al., 2011; Das et al., 2012; Liu et al., 2013; Kranjc et al., 2015; Kempter et al., 2014; Munezero et al., 2015; Gamon et al., 2008; Wecker et al., 2014*
  - ❑ Τέλος, η οπτικοποίηση με χρήση γλύφων, που χρησιμοποιούνται στην παρούσα υλοποίηση:
    - *Σχετική βιβλιογραφία (Liu et al., 2003; Gobron et al., 2010; Lee et al., 2010; Cresci et al., 2014; Li & Ren, 2009; Almutairi, 2013; Kempter et al. 2014; Munezero et al., 2015).*

- Ανάλυση και οπτικοποίηση συναισθημάτων.
- Ευρεία βιβλιογραφική ανασκόπηση θεματικού πεδίου.
- Λεπτομερής κατηγοριοποίηση τεχνικών οπτικοποίησης συναισθημάτων.
- Αναλυτική καταγραφή εργαλείων και εφαρμογών (εμπορικών και ανοιχτού κώδικα – open source).
- Μελέτη περίπτωσης (Case Study), με την ανάπτυξη διαδραστικής διεπαφής ανάλυσης και οπτικοποίησης συναισθημάτων.

# Όρια και Περιορισμοί της Έρευνας – Μελλοντικές Επεκτάσεις

- Εστίαση στην έρευνα και στις εφαρμογές οπτικοποίησης της ανάλυσης συναισθημάτων.
- Όρια και περιορισμοί από χρησιμοποιηθέντα εργαλεία έρευνας:
  - *Βιβλιογραφική αναζήτηση*
  - *Εργαλεία ανάπτυξης σχετικών εφαρμογών*
- Μελλοντική επέκταση της παρούσας έρευνας:
  - *Περαιτέρω εμβάθυνση στη θεωρητική θεμελίωση της ανάλυσης και οπτικοποίησης συναισθημάτων.*
  - *Περαιτέρω επέκταση της παρούσας υλοποίησης σε επόμενες εκδόσεις.*
  - *Χρήση επιπλέον γραφικών βιβλιοθηκών της Python.*

# Ευχαριστώ Πολύ για τον Χρόνο σας!

Υπάρχουν παρακαλώ ερωτήσεις;;;