



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΛΟΠΟΝΝΗΣΟΥ
ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΑΣ, ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ



Ερευνητική Ομάδα Γνώσης και Αβεβαιότητας

Πτυχιακή εργασία

Ολοκληρωμένο περιβάλλον ανάπτυξης ασαφούς tag cloud

Ευγενία - Ευαγγελία Ξάνδρη
2024201000010

Επιβλέπων:

Μανόλης Γουάλλες
Επίκουρος Καθηγητής

Τρίπολη, Ιούνιος 2016

Εγκρίθηκε από την εξεταστική επιτροπή την 22η Ιουνίου 2016.

Εμμανουήλ Γουάλλες
Επίκουρος Καθηγητής

Κωνσταντίνος Πέππας
Λέκτορας

.....

Ξάνδρη Ευγενία - Ευαγγελία
Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστήμιο Πελοποννήσου

Copyright © Ξάνδρη Ευγενία - Ευαγγελία, 2016
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευτεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Πανεπιστημίου Πελοποννήσου.

Η παρούσα εργασία εκπονήθηκε σε συνεργασία με την Ερευνητική Ομάδα Γνώσης και Αβεβαιότητας (ΓΑΒ LAB)

Περίληψη

Τα tag clouds είναι οπτικοποιήσεις λέξεων σε συνδυασμό με τη βαρύτητα ή τη σημασία τους. Σε ποικίλα σχήματα, χρώματα και μεγέθη σχηματίζονται tag clouds που αναφέρονται σε πρόσωπα, κείμενα κ.α. Σύμφωνα με το αντικείμενο στο οποίο αναφέρονται μεταβάλλονται τα παραπάνω χαρακτηριστικά τους και σχηματίζεται κάθε φορά διαφορετικό tag cloud. Σε αυτή την οπτικοποίηση εύκολα μπορεί να προστεθεί, ως βασικό χαρακτηριστικό, και η έννοια της αβεβαιότητας ορίζοντας αντίστοιχα το βαθμό διαφάνειας της κάθε λέξης.

Για τη δημιουργία των tag cloud υπάρχουν πολυάριθμα εργαλεία καθώς και αρκετές βιβλιοθήκες ανοιχτού κώδικα. Η παρούσα εργασία αφορά την ανάπτυξη μίας εφαρμογής που θα βασίζεται σε υπάρχουσα βιβλιοθήκη ανοιχτού κώδικα και θα επιτρέπει με φιλικό προς το χρήστη γραφικό περιβάλλον τον ορισμό των παραμέτρων του tag cloud, περιλαμβανομένων στοιχείων αβεβαιότητας, και στη συνέχεια τη δημιουργία του.

Η εφαρμογή αυτή θα επιτρέπει στον χρήστη να αξιοποιεί ένα απλό αρχείο txt με τα δεδομένα που επιθυμεί να οπτικοποιήσει, καθώς θα μπορεί να το εισάγει στην εφαρμογή και να το επεξεργαστεί μέσω αυτής. Ο χρήστης θα μπορεί να ορίσει συγκεκριμένα χαρακτηριστικά και τελικά να δημιουργήσει μία εικόνα με το επιθυμητό tag cloud.

Τέλος, περιγράφονται και αναλύονται όλα τα στάδια της δημιουργίας της εφαρμογής, από την μελέτη των λειτουργικών απαιτήσεων και την τεχνική υλοποίηση, μέχρι την παρουσίαση των συμπερασμάτων και των μελλοντικών βελτιώσεων.

Abstract

Tag clouds are visualizations of words in relation to the severity or importance. Tag clouds take the shape of various colors and sizes, which refer to persons, texts etc. According to the object, that tag clouds are referred to, there is a change in their characteristics, so each time a different tag cloud is created. In this visualization can easily be added as a standard feature, and the concept of uncertainty defining respectively the degree of transparency of each word.

There are numerous tools and several open source libraries to create tag clouds. This thesis involves the development of an application based on existing open source library and will allow for a GUI user friendly definition of tag cloud parameters including uncertainty information, and then create it.

This application will allow user to utilize a simple txt file with the data who wishes to visualize, as he is going to have the ability to introduce it in the application and process it, through this. The user can define specific characteristics and eventually create an image with the desired tag cloud.

Finally, all creation stages of the application are described and analyzed, from the study of the functional requirements and technical implementation, until the presentation of the conclusions and future improvements.

*Αφιερώνεται σε όσους με βοήθησαν
σε αυτή την εργασία.*

Περιεχόμενα

Ευχαριστίες	ix
1 Εισαγωγή	1
2 Tag Clouds	3
2.1 Μορφές ανάπτυξης	3
3 Βιβλιοθήκη Kumo	9
3.1 Βασικά στοιχεία	9
3.2 Εξετάζοντας τις παραμέτρους	12
4 Αβεβαιότητα	17
5 Σχεδίαση εφαρμογής	21
6 Υλοποίηση	25
6.1 Σύνδεση	27
6.2 Δομή αρχείου λέξεων	27
6.3 Μετατροπή σε jar	28
7 Δοκιμαστική λειτουργία	33
7.1 Σενάριο πρώτο	33
7.2 Σενάριο δεύτερο	36
8 Συμπεράσματα	39
8.1 Επεκτάσεις και βελτιώσεις	39

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον διδάσκοντα και επιβλέποντα της πτυχιακής μου εργασίας κ. Γουάλλες Εμμανουήλ για την ευκαιρία που μου έδωσε να εργαστώ πάνω σε αυτό το ενδιαφέρον θέμα, αλλά και για την πάντα άμεση και πολύτιμη καθοδήγησή του.

Ένα μεγάλο ευχαριστώ σε όλους όσους με στήριξαν και με βοήθησαν στην ολοκλήρωση αυτής της πτυχιακής εργασίας.

Και τέλος, θα ήθελα να εκφράσω την ευγνωμοσύνη μου προς την οικογένειά μου η οποία με στήριξε καθ' όλη τη διάρκεια των σπουδών μου.

Κεφάλαιο 1

Εισαγωγή

Το tag cloud είναι μια οπτική αναπαράσταση από δεδομένα ενός κειμένου ή με λέξεις-κλειδιά λαμβάνοντας υπ'οψην την συχνότητα εμφάνισης ή κάποιο άλλο μέτρο συγκρισης, που βοηθάει τον αναγνώστη με μια ματιά να αντιληφθεί το περιεχόμενό του [1]. Κύριο χαρακτηριστικό των λέξεων είναι το χρώμα και το μέγεθος των γραμμάτων, τα οποία ποικίλουν ανάλογα με την βαρύτητα της λέξης στο κείμενο το οποίο αναφέρονται. Μερικές φορές κατηγοριοποιούνται και αλφαβητικά. Αυτή η οπτικοποίηση της συχνότητας εμφάνισης κάποιας λέξης δίνει εύκολα και ευχάριστα συγκεντρωτική πληροφορία στον αναγνώστη που δύσκολα αποκτούσε με άλλον τρόπο ανάλυσης.

Η πληροφορία δεν είναι πάντα βέβαιη και συχνά αμφισβητείται η αναφορά της (η ύπαρξή της) σε μία συλλογή λέξεων. Η συμβατική προσέγγιση για τη δημιουργία tag cloud δεν συμπεριλαμβάνει την παράμετρο της αβεβαιότητας [2]. Στόχος της παρούσας εργασίας είναι η υλοποίηση μιας διεπαφής χρήστη που θα επιτρέπει με φιλικό γραφικό περιβάλλον τον ορισμό των παραμέτρων του tag cloud, περιλαμβανομένων στοιχείων αβεβαιότητας, και στη συνέχεια τη δημιουργία του. Χρησιμοποιώντας τον βαθμό της αβεβαιότητας στην δημιουργία ενός tag cloud απεικονίζονται λέξεις-κλειδιά για τις οποίες υπάρχει αμφιβολία αν αναφέρονται στο κείμενο το οποίο βασίζεται.

Συγκεκριμένα, αξιοποιείται η έτοιμη βιβλιοθήκη Kumo [3] της Java η οποία δέχεται τις λέξεις κλειδιά και τις απαραίτητες παραμέτρους (βάρος, γραμματοσειρά και χρώμα). Στην παρούσα εργασία αναπτύσσεται η επέκτασή της στην οποία γίνεται χρήση μιας επιπλέον παραμέτρου, της αβεβαιότητας [4], και δημιουργείται το αρχείο εικόνας χωρίς την ανάγκη για δημιουργία ενός applet, όπως συμβαίνει με πολλές άλλες βιβλιοθήκες [5] [6] [7] [8] [9].

Η δομή του υπόλοιπου κειμένου είναι ως ακολούθως: Στο κεφάλαιο 2 παρουσιάζουμε την ιστορία των tag clouds καθώς και τον τρόπο ανάπτυξής τους με την επιλογή των κατάλληλων χαρακτηριστικών για κάθε λέξη-κλειδί που τα αποτελεί. Συνεχίζοντας στο κεφάλαιο 3 εξηγούμε αναλυτικότερα τον τρόπο λειτουργίας την βιβλιοθήκης Kumo η οποία είναι το εργαλείο που μετατρέπει μία συσσωρευμένη πληροφορία στην μορφή των tag clouds και τις απαραίτητες παραμέτρους τις οποίες ορίζουμε κατά την εκτέλεση της. Στο κεφάλαιο 4 αναλύουμε εκτενέστερα την έννοια της αβεβαιότητας. Στο επόμενο βήμα στο κεφάλαιο 5 περιγράφουμε τον τρόπο με τον οποίο σκεφτήκαμε και υλοποιήσαμε τον σχεδιασμό της εφαρμογής μας και στο κεφάλαιο 6 αναλύουμε τον τρόπο με τον οποίο έγινε η σύνδεση της εφαρμογής με την βιβλιοθήκη Kumo και την διαδικασία μετρατοπής της εφαρμογής σε jar αρχείο. Στο κεφάλαιο 7 παραθέτουμε τα αποτελέσματα από την χρήση της διεπαφής που δημιουργήσαμε. Κλείνοντας στο κεφάλαιο 8 παραθέτουμε επιπλέον προτάσεις για περαιτέρω ανάπτυξη και επέκταση των δυνατοτήτων της εφαρμογής.

Κεφάλαιο 2

Tag Clouds

Τα tag clouds (σύννεφα ετικετών) χρησιμοποιήθηκαν για πρώτη φορά σε δικτυακό τόπο υψηλής κυκλοφορίας, το Flickr [10]. Ήταν μια διαφορετική φιλοσοφία για την οργάνωση πληροφοριών, όπου οι πιο συχνά χρησιμοποιούμενες ετικέτες εμφανίζονταν με μεγαλύτερη γραμματοσειρά [11]. Ο πιο συνηθισμένος τρόπος εμφάνισής τους ήταν με αλφαβητική σειρά. Η πρώτη δημοσίευση έγινε στη γερμανική γλώσσα. Στις 24 Μαρτίου 2009 το CNN ισχυρίστηκε πως δημιούργησε το μεγαλύτερο σύννεφο λέξεων για τη διεύθυνση του προέδρου Μπαράκ Ομπάμα [12].

Τα σύννεφα λέξεων κέρδισαν μεγάλη δημοτικότητα εξαιτίας του ρόλου τους στη βελτιστοποίηση των μηχανών αναζήτησης των ιστοσελίδων. Σε κάποιες περιπτώσεις το να απαντήσεις σε συγκεκριμένα ερωτήματα ψάχνοντας σε ένα tag cloud είναι προτιμότερο από το να κάνεις συμβατική αναζήτηση [13]. Όταν η πληροφορία που ψάχνουμε είναι ακριβής είναι πιο αποδοτική η αναζήτηση, ενώ στην περίπτωση των tag clouds προτιμούνται αν η αναζήτηση της πληροφορίας είναι πιο γενική. Κατάλληλα παραμετροποιημένα σύννεφα λέξεων βοηθούν τους δικτυακούς τόπους να βελτιώσουν την κατάταξή τους στον ευρετηριασμό των μηχανών αναζήτησης [14].

Επίσης χρησιμοποιούνται από ιστολόγια για να περιγράψουν το περιεχόμενο τους ή ακόμα και σε επιστημονικές δημοσιεύσεις. Πιο πρόσφατα όμως το tag cloud έχει αναδειχθεί ως ένας πολύ έξυπνος τρόπος οπτικοποίησης σημασιολογικής πληροφορίας [12]. Η χρήση τους επεκτάθηκε στην απεικόνιση ενός κειμένου με λέξεις ανάλογα με την συχνότητα εμφάνισής τους σε αυτό, σημαίνοντας ότι οι λέξεις που περιέχονται πολλές φορές εμφανίζονται μεγάλες και στη συνέχεια οι υπόλοιπες όλο και πιο μικρές.

2.1 Μορφές ανάπτυξης

Ένας τρόπος δημιουργίας των tag clouds βασίζεται στην γραμμική προσέγγιση της σχέσης 2.1 για τον υπολογισμό της βαρύτητας των λέξεων [15].

$$S_i = \left\lceil \frac{f_{max} * (t_i - t_{min})}{t_{max} - t_{min}} \right\rceil \quad (2.1)$$

Όπου:

S_i : μέγεθος γραμματοσειράς

f_{max} : μέγιστο μέγεθος γραμματοσειράς

t_i : συχνότητα

t_{min} : ελάχιστη συχνότητα

t_{max} : μέγιστη συχνότητα

Άλλος δημοφιλής τρόπος είναι αυτός της σχέσης 2.2, με τον οποίο υπολογίζεται ο λογάριθμος της χρήσης κάθε ετικέτας. Χρησιμοποιώντας την λογαριθμική τιμή αντί της αρχικής έχουμε σαν αποτέλεσμα μια ομαλή διαβάθμιση από τη λιγότερο στην περισσότερο χρησιμοποιούμενη ετικέτα [16].

$$l_i = \frac{\log u_i - \log \min \mathbf{u}}{\log \max \mathbf{u} - \log \min \mathbf{u}} \quad (2.2)$$

Όπου θεωρούμε τις ετικέτες το διάνυσμα $\mathbf{u} = [u_1, u_2, \dots, u_n]$

$\min \mathbf{u}$: ελάχιστη τιμή

$\max \mathbf{u}$: μέγιστη τιμή

l_i : κανονικοποιημένη τιμή για κάθε ετικέτα (0-1)

Υπάρχει ποικιλία τρόπων για την σχεδίαση και την υλοποίηση των tag clouds [17]. Οι πιο συνηθισμένοι μέθοδοι δίνουν μεγαλύτερη βαρύτητα στο μέγεθος των λέξεων και άλλοι στο συνδυασμό χρωμάτων, την γραμματοσειρά και τέλος στην τοποθέτηση των λέξεων. Ενδεικτικά:

1. Tag cloud με τις λέξεις-κλειδιά ταξινομημένες αλφαβητικά στο σχήμα 2.1.
2. Tag cloud με τις λέξεις-κλειδιά ταξινομημένες αλφαβητικά. Το μέγεθος της γραμματοσειράς αλλάζει ανάλογα με την βαρύτητα της κάθε ετικέτας στο σχήμα 2.2, ή προσθέτοντας άλλο χρώμα στην κάθε σημαντική λέξη στο σχήμα 2.3.
3. Tag cloud με ετικέτες ταξινομημένες ανάλογα με τη σημασία ή τη συχνότητα τους. Τόσο το μέγεθος της γραμματοσειράς όσο και τα χρώματα μπορούν να χρησιμοποιηθούν για να τονισθεί η σημασία των όρων στο σχήμα 2.4.
4. Tag cloud που δεν είναι ταξινομημένα αλλά περικλείονται σ' ένα συγκεκριμένο σχήμα. Οι λέξεις-κλειδιά τοποθετούνται διαφορετικά στο καθ' ένα από αυτά στα σχήματα 2.5 και 2.6.
5. Tag cloud με συγκεκριμένο σχήμα στο οποίο οι ετικέτες με το μεγαλύτερο βάρος τοποθετούνται στο κέντρο της εικόνας στο σχήμα 2.7.

Τέλος, οι περισσότεροι αλγόριθμοι, όπως η βιβλιοθήκη Kumο της Java πραγματοποιούν τυχαία τοποθέτηση των λέξεων στο φόντο της εικόνας με μόνη δυνατότητα αλλαγής του Collision mode, δηλαδή αν οι λέξεις θα τοποθετούνται η μία μέσα στην άλλη ή τελείως ξεχωριστά.



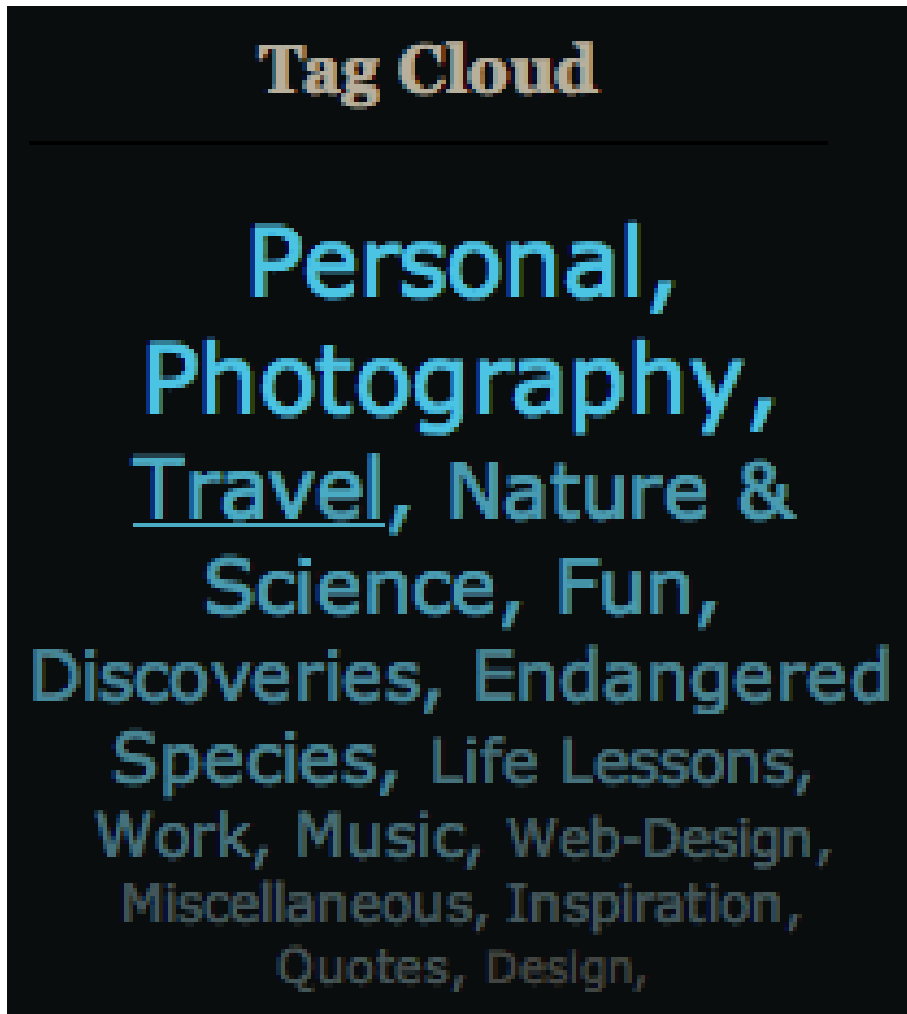
Σχήμα 2.1: Χωρίς διαφορετικά χαρακτηριστικά [18]



Σχήμα 2.2: Διαφορετικό μέγεθος [19]



Σχήμα 2.3: Διαφορετικό μέγεθος & χρώμα [20]



Σχήμα 2.4: Ταξινόμηση με βάση την βαρύτητα [18]



Σχήμα 2.5: Tag Cloud “Head” [21]



Σχήμα 2.6: Tag “Cloud” [22]

Κεφάλαιο 3

Βιβλιοθήκη Kumo

Η Kumo είναι μία δημοφιλής βιβλιοθήκη της Java που αναπτύχθηκε από τον Kenneth Cason [3]. Ο στόχος της είναι η δημιουργία μιας open source φιλικής προς το χρήστη Word Cloud βιβλιοθήκης. Κύριες δυνατότητες που παρέχει για την δημιουργία ενός tag cloud είναι οι ακόλουθες.

- Υποστηρίζει την αλληλοκάλυψη σε επίπεδο λέξεων και pixel
- Ευέλικτο μέγεθος εικόνας
- Ευέλικτο μέγεθος γραμματοσειράς
- Υποστηρίζει συγκεκριμένες ή τυχαίες παλέτες χρωμάτων.

Αρχικά, όπως παρατηρούμε στις εικόνες 3.1 και 3.2, κάθε λέξη αντιστοιχεί σε ένα νοητό ορθογώνιο όταν η αλληλοκάλυψη είναι σε επίπεδο λέξεων. Αντίθετα σε επίπεδο pixel υπάρχει αλληλοκάλυψη ανάλογα με το σχήμα, το μέγεθος και τη θέση του κάθε γράμματος σε κάθε λέξη όπως φαίνεται στις εικόνες 3.3 και 3.4.

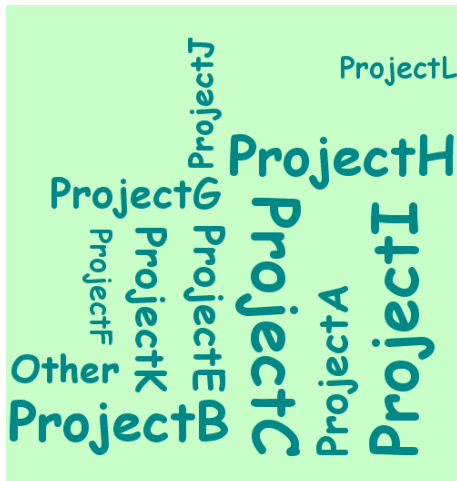
Στην βιβλιοθήκη αυτή ο χρήστης εύκολα μπορεί να μεταβάλλει το μέγεθος της εικόνας που θα παραχθεί. Ένα πιο μικρό Tag Cloud μπορεί να είναι αυτό της εικόνας 3.5.

Επιπλέον ο χρήστης μπορεί να αλλάζει και να προσαρμόζει κατάλληλα το μέγεθος της γραμματοσειράς.

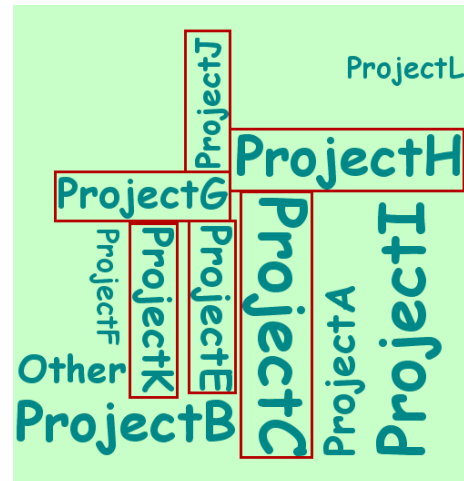
Τα χρώματα των λέξεων μπορούν να δημιουργηθούν τυχαία απο παλέτα χρωμάτων ή να οριστούν συγκεκριμένα απο τον χρήστη.

3.1 Βασικά στοιχεία

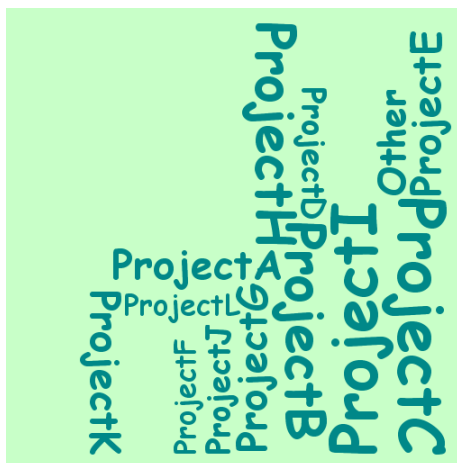
Τα πακέτα που αποτελούν την συγκεκριμένη βιβλιοθήκη είναι άμεσα διαθέσιμα [24] σε όλους τους χρήστες. Για να τα χρησιμοποιηθεί πρέπει να δημιουργηθεί ένα project με τα packages της Kumo. Για να λειτουργήσει το project είναι απαραίτητο να συμπεριληφθούν τα ακόλουθα jar αρχεία:



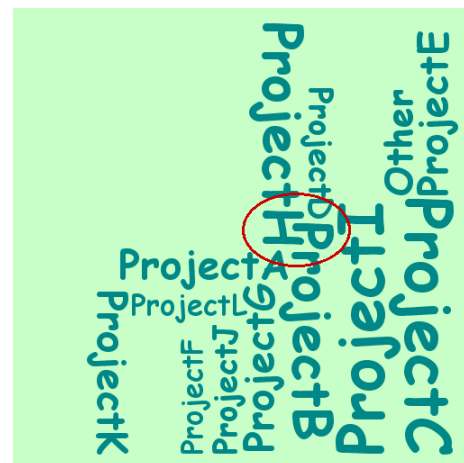
Σχήμα 3.1: Επίπεδο Λέξεων



Σχήμα 3.2: Επεξήγηση στο επίπεδο λέξεων



Σχήμα 3.3: Επίπεδο Pixel



Σχήμα 3.4: Επεξήγηση στο επίπεδο των pixel



Σχήμα 3.5: Μικρότερο μέγεθος tag cloud

- com.springsource.net.sf.cglib-2.2.0
- common-lang3
- commons-io-2.4
- hamcrest-all-1.3
- jsoup-1.8.1
- junit-4.3
- lambdaj-2.3.3
- lambdaj-2.3.3-sources
- log4j-1.2.17
- objenesis-2.1

Για παράδειγμα αν χρησιμοποιεί κάποιος NetBeans πρέπει για να τα προσθέσει στις βιβλιοθήκες (Libraries->Add Jars), αν χρησιμοποιεί κάποιο άλλο IDE θα πρέπει να ψάξει στο help για περισσότερες πληροφορίες.

Στην Kumo, από πιο ενδιαφέροντα πακέτα που εμπεριέχουν τις δυνατότητες τις οποίες περιγράψαμε είναι το collide, που περιέχει κλάσεις μέσω των οποίων υπολογίζει την θέση των λέξεων και τις τοποθετεί. Αμέσως επόμενο στοιχείο είναι το font και το font.scale. Στο font.scale υπάρχουν δύο κλάσεις με τους δύο τύπους γραμματοσειράς που μπορούν να χρησιμοποιηθούν, από τις οποίες μπορεί να αλλάξει και το μέγεθος των γραμμάτων. Σημαντικό είναι και το πακέτο palette. Με την κλάση ColorPalette και με τη δημιουργία αντικειμένων τύπου Color, που θέτουμε ως όρισμα, πραγματοποιούνται αλλαγές στα χρώματα των λέξεων και του φόντου με την βοήθεια των μεθόδων setColorPalette και setBackground αντίστοιχα.

Για να φτιάξουμε ένα καινούριο tag cloud πρέπει αρχικά να δημιουργήσουμε ένα αντικείμενο τύπου WordCloud καλώντας τη μέθοδο με ορίσματα τις διαστάσεις τις εικόνας και την επιλογή για το Collision Mode. Στη συνέχεια γίνεται κλήση της μεθόδου setBackground όπου επιλέγουμε το σχήμα του tag cloud με τις κατάλληλες διαστάσεις. Έπειτα ορίζουμε τα χρώματα χρησιμοποιώντας τις setColorPalette(new ColorPalette(colorLetter) και setBackgroundColor(colorBackgr), όπου colorLetter και colorBackgr μεταβλητές τύπου Color. Όμοια με την γραμματοσειρά επιλέγουμε ως όρισμα *SqrtFontScalar(10, 40)* ή *LinearFontScalar(10, 70)* στην μέθοδο setFontScalar. Καταλήγουμε την μέθοδο *writeToFile("new.png")*, για να αποθηκεύσουμε την png εικόνα με όνομα new στο φάκελο που υπάρχει το project. Συνδυάζοντας τις εντολές αυτές προκύπτει:

```
public static void main(String[] args){
    WordCloud wordCloud = new WordCloud(1024, 1024, CollisionMode.RECTANGLE);
    java.util.List<WordFrequency> wordFrequencies;
    wordFrequencies = frequencyAnalyzer.load(getInputStream("text/datarank.txt"));
    wordCloud.setBackground(new RectangleBackground(500,500));
    wordCloud.setColorPalette(new ColorPalette(new Color(0x008888)));
    wordCloud.setBackgroundColor(new Color (0xc8ffc8));
    wordCloud.setFontScalar(new SqrtFontScalar(10, 70));
    wordCloud.build(wordFrequencies);
}
```


Κεφάλαιο 4

Αβεβαιότητα

Θεωρούμε πολύ σημαντικό ξεκινώντας αυτό το κεφάλαιο να δωθεί η γενική ιδέα της αβεβαιότητας. Είναι ένα έμφυτο χαρακτηριστικό του ανθρώπου, το οποίο έχουμε συνηθίσει να αντιμετωπίζουμε σε οποιαδήποτε πληροφορία συναντάμε. Αυτό συμβαίνει τόσο συχνά που η ύπαρξή της είναι ανεπαίσθητη. Το πιο σημαντικό είναι ότι συνήθως εμείς θεωρούμε ότι ο όρος “αβεβαιότητα” αναφέρεται σε οποιαδήποτε κατάσταση στην οποία κάτι δεν είναι σίγουρα βέβαιο και ακριβές, χωρίς να συνειδητοποιούμε ότι πολλοί ανομοιογενείς τύποι πληροφοριών περιλαμβάνονται σε αυτή την κατηγορία [2]. Παρακάτω θα αναλύσουμε πιο συγκεκριμένα κάποια είδη της αβεβαιότητας.

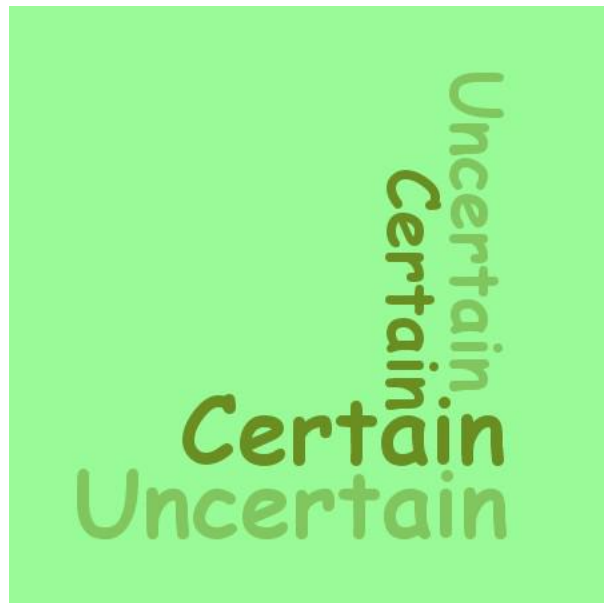
Πιθανοτική πληροφορία (Probabilistic information). είναι όταν δεν είναι βέβαιο τότε ένα ενδεχόμενο θα συμβεί, αλλά η ακριβής πιθανότητα γι' αυτό είναι γνωστή. Σ' αυτήν την περίπτωση η πιθανότητα είναι συγκεκριμένη δηλαδή δεν υπάρχει αβεβαιότητα ως προς αυτήν, αλλά μόνο στον χρόνο. Ένα απλό παράδειγμα αυτής της κατηγορίας είναι η πιθανότητα να τραβήξεις, τυχαία, ντάμα σε μία τράπουλα, η οποία είναι ακριβώς 1/13. Ευρέως γνωστό παράδειγμα είναι και η πιθανότητα να τύχει κάποιος ένα συγκεκριμένο αριθμό στο ζάρι.

Δυνατοτική πληροφορία (Possibilistic information). Ομοίως με την παραπάνω περίπτωση, δεν είναι βέβαιο τότε θα συμβεί το ενδεχόμενο. Επιπρόσθετα, όμως, ούτε η πιθανότητα είναι γνωστή. Μόνο κάποια άνω (αληθοφάνεια) και κάτω (αναγκαιότητα) όρια είναι γνωστά. Ένα παράδειγμα μπορεί να είναι οι πιθανότητες να κερδίσει κάποιος σε ένα ραδιοφωνικό διαγωνισμό με κλήρωση, όταν δεν ξέρουμε πόσοι είναι οι συμμετέχοντες.

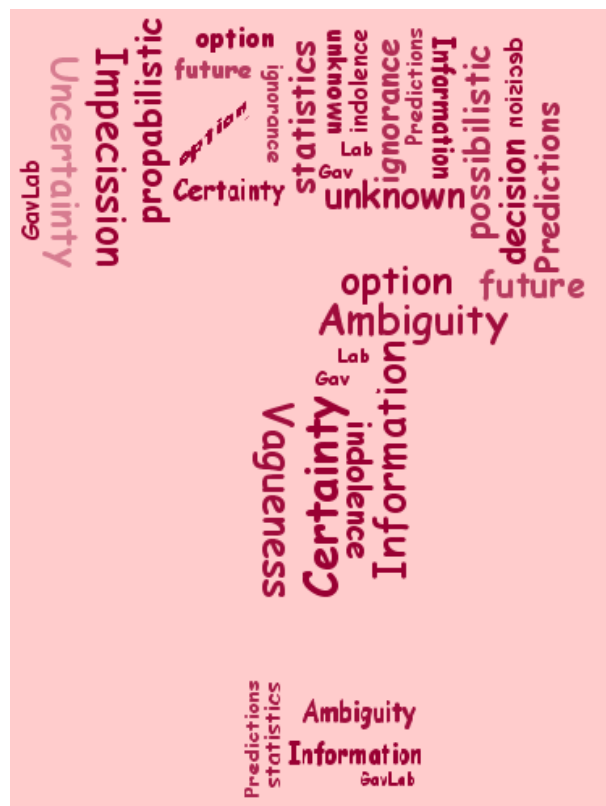
Αμφισημία (Ambiguity). Όταν η καταγωγή της αβεβαιότητας δεν βρίσκεται εντός των δεδομένων αλλά μάλλον εντός της ερμηνείας τους. Αυτό είναι πιο σύνηθες όταν χρησιμοποιούμε λέξεις που αντιπροσωπεύουν τα δεδομένα. Αναλυτικότερα, δημιουργείται ασάφεια όταν (ο ορισμός μίας λέξης) δεν είναι απόλυτα σαφές(σαφής) τι σημαίνει μία λέξη. Αυτό συμβαίνει όταν η λέξη έχει δύο έννοιες και δεν ξέρουμε ποια από τις δύο είναι επικρατέστερη. Δηλαδή, όταν δεν είναι σαφές αν το φύλο αναφέρεται στο sex(άνδρας - γυναίκα) ή το leaf(φύλλο δέντρου).

Ανακρίβεια (Imprecision). Σε αυτήν την κατηγορία, συνήθως προέρχονται από την πεπερασμένη ακρίβεια ενός εργαλείου μέτρησης. Συγκεκριμένα όταν θέλουμε να μετρήσουμε το ακριβές βάρος ενός ατόμου και δεν ξέρουμε αν η μέτρηση είναι ακριβής. Όμοια, αν χρησιμοποιούμε ένα ψηφιακό θερμόμετρο που οι ενδείξεις του είναι με ένα δεκαδικό ψηφίο. Συνεπώς, ο πυρετός μπορεί να είναι πάνω ή κάτω από την ένδειξη 37.2.

Ασάφεια (Vagueness). Όταν τα περιθώρια της έννοιας ενός όρου δεν είναι ορισμένα με ακρίβεια και διεθνώς. Εδώ οι λέξεις δεν έχουν δύο έννοιες, ωστόσο υπάρχει ασάφεια ως προς το τι ακριβώς σημαίνει. Σκεφτείται την έννοια του όρου “ψηλός άνθρωπος” και την δυσκολία στο να κατηγοριοποιήσεις κάποιους ανθρώπους ως μέλη ή όχι των ψηλών ανθρώπων (ή του όρου ψη-



Σχήμα 4.1: Uncertainty



λοί άνθρωποι). Ένα ακόμα παράδειγμα που μπορούμε να αναφέρουμε είναι όταν χαρακτηρίζουμε την θερμοκρασία “υψηλή”. Σε αυτήν την περίπτωση δεν γίνεται αναφορά των βαθμών, οπότε δεν γνωρίζουμε ακριβώς πόσο “υψηλή”.

Option uncertainty είναι ένα ενδημικό χαρακτηριστικό αποφάσεων που μπορούμε να προβλέψουμε τις συνέπειες των ενεργειών μας με κάθε λεπτομέρεια (ή αλλιώς να είμαστε σίγουροι για το ποιες είναι οι επιλογές μας). Παρόλο ότι σε πολλές περιπτώσεις αποφάσεων δεν έχει τόσο μεγάλη σημασία ποια είναι ακριβώς η συνέπεια της κάθε πράξης, σε μερικές οι λεπτομέρειες έχουν πολύ μεγάλη σημασία. Ένα παράδειγμα είναι όταν επιλέγουμε μεταξύ διάφορων πακέτων διακοπών η ικανοποίησή μας εξαρτάται από κάποιες λεπτομέρειες όπως το αν θα πάθουμε τροφική δηλητηρίαση στο ξενοδοχείο, ή το αν θα υπάρχει μια θορυβώδης κατασκευή έξω από το δωμάτιο. Όπως επίσης δεν μπορούμε να είμαστε σίγουροι ότι θα παραμείνουμε στεγνοί σε μία βροχή, ακόμα και αν πάρουμε ομπρέλα. Αυτό γιατί μπορεί να υπάρξει δυνατός αέρας και να μας καταστρέψει την ομπρέλα. Έτσι το να καταλάβουμε το πως να χειριστούμε το option uncertainty είναι πολύ σημαντικό [26].

Κεφάλαιο 5

Σχεδίαση εφαρμογής

Ξεκινήσαμε με την ιδέα να δημιουργήσουμε μία διεπαφή χρήστη που θα παράγει tag clouds με την βοήθεια της Kumo, έχοντας προσαρμόσει και την παράμετρο της αβεβαιότητας. Για την δημιουργία της εφαρμογής προσαρμόσαμε τις δυνατότητες της βιβλιοθήκης, αναλύσαμε τις απαιτήσεις των χρηστών και καταγράψαμε τις εξής λειτουργικές προδιαγραφές [27]:

A/A	Προδιαγραφή	Δεδομένα εισόδου	Επίπεδο Αναγκαιότητας	Αποτέλεσμα
1	Λειτουργία δημιουργίας Tag Cloud	λέξεις από αρχείο και ορισμός παραμέτρων	Απαραίτητη	δημιουργία εικόνας νέου Tag cloud
2	Λειτουργία άνοιγμα-επεξεργασία λέξεων	αρχείο txt, εισαγωγή σε πίνακα	Απαραίτητη	πίνακας λέξεων
3	Λειτουργία χρώματος background	παλέτα διαθέσιμων χρωμάτων	Απαραίτητη	ορισμός του χρώματος για το background
4	Λειτουργία χρώματος γραμμάτων	παλέτα διαθέσιμων χρωμάτων	Απαραίτητη	ορισμός του χρώματος για τα χρώματα
5	Λειτουργία επιλογής σχήματος	σχήμα επιλογής	Απαραίτητη	ορισμός σχήματος
6	Λειτουργία επιλογής γραμματοσειράς	επιλογή γραμματοσειράς	Απαραίτητη	ορισμός γραμματοσειράς
7	Λειτουργία collision mode	διαθέσιμες επιλογές	Απαραίτητη	ορισμός επιλογής collision mode
8	Λειτουργία επιλογής διαστάσεων εικόνας	διαστάσεις εικόνας	Απαραίτητη	ορισμός μεγέθους εικόνας
9	Λειτουργία προσθήκης λέξεων στην λίστα	λέξη και χαρακτηριστικά	Υψηλή	εισαγωγή στοιχείων σε πίνακα
10	Λειτουργία αποθήκευσης λίστας λέξεων	νέα λίστα δεδομένων	Υψηλή	αποθήκευση σε αρχείο txt
11	Λειτουργία αποθήκευσης εικόνας σε συγκεκριμένο path	νέα εικόνα	Επιθυμητή	δημιουργία εικόνας png

Αναλύοντας τα δεδομένα που συλλέξαμε και προσπαθώντας να σχεδιάσουμε την εφαρμογή αποφασίσαμε να προσφέρει τα ακόλουθα: Ο χρήστης θα μπορεί να χρησιμοποιήσει ένα έτοιμο αρχείο txt που θα περιέχει λέξεις-κλειδιά, τις οποίες μπορεί να επεξεργαστεί. Στη συνέχεια θα έχει τη δυνατότητα να ορίσει τις παραμέτρους χρωμάτων, γραμματοσειράς, collision mode, σχήματος και μεγέθους. Η επιλογή των χρωμάτων θα γίνεται από συγκεκριμένη παλέτα, όπως και η γραμματοσειρά, σχήμα και collision mode θα περιορίζονται σε δύο επιλογές στο καθένα. Οι επιλογές που αποφασίσαμε να δωθούν στον χρήστη είναι οι δυνατότητες της βιβλιοθήκης που θα χρησιμοποιήσουμε.

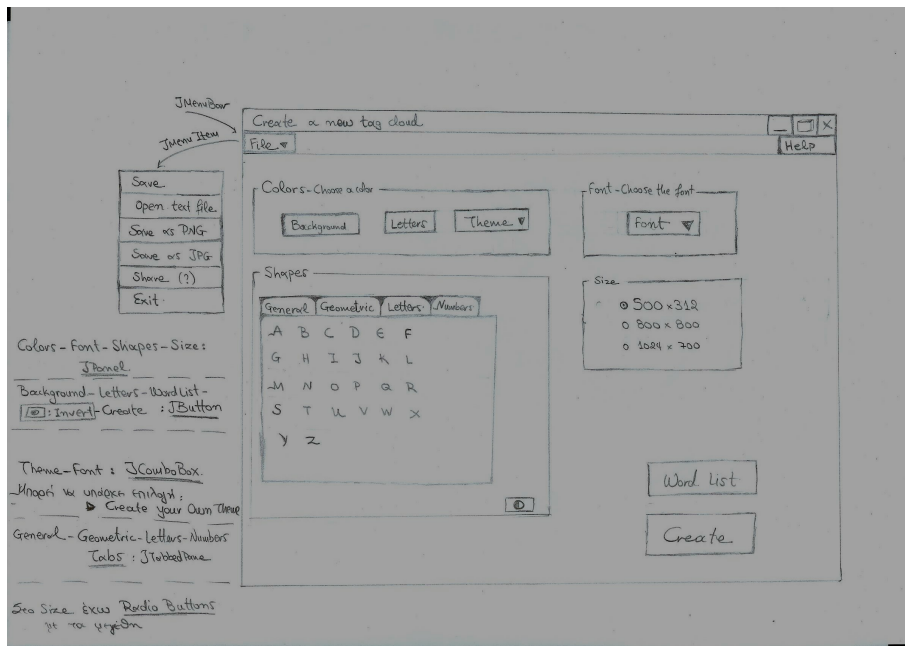
Σκοπό μας ήταν μία όσο δυνατό πιο λειτουργική διεπαφή όπου θα είναι ξεκάθαρες οι επιλογές στον χρήστη. Έτσι σκεφτήκαμε πως μπορούν να κατανεμηθούν σ'ένα user interface οι δυνατότητες ώστε να σχεδιάσουμε ένα εύχρηστο περιβάλλον προς τον χρήστη. Έχοντας τον κατάλογο των επιλογών καταλήξαμε η εμφάνιση της διεπαφής να είναι όσο πιο κοντά γίνεται στην Εικόνα 5.1.

Μεγαλύτερη βαρύτητα δώσαμε στα δεδομένα όπως λέξεις κλειδιά, βάρος και μέτρηση βεβαιότητας τα οποία είναι αποθηκευμένα σε ένα αρχείο (txt). Στο αρχείο αυτό, ο χρήστης αποκτά πρόσβαση επιλέγοντας το κουμπί “Word List”. Πατώντας το εμφανίζεται το παράθυρο της λίστας (Εικόνα 5.2) στο οποίο δίνονται οι επιλογές πρόσθεσης, διαγραφής ακόμα και επεξεργασίας των δεδομένων. Για ευκολία του χρήστη, η λίστα αρχικά είναι κενή, δίνοντας ένα απλό παράδειγμα σύνταξης των δεδομένων. Διαλέγοντας στο μενού το “Open text” φορτώνεται το περιεχόμενο του στον πίνακα του παραθύρου.

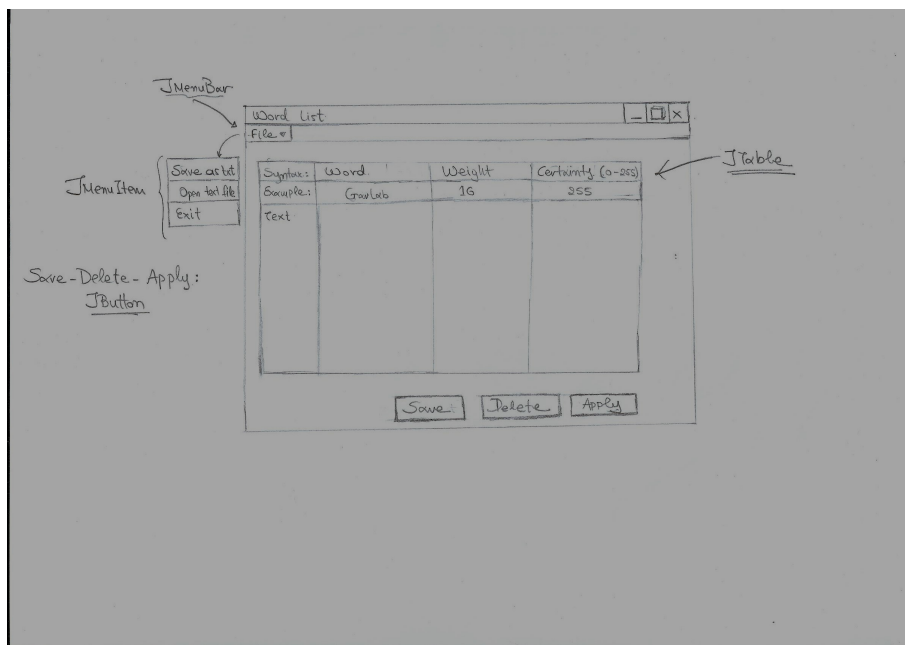
Συνεχίζοντας στην πρώτη οθόνη της διεπαφής, ο χρήστης μπορεί να ορίσει τον συνδυασμό χρωμάτων για το νέο tag cloud. Τα κουμπιά Background και Letters οδηγούν στο παράθυρο που βλέπουμε στην Εικόνα 5.3, από το οποίο ο χρήστης μπορεί να διαλέξει από την παλέτα χρωμάτων τον επιθυμητό συνδυασμό. Στη συνέχεια, με απλό μενού ορίζει την γραμματοσειρά.

Στην εφαρμογή που δημιουργήσαμε ο χρήστης έχει την δυνατότητα να ορίσει τις διαστάσεις για κάθε νέο tag cloud. Στην περιοχή “Size” βρίσκονται δύο text areas με προτεινόμενο μέγεθος “1024*700”. Στην αριστερή πλευρά της αρχικής οθόνης με μενού πολλαπλής επιλογής, το οποίο περιέχει τις pixel perfect και rectangle, ο χρήστης ορίζει το collision mode που είναι απαραίτητο πεδίο για την δημιουργία του νέου tag cloud. Με όμοιο τρόπο έχει υλοποιηθεί και η περιοχή σχήματος. Η βιβλιοθήκη Kumo που χρησιμοποιήσαμε παράγει tag clouds σε δύο σχήματα, κύκλο και τετράγωνο, οι οποίες είναι και δύο επιλογές που χρησιμοποιήσαμε στην διεπαφή.

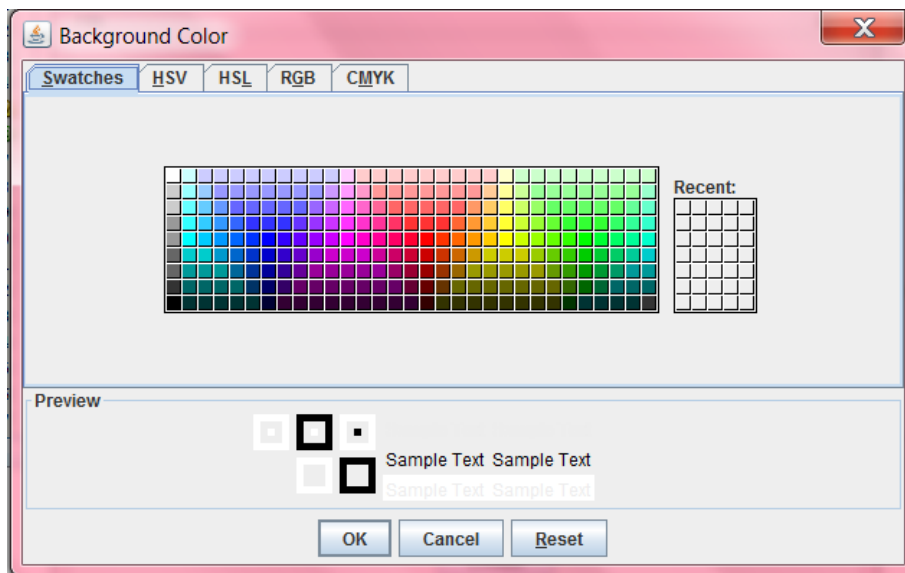
Το κουμπί “Create” είναι αυτό που θα δώσει την τελική εικόνα σε μορφή png. Επιλέγοντας το, ο χρήστης μπορεί να ορίσει τον φάκελο αποθήκευσης. Στη συνέχεια εμφανίζεται το tag cloud που παράχθηκε από όπου μπορεί εύκολα να το σβήσει.



Σχήμα 5.1: Πρώτο Σχέδιο



Σχήμα 5.2: Word List



Σχήμα 5.3: Color Palette

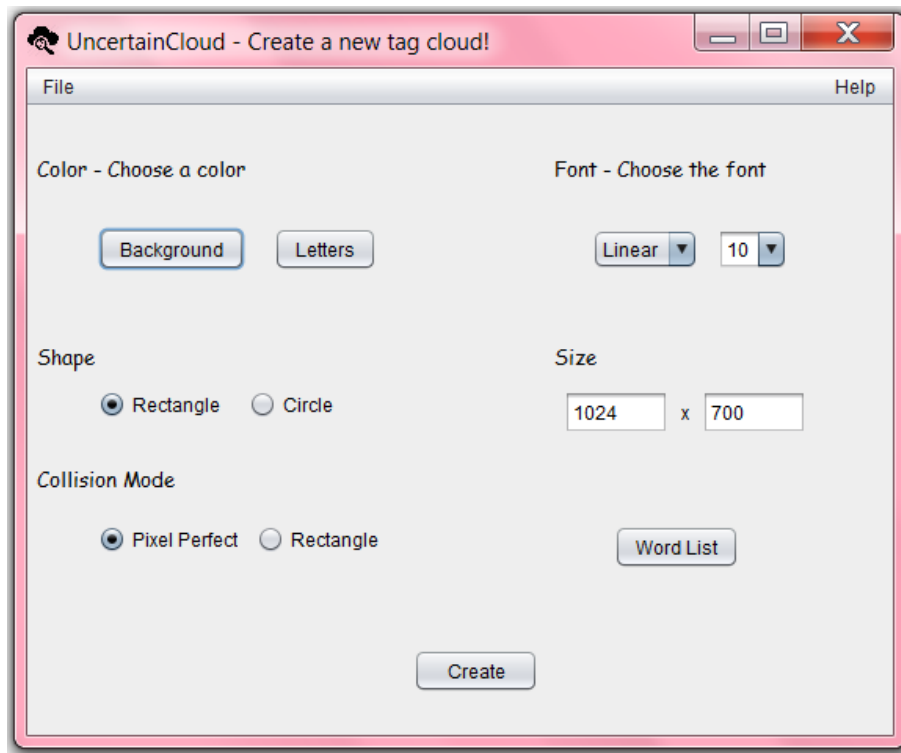
Κεφάλαιο 6

Υλοποίηση

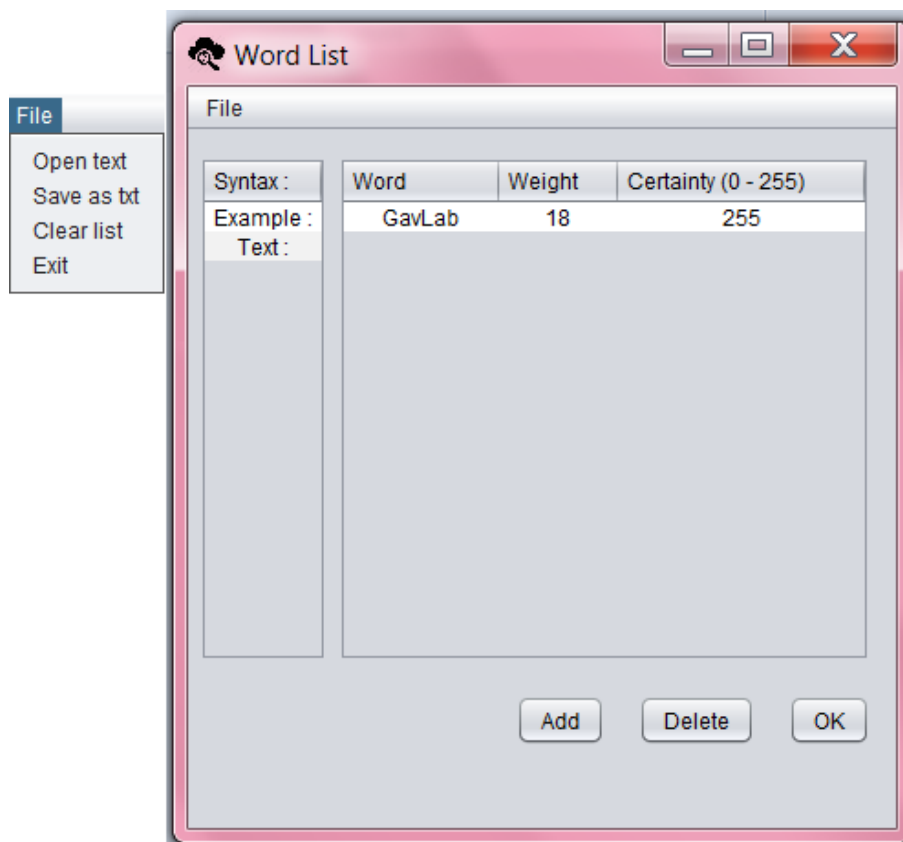
Έχοντας την βιβλιοθήκη Kumo σε Java, για την δημιουργία διεπαφής χρησιμοποιήσαμε την Swing που είναι ένα Gui (Graphical User Interface) widget toolkit για την Java. Παρέχει ένα γραφικό περιβάλλον για τον χρήστη με ισχυρά και ευέλικτα εξαρτήματα όπως κουμπιά και ετικέτες. Παράλληλα με τα γνωστά εργαλεία παρέχει και προηγμένα εξαρτήματα, για παράδειγμα καρτέλες, δέντρα, πίνακες και λίστες [28].

Στην εφαρμογή που υλοποιήσαμε δημιουργήσαμε τρεις κλάσεις (στις οποίες έχουμε ενσωματώσει πακέτα της Swing). Αρχικά στην πρώτη κλάση, η οποία είναι η “page1.java”, σχεδιάσαμε την αρχική οθόνη (Εικόνα 6.1) με τις παραμέτρους για τα tag clouds όπως είχαμε σκεφτεί στο κεφάλαιο 5. Συγκεκριμένα χρησιμοποιήσαμε το πακέτο JButton για να δημιουργήσουμε τα κουμπιά “Background”, “Letters”, “Word List” και “Create”. Το κάθε κομμάτι της σελίδας, δηλαδή οι παράμετροι χρώμα, γραμματοσειρά, σχήμα, μέγεθος και τέλος collision mode βρίσκονται σε ξεχωριστό JPanel ώστε να υπάρχει η επιθυμητή στοίχιση και η τελική εμφάνιση. Στο Panel για το Shape και το Collision Mode, δημιουργήσαμε JRadioButton για τις επιλογές που δίνονται στο χρήστη. Με την χρήση των κουμπιών πολλαπλής επιλογής περιορίζουμε της δυνατότητες της εφαρμογής, σε δύο σχήματα ορθογώνιο και κύκλο και το collision mode σε rectangle και pixel perfect. Στη συνέχεια, στο Panel για το Size χρησιμοποιήσαμε δύο text areas, ώστε ο χρήστης να έχει την ευχέρεια αλλαγής σε οποιοδήποτε ορισμό των διαστάσεων της εικόνας. Τέλος, στο πάνω μέρος μέσω του πακέτου JMenuBar, εμφανίσαμε την μπάρα μέσω της οποίας παρέχονται κάποιες πληροφορίες και η δυνατότητα εξόδου.

Σε ξεχωριστή κλάση υλοποιήσαμε την wordList για την καλύτερη διαχείριση των δεδομένων (Εικόνα 6.2). Στην κλάση αυτή περιέχεται ο πίνακας των λέξεων με τις τρεις εξής στήλες: word, weight και certainty. Για την σχεδίαση του πίνακα χρησιμοποιήσαμε δύο JTable, στο πρώτο παρουσιάζουμε την δομή του δεύτερου, όπως παρατηρείτε υπάρχει η επεξήγηση για τις δύο πρώτες γραμμές, Syntax και Example, και στις υπόλοιπες γραμμές, όπως ορίζεται, θα εμφανιστεί το υπόλοιπο text. Στο JMenuBar αυτής της κλάσης επιλέγοντας το JMenuItem:File ο χρήστης μπορεί να “ανεβάσει” ένα έτοιμο αρχείο με τα δεδομένα που θα περιέχονται στο tag cloud. Στο κάτω μέρος του παραθύρου έχουμε τρία JButton για την προσθήκη και την διαγραφή λέξεων του πίνακα, και το ok μέσω του οποίου μπορεί να κλήσει το παράθυρο και να αποθηκεύσει και τις αλλαγές σε ένα αρχείο txt.



Σχήμα 6.1: Αρχική οθόνη εφαρμογής



Σχήμα 6.2: Λίστα λέξεων

6.1 Σύνδεση

Για την σύνδεση της εφαρμογής με την βιβλιοθήκη παραγωγής tag cloud κάναμε import τα απαραίτητα packages wordcloud και δημιουργήσαμε ένα αντίστοιχο αντικείμενο. Για εύκολη διαχείριση των δεδομένων εισόδου που θα δωθούν απο τον χρήστη δημιουργήσαμε τις ακόλουθες μεθόδους.

- collisionCall();
- shapeCall();
- colorCall();
- fontCall();

Αναλυτικότερα, στην πρώτη μέθοδο ελέγχουμε ποιο radio button επέλεξε ο χρήστης για το collision mode. Επιπλέον, παίρνουμε τα δεδομένα από τα text areas για τον ορισμό του μεγέθους. Με τα δεδομένα εισόδου που συλλέξαμε πραγματοποιούμε την δημιουργία του αντικείμενο τύπου wordcloud της Kumo με τα ορίσματα του χρήστη.

Στην συνέχεια, για να οριστεί το σχήμα σύμφωνα με την επιλογή που έχει γίνει, μέσω της εφαρμογής, πραγματοποιείται έλεγχος, στην συνάρτηση shapeCall. Ανάλογα με το ποιο radio button έχει επιλεγθει, ακολουθεί η κλήση της μεθόδου setBackground, μέσω του αντικειμένου wordcloud. Όπως έχουμε αναφέρει στο κεφάλαιο 3, στη μέθοδο αυτή δίνονται τα ορίσματα CircleBackground και RectangleBackground για τον κύκλο και το τετράγωνο αντίστοιχα.

Στη μέθοδο colorCall της κλάσης αυτής συλλέγουμε τα χρώματα που έχει επιλέξει ο χρήστης από την παλέτα χρωμάτων για το φόντο και τα γράμματα τα οποία έχουν αποθηκευτεί σε δύο πεδία τύπου color. Κατά την κλήση οι μέθοδοι της Kumo μέσω των οποίων θέτουμε τα χρώματα είναι setBackgroundColor και setColorPalette, όπου δίνουμε και ως όρισμα της επιλογές του χρήστη.

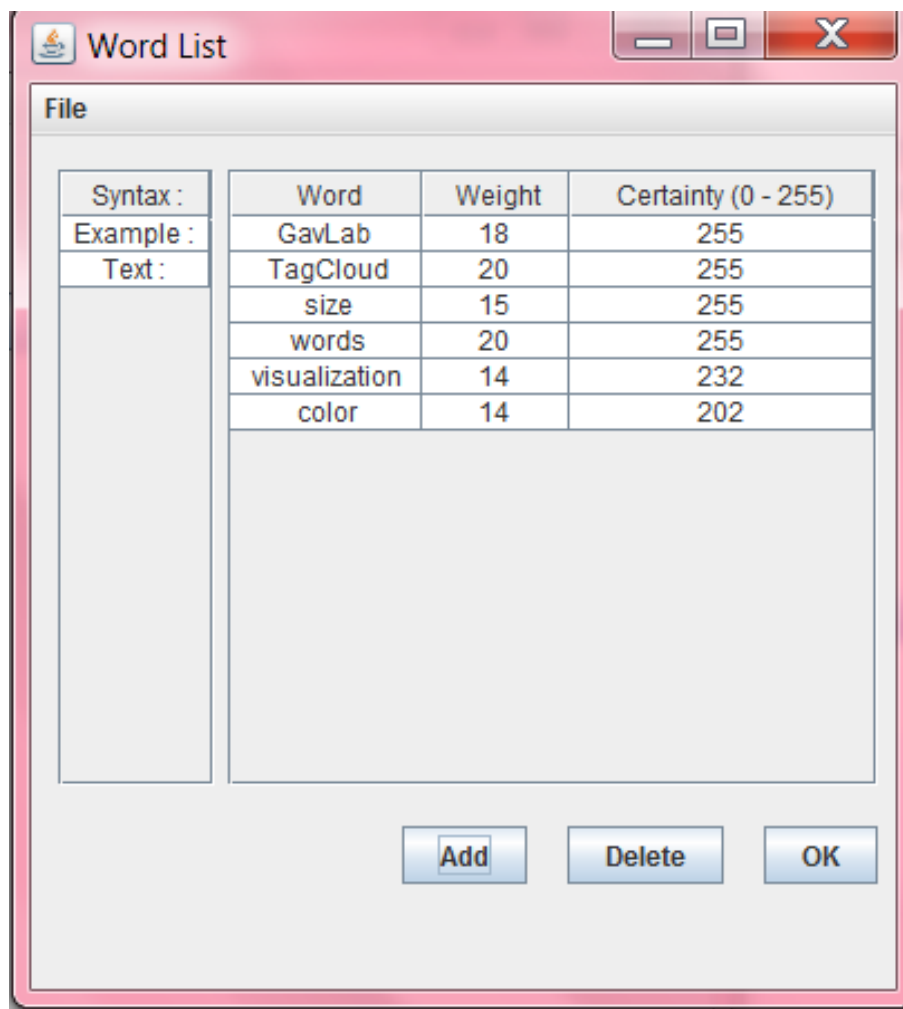
Μέσω της τελευταίας συνάρτησης αλλάζουμε και ορίζουμε την γραμματοσειρά. Έχουμε ορίσει ως προεπιλογή την linear γραμματοσειρά. Ελέγχουμε αν ο χρήστης έχει επιλέξει την Sqrl και καλούμε την μέθοδο setFontScalar με το αντίστοιχο όρισμα και ένα καθορισμένο μέγεθος.

Τελευταίο βήμα είναι ο ορισμός της θέσης που θα αποθηκευθεί η τελική εικόνα. Κατά την κλήση της μεθόδου writeToFile, που όπως έχουμε εξηγήσει δημιουργεί την εικόνα, θέτουμε ως όρισμα το path που έχει επιλέξει ο χρήστης μέσω του πακέτου JFileChooser της swing.

6.2 Δομή αρχείου λέξεων

Κάθε λέξη συνοδεύεται από το βάρος της και τον βαθμό βεβαιότητας, απαραίτητα χαρακτηριστικά για την εμφάνιση της κάθε λέξης-κλειδί στο tag cloud. Ο χρήστης μπορεί να εισάγει τις επιθυμητές λέξεις-κλειδιά στον πίνακα και τέλος να τις αποθηκεύσει σε ένα αρχείο txt. Το αρχείο “words.txt”, που θα περιέχει τα δεδομένα, για εύκολη διαχείριση επιλέξαμε να μην είναι binary έτσι ώστε ο χρήστης να μπορεί να το επεξεργαστεί απλά, επιπλέον να έχει την δυνατότητα να χρησιμοποιήσει τα δεδομένα αλλού.

Η εσωτερική μορφή του αρχείου λέξεων είναι νοητά ο πίνακας της εφαρμογής. Αναλυτικότερα η δομή του πρέπει να είναι αυστηρά πρώτα η λέξη, θα ακολουθείται από μόνο ένα κενό. Στη συνέχεια θα γράφεται το βάρος, το οποίο περιγράφεται με έναν ακέραιο αριθμό, και θα ακολουθείται από ένα κενό. Το επόμενο στοιχείο είναι ο βαθμός βεβαιότητας, ο οποίος είναι ένας ακέραιος αριθμός στο διάστημα 0-255. Ένα μικρό παράδειγμα του αρχείου είναι το ακόλουθο:



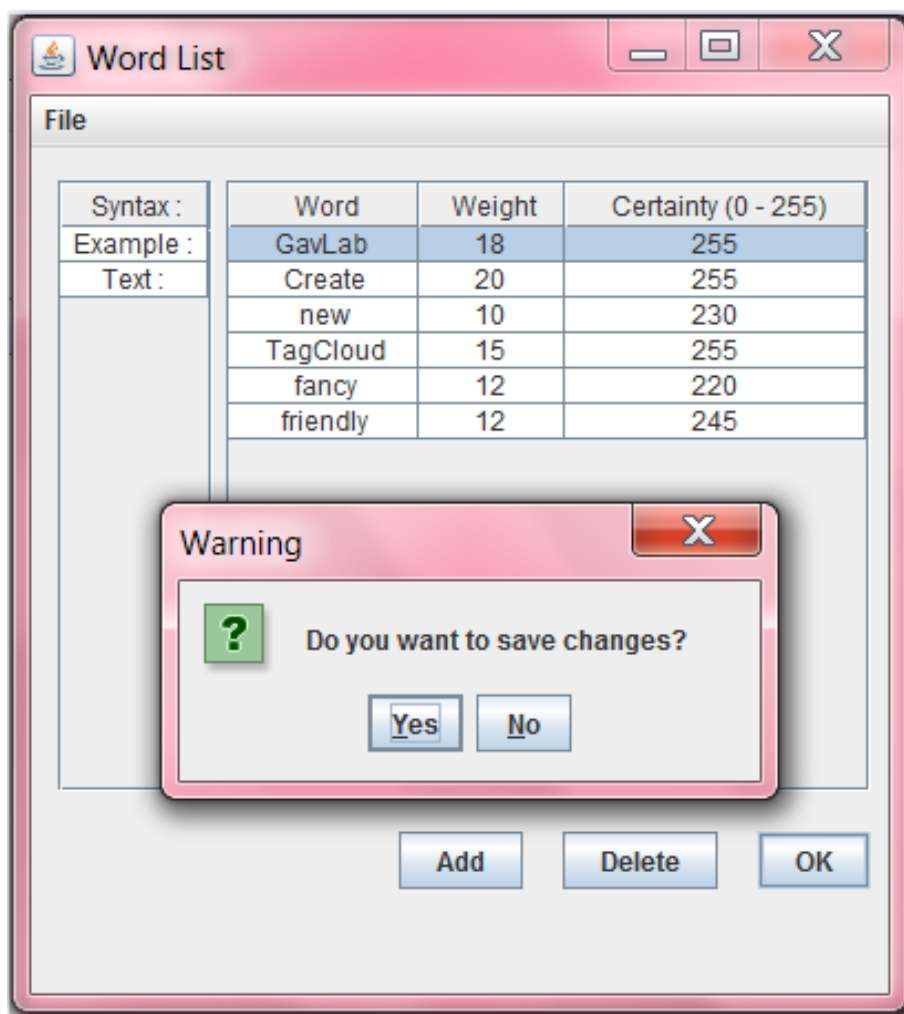
Σχήμα 6.3: Παράδειγμα λίστας λέξεων

```
GavLab 25 255
TagCloud 20 255
size 15 255
words 20 255
visualization 14 232
color 14 202
```

Στην εφαρμογή το ίδιο κομμάτι λέξεων εμφανίζεται στον πίνακα όπως φαίνεται στην εικόνα 6.3.

6.3 Μετατροπή σε jar

Η εφαρμογή υλοποιήθηκε στο πρόγραμμα NetBeans, μέσω του οποίου η μετατροπή σε jar εκτελέσιμο αρχείο είναι πολύ απλή. Στο περιβάλλον αυτό, όταν το project έχει ολοκληρωθεί πλήρως, υπάρχει η δυνατότητα “Clean and build Project” (Shift + F11) στην καρτέλα Run. Όταν ολοκληρωθεί, όπως φαίνεται στην εικόνα 6.5, εμφανίζει το path στο οποίο έχει αποθηκευθεί το αρχείο jar. Στον φάκελο dist του project έχει δημιουργηθεί ένα εκτελέσιμο jar, ένας φάκελος lib (έαν έχουμε εισάγει πρόσθετες βιβλιοθήκες στην εφαρμογή) και ένα αρχείο README.txt. Το αρχείο κειμένο με το όνομα readme μπορεί να διαγραφεί χωρίς να δημιουργηθεί σφάλμα.



Σχήμα 6.4: Αποθήκευση σε νέο αρχείο txt

```

Output - KumoTagCloud (clean.jar)
ant -f C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\Dnb.internal.action.name=rebuild clean jar
init:
deps-clean:
Updating property file: C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\build\build-clean.properties
Deleting directory C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\build
clean:
init:
deps-jar:
Created dir: C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\build
Updating property file: C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\build\build-jar.properties
Created dir: C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\build\classes
Created dir: C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\build\empty
Created dir: C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\build\generated-sources\ap-source-output
Compiling 48 source files to C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\build\classes
Note: Some input files use unchecked or unsafe operations.
Note: Recompile with -Xlint:unchecked for details.
Copying 6 files to C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\build\classes
compile:
Created dir: C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\dist
Copy libraries to C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\dist\lib.
Building jar: C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\dist\KumoTagCloud.jar
To run this application from the command line without Ant, try:
C:\Program Files\Java\jdk1.8.0_65\bin/java -jar "C:\Users\Evi\Documents\NetBeansProjects\KumoTagCloud\dist\KumoTagCloud.jar"
jar:
BUILD SUCCESSFUL (total time: 4 seconds)

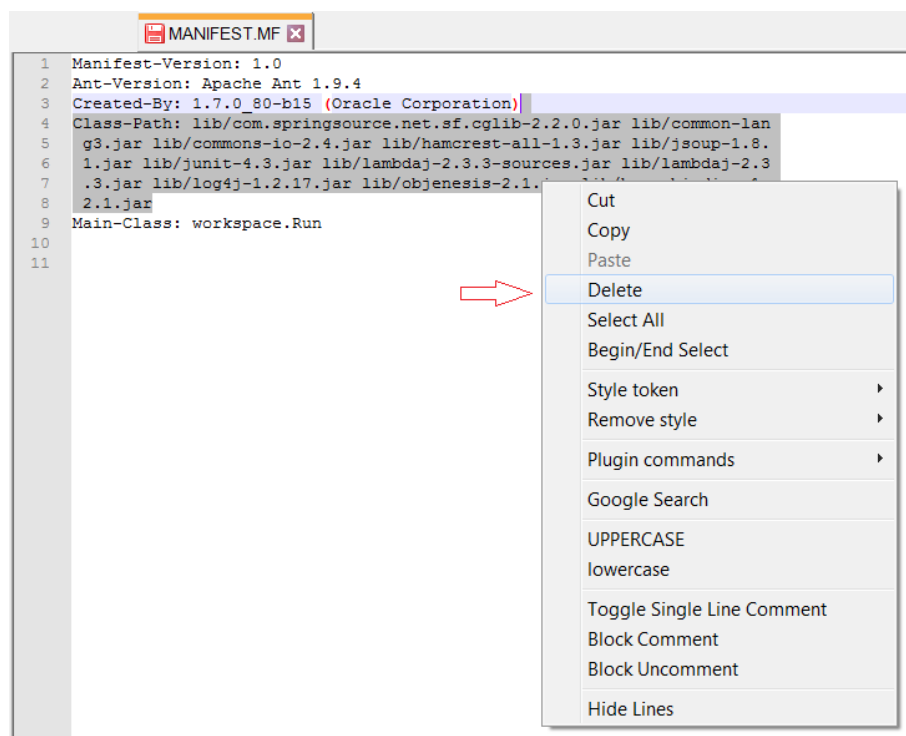
```

Σχήμα 6.5: Output

Η εφαρμογή έχει πλέον μετατραπεί σε μορφή jar, και μπορεί να μεταφερθεί, έχοντας και τον φάκελο lib, οπουδήποτε. Για να αποφύγουμε τυχόν σφάλματα κατά την μετακίνηση, μπορούμε να ενσωματώσουμε τις βιβλιοθήκες στο εκτελέσιμο αρχείο ακολουθώντας τα παρακάτω βήματα.

Αρχικά μετατρέψαμε τις βιβλιοθήκες από αρχεία τύπου jar σε zip μορφή. Η μετατροπή μπορεί να γίνει είτε χειροκίνητα, αλλάζοντας την κατάληξη του αρχείου, είτε διαδικτυακά [29]. Κάθε αρχείο περιέχει διάφορους φακέλους τους οποίους μετακινήσαμε, εκτός του φακέλου META-INF, στον φάκελο που έχουμε το εκτελέσιμο αρχείο. Η κίνηση αυτή πραγματοποιήθηκε για όλες τις βιβλιοθήκες ξεχωριστά. Οι ονομασία κάποιων φακέλων μπορεί να είναι κοινή, αλλά το περιεχόμενο διαφορετικό, για τον λόγο αυτό επιλέξαμε την συγχώνευση των φακέλων.

Επόμενο βήμα, είναι η μετατροπή της εφαρμογής (jar) σε συμπιεσμένη μορφή zip. Πραγματοποιήσαμε αποσυμπίεση στο μέρος που είχαμε και του φακέλους που πήραμε από τις βιβλιοθήκες. Στον φάκελο META-INF που εμφανίστηκε, υπάρχει ένα αρχείο το οποίο επεξεργαστήκαμε όπως φαίνεται στην εικόνα 6.6. Συμπιέζοντας όλους του φακέλους και μετατρέποντας το αρχείο ξανά σε μορφή jar, ενσωματώσαμε και τις βιβλιοθήκες.



Σχήμα 6.6: Επεξεργασία αρχείου Manifest

Κεφάλαιο 7

Δοκιμαστική λειτουργία

Στο κεφάλαιο αυτό θα παρουσιάσουμε διαφορετικά σενάρια χρήσης της εφαρμογής μας. Αναλυτικότερα θα παραθέσουμε την βασική λειτουργικότητα που προσφέρει η εφαρμογή UncertainCloud, μέσα από μία σειρά εικόνων και σεναρίων.

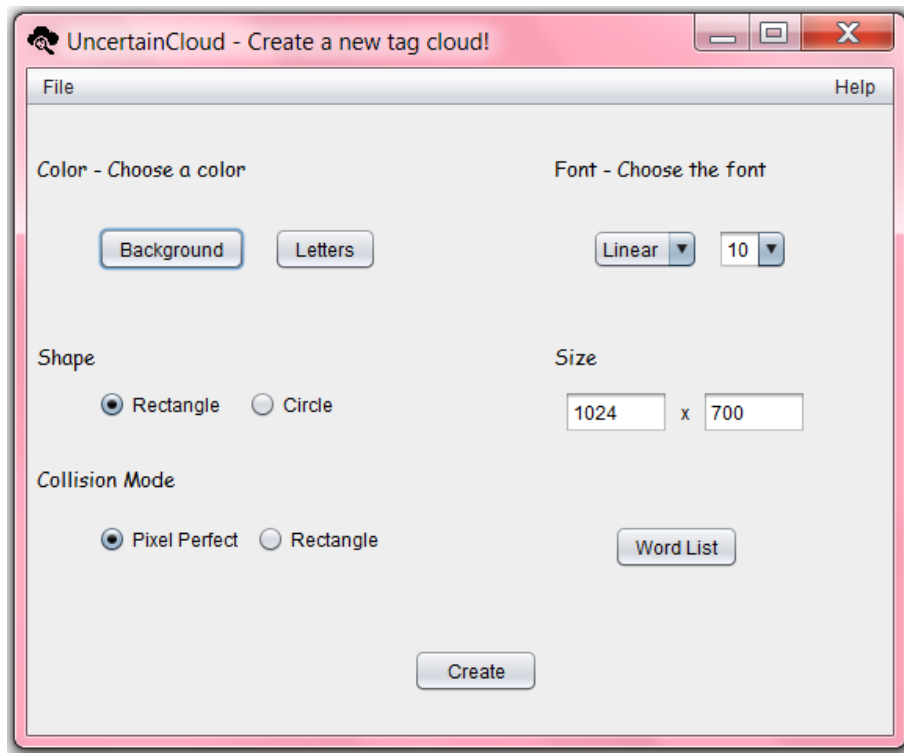
7.1 Σενάριο πρώτο

Ο κ. Γιάννης είναι ένας καθηγητής και θέλει να παρουσιάσει στους μαθητές του σημαντικά στοιχεία του κύκλου μέσω ενός tag cloud. Πρώτη του κίνηση είναι να ανοίξει την εφαρμογή. Όταν εμφανιστεί η Οθόνη 1 (Εικόνα 7.1): *UncertainCloud - Create a new tag cloud*, μέσω της οποίας δίνεται η επιλογή αλλαγής των παραμέτρων ενός tag cloud, ο κ. Γιάννης επιλέγει το κουμπί Background για να επιλέξει το φόντο της εικόνας του. Στην Οθόνη 2 (Εικόνα 7.2): *Background Color*, που εμφανίζεται, ο κ. Γιάννης διαλέγει ένα από τα διαθέσιμα χρώματα, ο κωδικός (HEX) του είναι 006666, και πατά το OK. Αφού επιστρέφει στην αρχική οθόνη 1 (Εικόνα 7.1), μετατρέποντας το χρώμα από το κουμπί σε αυτό που επέλεξε ο κ. Γιάννης, επιλέγει το κουμπί Letters. Στο ίδιο παράθυρο για την επιλογή χρώματος, αυτή τη φορά για το χρώμα της γραμματοσειράς του tag cloud η επιλογή του θα είναι το χρώμα με κωδικό CCF33.

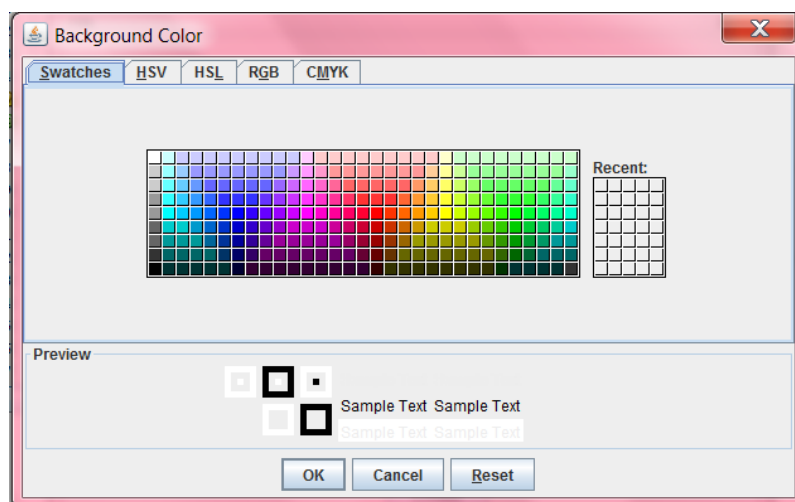
Στη συνέχεια διαλέγει την γραμματοσειρά Sqrt και μέγεθος 9, ώστε να φαίνονται όλες οι λέξεις της λίστας του. Ορίζοντας το σχήμα να είναι Circle και το Collision Mode να είναι Rectangle θα παρουσιάσει σχηματικά τον κύκλο του με τα στοιχεία αραιά τοποθετημένα. Ο κ. Γιάννης θεωρεί ως ιδανικό μέγεθος για την εικόνα του το 700x700, και το πληκτρολογεί στα πεδία ορισμού στην περιοχή για τις διαστάσεις.

Ο κ. Γιάννης έχει έτοιμο το αρχείο των λέξεων που θέλει να εμπεριέχονται στο tag cloud. Συνεπώς πατά το κουμπί wordlist και στην Οθόνη 3 (Εικόνα 7.3): *Word List*, που ανοίγει ως νέο παράθυρο επιλέγει File/Open text. Στον πίνακα λέξεων φορτώνεται και εμφανίζεται το περιεχόμενο του text αρχείου που είχε ήδη δημιουργήσει. Δεν θέλει να αλλάξει κάτι από το περιεχόμενο του οπότε επιλέγει OK και επιστρέφει στην αρχική οθόνη (Εικόνα 7.4).

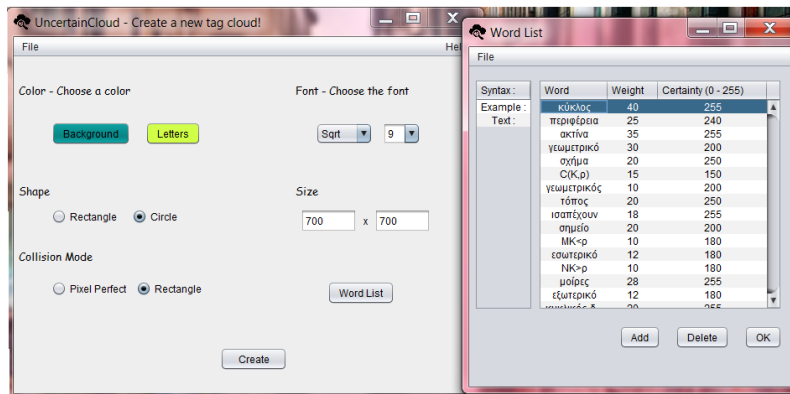
Τέλος, θα επιλέξει το Create, ορίζοντας το Desktop ως προορισμό αποθήκευσης της εικόνας. Η εφαρμογή αφού επεξεργάσει τα δεδομένα, δημιουργεί το tag cloud. Αποθηκεύει την εικόνα στην επιφάνεια εργασίας που ορίστηκε και την εμφανίζει (Εικόνα 7.5).



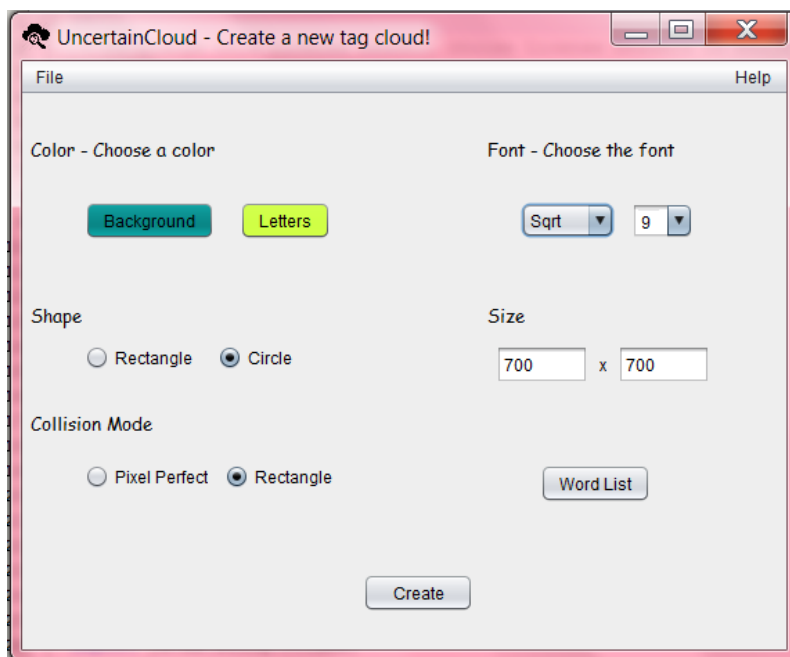
Σχήμα 7.1: Οθόνη 1



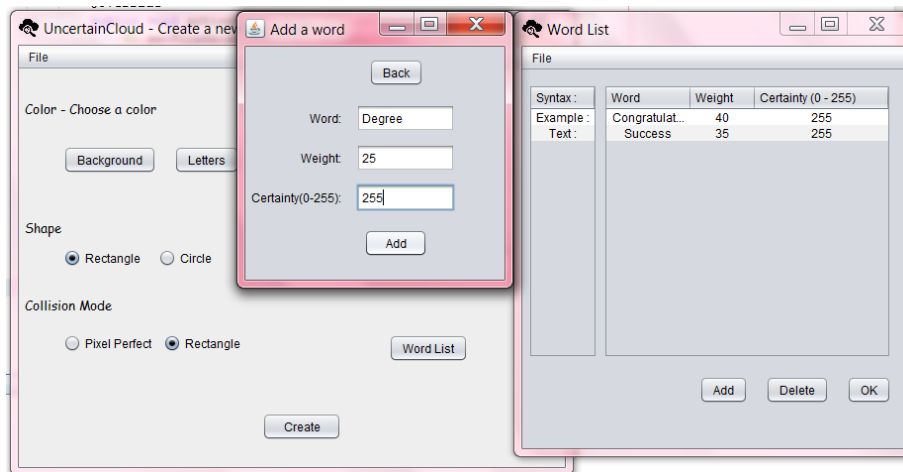
Σχήμα 7.2: Οθόνη 2



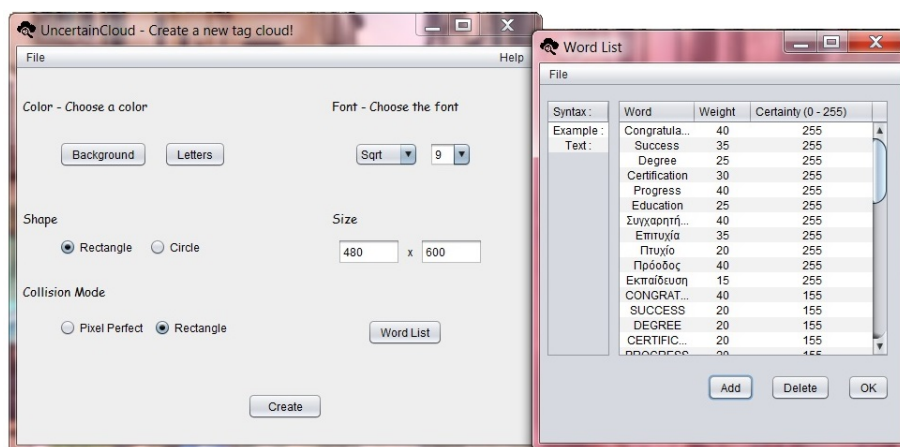
Σχήμα 7.3: Οθόνη 3



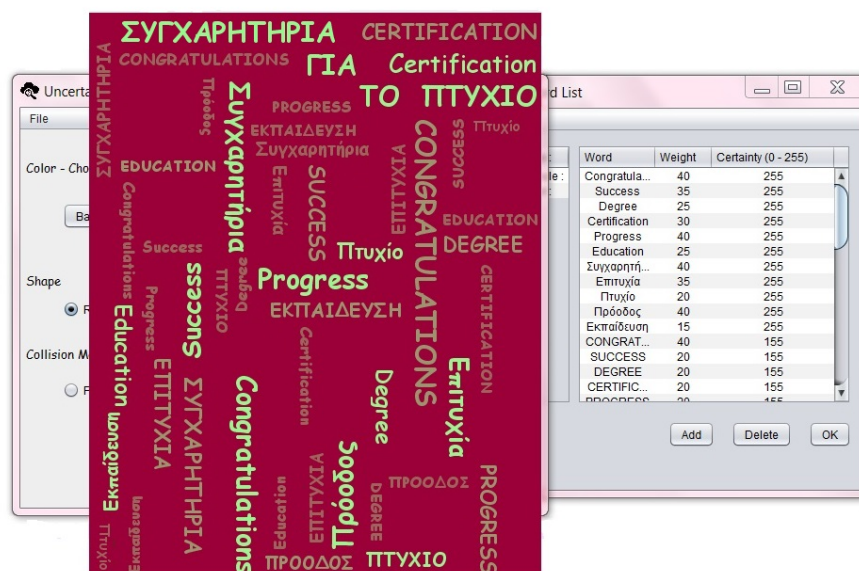
Σχήμα 7.4: Τελική οθόνη



Σχήμα 7.6: Παράθυρο προσθήκης λέξεων



Σχήμα 7.7: Τελική οθόνη



Σχήμα 7.8: Σενάριο 2: Image Pop up

Κεφάλαιο 8

Συμπεράσματα

Στην εργασία αυτή ασχοληθήκαμε με τον ορισμό διαφορετικών εννοιών σχετικά με τα tag clouds και την έννοια της αβεβαιότητας. Στην συνέχεια προχωρήσαμε στον σχεδιασμό και υλοποιήσαμε μια διεπαφή όσο πιο λειτουργική και εύχρηστη ως προς τον χρήστη. Βασικό εργαλείο για την υλοποίηση της εφαρμογής μας ήταν η επέκταση της βιβλιοθήκης Kumo της Java. Παρουσιάστηκαν λεπτομερώς τα πακέτα που την αποτελούν ώστε να γνωρίσουμε όλες τις δυνατότητες που μας παρέχει και να τις ενσωματώσουμε στην εφαρμογή που δημιουργήσαμε.

Στηριζόμενοι κυρίως σε διεθνή άρθρα και επιστημονικές δημοσιεύσεις αναπτύξαμε την έννοια της αβεβαιότητας καθώς επικεντρωθήκαμε στις αβέβαιες πληροφορίες και στα είδη, σύμφωνα με τα χαρακτηριστικά της πληροφορίας, που μπορούμε να την κατηγοριοποιήσουμε. Αναλύσαμε την χρήση της διαφάνειας στα tag clouds και πως μέσω αυτής οπτικοποιείται ο βαθμός της αβεβαιότητας.

Δημιουργήσαμε και αναπτύξαμε την εφαρμογή Uncertain Cloud όπου ενσωματώθηκαν σε σημαντικό βαθμό οι λειτουργίες της Kumo. Συγκεκριμένα βασικά χαρακτηριστικά των tag clouds που δημιουργούνται είναι το χρώμα, η γραμματοσειρά, το μέγεθος, το collision mode και δύο βασικά σχήματα. Στην εφαρμογή Uncertain Cloud εφαρμόστηκε επιτυχώς η οπτικοποίηση της βεβαιότητας. Αναλυτικότερα, στη λίστα που αποτελείται από τις λέξεις που θα περιέχονται στο σύννεφο και θα μπορούν να εισαχθούν από έτοιμο αρχείο συμπεριλαμβάνουμε την παράμετρο της αβεβαιότητας, μέσω ενός εύρους ακεραίων (0-255).

Γενικά, μέσω της εφαρμογής Uncertain Cloud, που είναι διαθέσιμη πλέον, μπορεί εύκολα ένας απλός χρήστης να οπτικοποιήσει συγκεκριμένα δεδομένα, χάρη στην προσοχή που δώθηκε στις λεπτομέρειες κατά την δημιουργία της. Η βιβλιοθήκη Kumo είναι ένας open source κώδικας χωρίς να συνοδεύεται από κάποια διεπαφή για έναν απλό χρήστη. Θετικό είναι ότι τώρα οι χρήστες που θέλουν να δημιουργήσουν κάτι με βάση την Kumo μπορούν να το κάνουν χρησιμοποιώντας απλά το εργαλείο που δημιουργήσαμε, στο οποίο πραγματοποιείται και η οπτικοποίηση της βεβαιότητας.

8.1 Επεκτάσεις και βελτιώσεις

Παρότι έχουν γίνει τα αρχικά και σημαντικά βήματα της εφαρμογής Uncertain Cloud, πολλά ακόμα θα μπορούσαν να γίνουν προκειμένου να βελτιωθεί και να προσφέρει περισσότερες λειτουργίες στους χρήστες. Συνεπώς, στο πλαίσιο περαιτέρω ανάπτυξης της παρούσας πτυχιακής εργασίας προτείνεται η επέκταση των λειτουργιών της διεπαφής λαμβάνοντας υπόψη επιπλέον παραμέτρους.

Εκτενέστερα θα μπορούσαμε να ενσωματώσουμε περισσότερες επιλογές όσον αφορά τα σχήματα, δηλαδή να παρουσιάζεται ένα πλήθος επιλογών χωρισμένο σε κατηγορίες, παραδείγματος χάρη γράμματα, αριθμοί, γεωμετρικά σχήματα. Επιπλέον θα ήταν εφικτό να πραγματοποιηθεί ένας διαφορετικός τρόπος εισαγωγής δεδομένων, εντάσσοντας την αυτόματη επεξεργασία κειμένων ώστε να προκύψουν οι λέξεις-κλειδιά, το βάρος, και οι βαθμοί σημασίας και βεβαιότητας.

Στην περίπτωση που υλοποιούνταν η επιλογή διαφορετικών χρωμάτων για κάθε λέξη-κλειδί θα βελτιώνονταν η εμφάνιση του τελικού tag cloud που δημιουργεί η εφαρμογή. Με αυτή την πρόσθετη λειτουργία, θα δίνονταν έμφαση με την βοήθεια των χρωμάτων σε επιλεγμένες λέξεις-κλειδιά. Επεκτείνοντας κι άλλο την παραπάνω πρόταση θα μπορούσε να δίνεται η επιλογή στον χρήστη ανάμεσα από προκαθορισμένα θέματα συνδιασμένων χρωμάτων.

Κλείνοντας, μετά την υλοποίηση της διαφάνειας για την οπτικοποίηση της αβεβαιότητας, σημαντικό θα ήταν αν δημιουργούνταν η οπτικοποίηση context με έναν διαφορετικό και ευδιάκριτο για τον χρήστη τρόπο.

Βιβλιογραφία

- [1] Tag clouds https://en.wikipedia.org/wiki/Tag_cloud
- [2] Wallace Manolis and Platis Nikos, *The uncertain tag cloud* Semantic and Social Media Adaptation and Personalization (SMAP), 2015 10th International Workshop on (2015) 1-5
- [3] Kenny Cason, Kumo - Java Word Cloud, <http://kennycason.com/posts/2014-07-03-kumo-wordcloud.html>
- [4] Platis Nikos, Wallace Manolis and Triantos Thanos *Visualization of Uncertainty in Tag Clouds*, Semantic Keyword-based Search on Structured Data Sources (2015) 127-132
- [5] Wordle <http://www.wordle.net/>
- [6] Tagcloud Generator <http://www.tagcloud-generator.com/>
- [7] Word Cloud Generator <http://www.jasondavies.com/wordcloud/>
- [8] Tagxedo <http://www.tagxedo.com/>
- [9] TagCrowd <http://tagcrowd.com/>
- [10] Flickr <https://en.wikipedia.org/wiki/Flickr>
- [11] Flickr Tag <https://www.flickr.com/photos/tags>
- [12] Σύννεφο ετικετών https://el.wikipedia.org/wiki/%CE%A3%CF%8D%CE%BD%CE%BD%CE%B5%CF%86%CE%BF_%CE%B5%CF%84%CE%B9%CE%BA%CE%B5%CF%84%CF%8E%CE%BD
- [13] Bogárdi-Mészöly, Ágnes and Hashimoto, Takeshi and Yokoyama, Shohei and Ishikawa, Hiroshi *Frequency-and Content-Based Tag Cloud Font Distribution Algorithm*
- [14] Tag clouds - History https://en.wikipedia.org/wiki/Tag_cloud#History
- [15] Creation of a Tag Cloud https://en.wikipedia.org/wiki/Tag_cloud#Creation_of_a_tag_cloud
- [16] Sgurev, Vassil and Hadjiski Mincho *Intelligent systems: From theory to practice* Vol. 299. Springer (2010)
- [17] Kaser, Owen and Lemire, Daniel, *Tag-cloud drawing: Algorithms for cloud visualization*, arXiv preprint cs/0703109 (2007)
- [18] Tag Clouds Gallery <https://www.smashingmagazine.com/2007/11/tag-clouds-gallery-examples-and-good-practices/>

- [19] <http://www.joelamantia.com/tag-clouds/10-best-practices-for-displaying-tag-clouds>
- [20] <https://www.fotolia.com/id/42297385>
- [21] <http://www.alamy.com/stock-photo-learning-concept-in-tag-cloud-of-human-head-shape-on-white-background-50313572.html>
- [22] <http://www.shutterstock.com/pic-127855526/stock-photo-where-who-why-how-when-info-text-question-word-cloud-concept-word-cloud-tag-cloud-text-business.html>
- [23] http://www.123rf.com/photo_23796130_big-data-concept-in-word-tag-cloud-on-white-background.html
- [24] Kumo packages <https://github.com/kennycason/kumo/tree/master/src/main/java/com/kennycason/kumo>
- [25] Kumo text <https://github.com/kennycason/kumo/tree/master/src/test/resources/text>
- [26] Bradley Richard and Drechsler Mareile *Types of uncertainty*, Springer(2014) 1225-1248
- [27] Γιάννης Ιωαννίδης, Γιώργος Λέπουρας *Σημειώσεις Επικοινωνίας Ανθρώπου-Μηχανής*
- [28] Swing Java [https://en.wikipedia.org/wiki/Swing_\(Java\)](https://en.wikipedia.org/wiki/Swing_(Java))
- [29] Jar to Zip <https://convertio.co/>
- [30] Sinclair, James and Cardew-Hall, Michael, *The folksonomy tag cloud: when is it useful?*, Journal of Information Science (2008) 15-29