



ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΘΕΜΑ: «Εξόρυξη Σημασιολογικών υπό-δικτύων από Κοινωνικά Δίκτυα»

ΣΠΟΥΔΑΣΤΗΣ: Κισσαμιτάκης Εμμανουήλ

ΕΠΙΒΛΕΠΟΝ ΜΕΛΟΣ Δ.Ε.Π: Μυλωνάς Φοίβος

ΑΘΗΝΑ 2020

**Στην οικογένεια μου
και στους φίλους μου Έλλη, Έλενα και Νίκο.**

.....

.....

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1^ο - ΕΙΣΑΓΩΓΙΚΑ

- 1.1.ΕΙΣΑΓΩΓΗ.....
- 1.2.ΠΡΟΗΓΟΥΜΕΝΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ.....

ΚΕΦΑΛΑΙΟ 2^ο – ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

- 2.1.ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ.....
- 2.2. ΚΟΙΝΩΝΙΚΑ ΔΙΚΤΥΑ.....
- 2.3. BOTS.....
- 2.4. HASHTAGS.....
- 2.5. GAMIFICATION.....
- 2.6. AUGMENTED REALITY.....

ΚΕΦΑΛΑΙΟ 3^ο – ΜΕΘΟΔΟΛΟΓΙΑ

- 3.1. ΜΕΘΟΔΟΛΟΓΙΑ.....
- 3.2. ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΧΡΗΣΗ CRAWLER.....

ΚΕΦΑΛΑΙΟ 4^ο – ΕΡΕΥΝΑ

- 4.1. ΕΡΕΥΝΑ.....

ΚΕΦΑΛΑΙΟ 5^ο – ΕΠΙΛΟΓΟΣ

- 5.1. ΓΕΝΙΚΕΣ ΠΑΡΑΤΗΡΗΣΕΙΣ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ

ΚΕΦΑΛΑΙΟ 6^ο – ΒΙΒΛΙΟΓΡΑΦΙΑ

- 6.1. ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΠΗΓΕΣ

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ – ΓΡΑΦΗΜΑΤΩΝ – ΕΙΚΟΝΩΝ:

ΠΙΝΑΚΕΣ

Πίνακας 1 – Augmented Reality Tweets Ανά Ημέρα.....σελ.24	σελ.24
Πίνακας 2 – Gamification Tweets Ανά Ημέρα σελ.26	σελ.26
Πίνακας 3 – Augmented Reality Top Users Ανά Ημέρα σελ.27	σελ.27
Πίνακας 4 – Gamification Reality Top Users Ανά Ημέρα σελ.29	σελ.29
Πίνακας 5 – Both Hashtags Top Users Ανά Ημέρα σελ.31	σελ.31
Πίνακας 6 – Augmented Reality Top Users Retweets σελ.34	σελ.34
Πίνακας 7 – Gamification Top Users Retweets σελ.35	σελ.35
Πίνακας 8 – Augmented Reality Top Users Retweeted σελ.37	σελ.37
Πίνακας 9 – Gamification Top Users Retweeted σελ.38	σελ.38

ΔΗΛΩΣΗ ΠΕΡΙ ΜΗ ΛΟΓΟΚΛΟΠΗΣ

Δηλώνω ότι είμαι ο η συγγραφέας της παρούσας εργασίας με τίτλο «Εξόρυξη Σημασιολογικών υπό-δικτύων από Κοινωνικά Δίκτυα». Η αναφερόμενη εργασία δεν αποτελεί αντιγραφή ούτε προέρχεται από ανάθεση σε τρίτους. Οι πηγές που χρησιμοποιήθηκαν αναφέρονται σαφώς στη βιβλιογραφία ενώ κάθε εξωτερική βοήθεια, αναγνωρίζεται ρητά.

Εμμανουήλ Κισσαμιτάκης - 15/09/2020

Κεφάλαιο 1: Εισαγωγικά

1.1 Εισαγωγή

Η παρούσα εργασία με θέμα: “Εξόρυξη Σημασιολογικών υπό-δικτύων από Κοινωνικά Δίκτυα” πραγματοποιήθηκε στα πλαίσια της διπλωματικής μου εργασίας. Η εργασία περιλαμβάνει τα παρακάτω τέσσερα μέρη

Στο πρώτο μέρος θα συμπεριλαμβάνεται το θεωρητικό πλαίσιο της εργασίας, όπου θα παρουσιαστούν σύντομα κάποιοι όροι ώστε να κατανοηθεί καλύτερα το πεδίο και οι έννοιες που θα χρησιμοποιηθούν στην παρούσα εργασία, με σκοπό να μπορεί να διαβαστεί και από έναν πιθανό αναγνώστη εκτός του πεδίου της Πληροφορικής.

Στο δεύτερο μέρος θα παρουσιαστεί η μεθοδολογία, όπου θα παρουσιαστούν αναλυτικά όλοι οι μέθοδοι που χρησιμοποιήθηκαν για να υλοποιηθεί η παρούσα εργασία. Αυτό περιλαμβάνει εργαλεία, τεχνικές και δημοφιλείς μεθόδους του επιστημονικού κλάδου.

Το τρίτο μέρος της εργασίας αφορά την έρευνα, όπου θα παρουσιαστούν δεδομένα από το dataset με γραφικές απεικονίσεις (πίνακες, διαγράμματα κ.α.) και θα αναλυθούν ως προς τη σημασία τους, με βάση τα δεδομένα που συλλέχθηκαν.

Στο τέταρτο και τελευταίο μέρος θα παρουσιαστούν τα συμπεράσματα που προέκυψαν για το παρόν dataset βάση της έρευνας του τρίτου κεφαλαίου, και θα αναλυθούν πιθανές χρήσεις αυτών σε μελλοντικές εφαρμογές είτε σε ακαδημαϊκά είτε σε εργασιακά περιβάλλοντα.

1.2 Προηγούμενες Προσεγγίσεις

Η ανάλυση κοινοτήτων μέσα από δεδομένα που προκύπτουν από το Twitter είναι κάτι που έχει γίνει στο παρελθόν, με παρόμοιες ή εντελώς διαφορετικές τεχνικές ανάλυσης.

Μερικές από αυτές περιλαμβάνουν ανάλυση κοινοτήτων μέσα από τη σχέση Κύριων Κόμβων (Prime Nodes) μέσα σε ένα δίκτυο, ανάλυση από Polarized Crowds (συνήθως αφορά ένα συγκεκριμένο ζήτημα όπως π.χ. ένα πολιτικό ζήτημα) σε Clusters Κοινοτήτων, ανάλυση μέσα από graph clusters και χρήση αλγόριθμου Clauset-Newman-Moore.

Επιπλέον, τεχνικές ανάλυσης βάση δημοσίων διαλόγων μέσα από post του Twitter, περιλαμβάνουν διαχωρισμό σε 6 βασικές δομές: divided, unified, fragmented, clustered, inward and outward hub και spoke structures.

Κεφάλαιο 2: Θεωρητικό Πλαίσιο

Περίληψη: Στο παρακάτω κεφάλαιο περιγράφονται σύντομα οι βασικοί όροι-λέξεις κλειδιά που θα χρησιμοποιηθούν στην εργασία (Εξόρυξη Δεδομένων, Κοινωνικά Δίκτυα, Bots, Hashtags, Gamification, Augmented Reality)

2.1 Εξόρυξη Δεδομένων

Εξόρυξη δεδομένων (ή ανακάλυψη γνώσης από βάσεις δεδομένων) (αγγλ. data mining) είναι η εξεύρεση μιας (ενδιαφέρουσας, αυτονόητης, μη προφανούς και πιθανόν χρήσιμης) πληροφορίας ή προτύπων από μεγάλες βάσεις δεδομένων ή κάποιο συγκεκριμένο dataset που παρουσιάζει ενδιαφέρον, με χρήση αλγορίθμων ομαδοποίησης ή κατηγοριοποίησης και των αρχών της στατιστικής, της τεχνητής νοημοσύνης, της μηχανικής μάθησης και των συστημάτων βάσεων δεδομένων. Στόχος της εξόρυξης δεδομένων είναι η πληροφορία που θα εξαχθεί και τα πρότυπα που θα προκύψουν να έχουν δομή κατανοητή προς τον άνθρωπο έτσι ώστε να τον βοηθήσουν να πάρει τις κατάλληλες αποφάσεις για την εκάστοτε κατάσταση.

Ο όρος εξόρυξη δεδομένων είναι μία έννοια που συνήθως παραπέμπει σε κάθε είδος φόρμας με μεγάλη ποσότητα δεδομένων ή επεξεργασία δεδομένων (συλλογή, εξαγωγή δεδομένων, warehouse, ανάλυση δεδομένων και στατιστικής) αλλά επίσης γενικεύεται σε κάθε είδος συστήματος υποστήριξης αποφάσεων συμπεριλαμβανομένου της τεχνητής νοημοσύνης, της εκμάθησης μηχανής και της επιχειρηματικής ευφυΐας. Στην ορθή χρήση του όρου η λέξη κλειδί είναι η ανακάλυψη, που ορίζεται ως η ανίχνευση κάτι καινούριου.

Ένα μεγάλο κομμάτι στην εξόρυξη δεδομένων αποτελούν τα εργαλεία που είναι διαθέσιμα για την εφαρμογή της και τον τρόπο που είναι δομημένα. Ένα από τα σημαντικότερα εργαλεία αποτελούν τα αρχεία csv (comma separated value). Τα csv αρχεία είναι απλά αρχεία κειμένου που μπορούν να ανοίξουν από έναν απλό επεξεργαστή κειμένου, κάτι που εξηγεί το μικρό τους μέγεθος, σημαντικός παράγοντας για διεργασίες που μπορεί να περιλαμβάνουν εκατομμύρια εγγραφές. Είναι δομημένα με τρόπο ώστε κάθε γραμμή του κειμένου να αποτελεί μία εγγραφή, και κάθε εγγραφή μπορεί να αποτελείται από πολλά πεδία, όπου το καθένα διαχωρίζεται με ένα κόμμα. Αυτή η δομή επιτρέπει σε αυτά τα αρχεία να έχουν τη δομή τραπέζιου, μπορούν δηλαδή να εμφανίζονται σε προγράμματα όπως το Microsoft Excel, σε μορφή που μπορεί να βοηθήσει κατά πολύ την οργάνωσή τους, είτε αρχικά μέσα στο αρχείο, είτε σε μία βάση δεδομένων.

2.2 Κοινωνικά Δίκτυα

Τα κοινωνικά δίκτυα είναι ένα σύνολο αλληλεπιδράσεων και διαπροσωπικών σχέσεων. Ο όρος σήμερα χρησιμοποιείται επίσης για να περιγράψει ιστοσελίδες οι οποίες επιτρέπουν την διεπαφή ανάμεσα στους χρήστες, πχ. με σχόλια, φωτογραφίες και προσωπικά μηνύματα. Μερικές από τις πιο γνωστές ιστοσελίδες είναι το Facebook, Twitter, Instagram και LinkedIn.

Οι ιστότοποι αυτοί αποτελούν εικονικές κοινότητες όπου οι χρήστες μπορούν να επικοινωνούν και να αναπτύσσουν επαφές μέσα από αυτές.

Ένα κοινωνικό δίκτυο είναι μια κοινωνική δομή που αποτελείται από ένα σύνολο παραγόντων, όπως άτομα ή οργανισμούς. Στο διαδίκτυο, τα κοινωνικά δίκτυα είναι μία πλατφόρμα που συντηρείται για την δημιουργία κοινωνικών σχέσεων μεταξύ των ανθρώπων, που συνήθως αποτελούν

ενεργά μέλη του κοινωνικού δικτύου, με κοινά ενδιαφέροντα ή δραστηριότητες.

Οι ιστότοποι κοινωνικής δικτύωσης είναι οργανωμένες ιστοσελίδες στο διαδίκτυο με περισσότερο ομαδοκεντρικό χαρακτήρα που παρέχουν, στην συντριπτική τους πλειοψηφία, μία σειρά από βασικές και δωρεάν υπηρεσίες όπως τη δημιουργία προφίλ, το ανέβασμα εικόνων και βίντεο, τον σχολιασμό σε ενέργειες που γίνονται από άλλα μέλη του δικτύου ή μίας ομάδας, την άμεση ανταλλαγή μηνυμάτων και πολλά άλλα.

2.3 Bots

Τα κοινωνικά δίκτυα αποτελούν από τη γέννησή τους πόλο έλξης για μικρά προγράμματα ή κομμάτια λογισμικού τα οποία κάνουν συγκεκριμένες λειτουργίες μέσα σε αυτά. Αυτά τα προγράμματα ονομάζονται bots. Οι βασικές τους κατηγορίες είναι δύο. Τα (νόμιμα) εμπορικά bots, και τα κακόβουλα bots.

Ο διαχωρισμός μας βοηθάει να ξεχωρίσουμε τι bots θα μπορούσαν να είναι αποδεκτά ως προς τη χρήση τους και ποια θα πρέπει να αποφεύγουν οι χρήστες.

Μερικά παραδείγματα εμπορικών bots στα social media αποτελούν bots τα οποία αυτοματοποιούν το ανέβασμα κάποιου υλικού που υποδεικνύει ο χρήστης, δίνουν πληροφορίες για την κίνηση ενός λογαριασμού (likes, επισκεψιμότητα κ.α.), chatbots που καθοδηγούν τους επισκέπτες αναλόγως τι ψάχνουν ή τους βοηθάνε να ολοκληρώσουν μία αγορά κ.α.

Από την άλλη πλευρά, υπάρχουν bots που μπορεί να αποστέλουν προσωπικά μηνύματα παρά τη θέληση του χρήστη, να διαγράφουν περιεχόμενο, να επηρεάζουν τάσεις, να κάνουν like, follow χωρίς να χρειάζεται η παρέμβαση του χρήστη κ.α

Σε ότι αφορά την επιρροή των τάσεων, μερικά bots μπορούν να προωθούν στους χρήστες πληροφορίες που δεν είναι αληθείς, σε τόσο μεγάλο βαθμό ώστε ένας μεγάλος αριθμός χρηστών να θεωρεί πως είναι αληθείς. Με αυτό τον τρόπο τα bots μπορούν να βοηθήσουν στη διάδοση fake news.

2.4 Hashtags

Hashtag είναι μία λέξη ή φράση που ακολουθείται μετά το σύμβολο «#» και χρησιμοποιείται σε Social Media sites και εφαρμογές, ειδικά στο Twitter, για τον εντοπισμό μηνυμάτων σε ένα συγκεκριμένο θέμα.

Στα Social Media και στις εφαρμογές που επιτρέπουν τη χρήση hashtag, οποιοσδήποτε χρήστης μπορεί να πληκτρολογήσει μέσα σε μια δημοσίευση ή σχόλιο ένα ή περισσότερα hashtags. Αυτό στη συνέχεια δημιουργεί αυτόματα έναν σύνδεσμο (link) που παραπέμπει τον χρήστη σε παρόμοιες δημοσιεύσεις που περιέχουν το συγκεκριμένο hashtag.

Στην ανάλυση των κοινωνικών δικτύων, τα hashtags έχουν μεγάλο σημασιολογικό βάρος, αφού αποτελούν ιδιαίτερο τρόπο φιλτραρίσματος και διάκρισης μεταξύ των δημοσιεύσεων των χρηστών. Η σημασιολογία αποτελεί βασική έννοια στην ανάλυση, διότι μας βοηθάει να κατανοήσουμε καλύτερα συμπεριφορές και δομές χρηστών.

2.5 Gamification

Gamification ονομάζεται η ενσωμάτωση μηχανισμών παιχνιδιού σε ένα περιβάλλον μη παιχνιδιού. Η πρακτική αυτή αξιοποιεί τα χαρακτηριστικά που κάνουν ένα game διασκεδαστικό, για να ενισχύσει το engagement του χρήστη, ανάλογα με το περιβάλλον στο οποίο αυτός βρίσκεται κάθε φορά. Ανάλογα λοιπόν με το αν ο χρήστης μιας gamified εφαρμογής αποτελεί μέλος ενός εργασιακού περιβάλλοντος, συμμετέχει στη μαθησιακή

διαδικασία ή αντιπροσωπεύει το κοινό μιας marketing στρατηγικής, η εμπειρία του ενισχύεται.

Ένα βασικό χαρακτηριστικό της στρατηγικής του παιχνιδιού είναι η επιβράβευση. Η συγκεκριμένη τεχνική εφαρμόζεται σε gamified εγχειρήματα μέσα από στόχους και επιβραβεύσεις. Κάθε φορά που ένας χρήστης ολοκληρώνει έναν στόχο, κερδίζει ένα badge. Τέτοια παραδείγματα επιβράβευσης εικονοποιούνται με καταμέτρηση πόντων, badges, progress bars, ακόμα και εικονικά νομίσματα. Η τεχνική αυτή ενθαρρύνει σημαντικά την εμπλοκή του χρήστη, καθώς καλύπτει την εσωτερική του ανάγκη για αναγνώριση της προσπάθειάς του, με αποτέλεσμα να αυξάνεται και ο βαθμός αφοσίωσής του.

2.6 Augmented Reality

Επαυξημένη πραγματικότητα (αγγλ. augmented reality) είναι η σε πραγματικό χρόνο άμεση ή έμμεση θέαση ενός φυσικού, πραγματικού περιβάλλοντος, του οποίου τα στοιχεία επαυξάνονται από στοιχεία αναπαραγώμενα από συσκευές υπολογιστών, όπως ήχος, βίντεο, γραφικά ή δεδομένα τοποθεσίας. Ο όρος εισήχθη το 1992 από τον Τομ Κάουντελ.

Τα τελευταία χρόνια έντονη είναι η εμφάνιση της επαυξημένης πραγματικότητας. Η τεχνολογία Augmented Reality ή αλλιώς επαυξημένη πραγματικότητα ή αλλιώς ενισχυμένη πραγματικότητα είναι μια τεχνολογία που χρησιμοποιείται στις κινητές συσκευές. Η τεχνολογία Augmented Reality επιτρέπει την ζωντανή προβολή ενός φυσικού περιβάλλοντος του οποίου όμως η πραγματικότητα είναι επαυξημένη με την προβολή πληροφοριών αλλά και εικονικών προσώπων ή χώρων σχεδιασμένων μέσα έναν ηλεκτρονικό υπολογιστή. Ο συνδυασμός της κάμερας με το σύστημα GPS ενός κινητού τηλεφώνου επιτρέπουν την προβολή επιπλέον πληροφοριών για ένα γεωγραφικό σημείο, διαμορφώνοντας ένα

επαυξημένο πληροφοριακά τελικό αποτέλεσμα. Παρέχονται πληροφορίες που περιλαμβάνουν κείμενα, ήχους και video και αφορούν ειδικά τη γεωγραφική θέση που βρίσκεται ο χρήστης και στοχεύει η κάμερα του. Οι προβολές δεδομένων είναι δυνατές είτε από τις οθόνες κινητών είτε από ειδικά γυαλιά προβολής Augmented Reality.

Με τεράστιο αριθμό εφαρμογών σε πολλές βιομηχανίες, ο ρόλος του Augmented Reality αρχίζει να γίνεται αυτός του απαραίτητου, και σε λίγα χρόνια αναμένεται να υπάρχει παντού. Μερικά παραδείγματα εφαρμογών περιλαμβάνουν απεικόνιση on-demand πληροφοριών σε συστήματα υγείας, βοήθεια στην εκπαίδευση προσωπικού μέσα από επαυξημένες εικόνες με μεγαλύτερη καθαρότητα πληροφορίας, εξομοιωτές πτήσεων, αλληλεπιδραστικά σενάρια για επισκέπτες μουσείων, χάρτες πόλεων ως επαυξημένα μοντέλα, ψηφιακή περιήγηση σε real estate για πιθανούς αγοραστές χωρίς να χρειαστεί να μεταβούν ως εκεί, μοντέλα αυτοκινήτων για καλύτερη σχεδίαση κ.α.

Κεφάλαιο 3: Μεθοδολογία

Περίληψη: Σε αυτό το κεφάλαιο παρουσιάζονται οι τρόποι συλλογής και ομαδοποίησης των δεδομένων (απόκτηση κλειδιών για Twitter App, εργαλεία όπως SQL Server και PyCharm, Office 365 και ομαδοποίηση σε πίνακες)

3.1 Μεθοδολογία

Στην παρούσα εργασία θα εξετάσω το dataset που συλλέχθηκε από το Twitter μέσω του Twitter Developers Site και θα προσπαθήσω να βγάλω συμπεράσματα για τους χρήστες και τα υποδίκτυα που δημιουργούνται βάση των hashtags #AugmentedReality και #Gamification.

Για την υλοποίηση της παραπάνω περιγραφής, χρησιμοποιήθηκε το IDE περιβάλλον PyCharm το οποίο περιγράφεται παρακάτω, όπου με τη χρήση της γλώσσας προγραμματισμού Python, δημιουργήθηκε ο βασικός Crawler ο οποίος αντλούσε δεδομένα από το Twitter χρησιμοποιώντας τα 4 κλειδιά της εφαρμογής που δημιουργήθηκε για την παρούσα εργασία όπως περιγράφεται και παρακάτω, και τα αποθήκευε σε ένα αρχείο .csv ώστε να είναι δυνατή η επεξεργασία τους αργότερα.

PyCharm:

Το PyCharm είναι ένα από τα κορυφαία Python IDE στον κόσμο. Υπάρχουν ποικίλοι λόγοι που το καθιστούν κορυφαίο περιβάλλον ανάπτυξης, συμπεριλαμβανομένου του γεγονότος ότι έχει αναπτυχθεί από την JetBrains, τον δημιουργό πίσω από το δημοφιλές IntelliJ IDEA IDE που

είναι ένα από τα 3 μεγαλύτερα Java IDE στον κόσμο. Έχοντας και δυνατότητα υποστήριξης για ανάπτυξη ιστοσελίδων, αξιοποιώντας το web framework Django είναι ένας ακόμη αξιόπιστος λόγος για να βρίσκεται εκεί που βρίσκεται.

Κάποια χαρακτηριστικά του PyCharm είναι:

- Περιλαμβάνει έναν graphical debugger
- Περιλαμβάνει integrated unit tester
- Περιλαμβάνει Integration support for version control systems (VCSs)
- Περιλαμβάνει υποστήριξη για data science με τη χρήση Anaconda
- Δίνει τη δυνατότητα δημιουργίας και διαχείρισης αρχείων κώδικα HTML, CSS και Javascript

Για επαγγελματικά περιβάλλοντα, στην professional έκδοσή του, το PyCharm περιλαμβάνει βιβλιοθήκες για Data Science και Machine Learning, όπως Matplotlib και SciPy, ωστόσο η συνήθης χρήση του είναι για ανάπτυξη web σελίδων και εφαρμογών.

Το PyCharm αποτελεί γενικότερα ένα ισχυρό εργαλείο για κάθε χρήση, από κάτι μικρό μέχρι επαγγελματικό στήσιμο εφαρμογών, περιλαμβάνει ωστόσο και έναν σχετικό βαθμό πολυπλοκότητας, χρειάζεται μία έστω μικρή εξοικείωση για να μπορέσει κάποιος να το χρησιμοποιήσει αποτελεσματικά. Λειτουργικά δεν έχει μεγάλες απαιτήσεις, και ένας σύγχρονος υπολογιστής μπορεί με ευκολία να το χρησιμοποιήσει.

Μετά τη συλλογή των tweet δεδομένων από το PyCharm, ακολούθησε η οργάνωση και διαχείρισή τους με τη χρήση του SQL Server Management Studio (SSMS), στο οποίο περνούσαν τα καθημερινά δεδομένα από τα .csv αρχεία σε μία βάση δεδομένων. Μέσα από τη βάση έγινε δυνατή η οργάνωση και η ομαδοποίηση των δεδομένων με τη χρήση της γλώσσας προγραμματισμού SQL, όπου μπόρεσα να ομαδοποιήσω τις μικρές κοινότητες βάση των λογαριασμών των χρηστών ανά ημερομηνία ανάρτησης. Το στήσιμο της βάσης έγινε με τη χρήση του SQL Server

Express 2017, όπου δημιουργήθηκε ένας τοπικός server και μέσα σε αυτόν η βάση δεδομένων.

SQL Server Management Studio (SSMS):

Το SQL Server Management Studio (SSMS) είναι μια εφαρμογή λογισμικού που ξεκίνησε για πρώτη φορά με τον Microsoft SQL Server 2005 και χρησιμοποιείται για τη διαμόρφωση και διαχείριση όλων των στοιχείων του Microsoft SQL Server, και αποτελεί ένα από τα δημοφιλέστερα εργαλεία της Microsoft.

Η βασική χρήση του συγκεκριμένου εργαλείου είναι η διαχείριση αντικειμένων μέσα σε βάσεις δεδομένων με τη χρήση του Object Explorer, ένα γραφικό εργαλείο που επιτρέπει στο χρήστη να περιηγηθεί, να επιλέξει και να ενεργήσει ελεύθερα σε οποιαδήποτε αντικείμενα του server.

Το SSMS έχει τη δυνατότητα να συνδέεται με μία πληθώρα άλλων εργαλείων της Microsoft, κάνοντας τη διασύνδεση και τη συνεργασία μεταξύ τεχνολογιών εξαιρετικά ευκολότερη και πιο αποτελεσματική. Η ευχέρεια χρήσης του συγκεκριμένου λογισμικού είναι υψηλή, αφού είναι αρκετά straightforward, και η χρήση SQL δεν είναι ο απαραίτητος τρόπος αλληλεπίδρασης του χρήστη με το λογισμικό, δίνοντας δυνατότητα και σε μη-προγραμματιστές να επωφεληθούν της χρήσης του.

Οι λειτουργικές απαιτήσεις του SSMS είναι αρκετά χαμηλές, με αποτέλεσμα να μπορεί να χρησιμοποιηθεί ακόμα και από παλαιότερους υπολογιστές και λειτουργικά συστήματα.

Ύστερα από τη συλλογή των δεδομένων στη βάση, έγινε η στατιστική ανάλυση των δεδομένων με τη χρήση του Microsoft Excel Office 365, και η οργάνωσή τους σε γραφήματα για την ορθή αναπαράσταση της πληροφορίας, ώστε να είναι σαφής η συσχέτιση μεταξύ των hashtags και των λογαριασμών που στηρίζουν τις άτυπες κοινότητες.

Λόγω της ιδιότητας του φοιτητή, υπήρξε η δυνατότητα να χρησιμοποιηθεί το Office 365 δωρεάν.

Παρακάτω παρουσιάζονται μερικά κομμάτια από τον κώδικα του Crawler, όπως και κομμάτια από τον κώδικα της οργάνωσης των δεδομένων μέσα στον SQL Server.

Κομμάτι από τον κώδικα του Crawler σε γλώσσα Python:

```
# Open/Create a file to append data
csvFile = open('DataAR.csv', 'a')
#Use csv Writer
csvWriter = csv.writer(csvFile)

for tweet in tweepy.Cursor(api.search, q="#AugmentedReality").items():
    csvWriter.writerow([tweet.user.screen_name, [tweet.created_at, tweet.text.encode('utf-8')]])
```

Εδώ ο κώδικας ανοίγει ένα .csv αρχείο το οποίο βρίσκεται στον φάκελο του Project, κι αν δεν βρει το αρχείο, τότε το δημιουργεί. Κατά τη συλλογή των δεδομένων, δημιουργήθηκαν 2 αρχεία, ένα για κάθε hashtag, και ο παραπάνω κώδικας έτρεχε κάθε μέρα ενημερώνοντας τα αρχεία, οπότε την πρώτη φορά που έτρεξε δημιούργησε τα αρχεία, και τις επόμενες φορές προσέθετε εγγραφές σε αυτά.

Παρακάτω παρουσιάζεται κομμάτι από τον κώδικα της οργάνωσης δεδομένων μέσα στο SQL Server με χρήση γλώσσας SQL:

```
select

    substring( substring([Message], CHARINDEX('@',[Message]), (CHARINDEX('
',[Message],CHARINDEX('@',[Message])) )) ,1,CHARINDEX(':', substring([Message],
CHARINDEX('@',[Message]), (CHARINDEX(' ',[Message],CHARINDEX('@',[Message])) )) ) ) as
Username

group by
    Username,
    Date
```

Το παραπάνω κομμάτι περιγράφει τις επεξεργασίες που γίνονταν στο αρχικό πεδίο του Message το οποίο περιέχει τις αρχικές αναρτήσεις όπως τις έφερνε ο crawler στα .csv αρχεία, ώστε να μπορέσουμε να βρούμε κάθε φορά μόνο το username του λογαριασμού που έκανε Retweet, ώστε να μπορέσουμε αργότερα να τα ταξινομήσουμε και να τα ομαδοποιήσουμε.

3.2 Συλλογή Δεδομένων και Χρήση Crawler

Δημιουργία και χρήση 4 κλειδιών για Twitter Crawler

Για τη σύνδεση της εφαρμογής με το Twitter ώστε να αποκτήσω τα δεδομένα των λογαριασμών, απαιτήθηκε η δημιουργία λογαριασμού στο Twitter Developers, όπου απαντήθηκαν μία σειρά από ερωτήσεις σχετικά με το τι είδους δεδομένα θα «τραβάει» η εφαρμογή, τι ακριβώς θα κάνει με αυτά και στο πλαίσιο ποιας δραστηριότητας ζητείται πρόσβαση σε αυτά.

Για να δημιουργηθούν τα κλειδιά, έπρεπε πρώτα να δημιουργήσω ένα App στο Twitter Developers, το οποίο λαμβάνει ένα μοναδικό ID, και ύστερα είχα το δικαίωμα να παράξω και να χρησιμοποιήσω τα εξής 4 σημαντικά κλειδιά:

- API key
- API secret key
- Access token
- Access token secret

Αυτά τα 4 μοναδικά κλειδιά είναι οι μοναδικοί κωδικοί της εφαρμογής μας και ήταν ο μόνος τρόπος να λάβω το δικαίωμα να ζητήσω δεδομένα του Twitter.

Τα δύο πρώτα κλειδιά αφορούν τη χρήση του API της εφαρμογής, δηλαδή του Application-Programming Interface, το οποίο δίνει πρόσβαση σε συγκεκριμένες βιβλιοθήκες, με τη χρήση των οποίων μπόρεσα να συνδεθώ με το Twitter και να διαχειριστώ τα δεδομένα του.

Τα δύο τελευταία κλειδιά δημιουργούνται μία μόνο φορά και σε περίπτωση που ο χρήστης χάσει την πρόσβασή του σε αυτά, πρέπει να γίνουν Revoke

μέσα από μια διαδικασία του Twitter Developers και ύστερα να ξαναγίνουν Generate ώστε να μπορεί ο χρήστης να ξαναχρησιμοποιήσει το App.

Κατά τη δημιουργία του εκάστοτε App, κλίθηκα να απαντήσω σε κάποια βασικά ερωτήματα προς το Twitter σχετικά με τη χρήση των δεδομένων τους. Ερωτήματα όπως το τι δεδομένα ακριβώς θα ήθελα να αιτηθώ από το Twitter, για πόσο μεγάλο χρονικό διάστημα θα χρειαστεί να έχω πρόσβαση στα συγκεκριμένα δεδομένα, υπό ποιο πρίσμα ή πλαίσιο αιτούμαι τα δεδομένα (π.χ. Διπλωματική εργασία).

Ύστερα από την απόκτηση των 4 κλειδιών της εφαρμογής, για να χρησιμοποιηθούν πρακτικά, τα συμπεριέλαβα μέσα στον κώδικα του Crawler. Έτσι, μπόρεσα να συνδέσω τον κώδικά με την εφαρμογή που έφτιαξα στο Developers περιβάλλον του Twitter, και να τρέξω τον κώδικα ώστε να τραβάει τα δεδομένα που χρησιμοποιήθηκαν στην παρούσα εργασία.

Η μόνη δυσκολία στα αλήθεια με το όλο εγχείρημα του να δημιουργήσω το App και να πάρω τα απαραίτητα για την εφαρμογή κλειδιά, ήταν η επιλογή ιστοσελίδας ή οποία αποτελεί υποχρεωτικό πεδίο στις ερωτήσεις σχετικά με την εφαρμογή και τη χρήση των δεδομένων. Δηλαδή, το Twitter απαιτεί υποχρεωτικά να δοθεί ένας σύνδεσμος από ένα site, είτε αυτό είναι του ίδιου του χρήστη, είτε κάποιας εταιρίας είτε ενός ιδρύματος. Στην προκειμένη περίπτωση, και αφού το Twitter το επιτρέπει, σαν σύνδεσμος χρησιμοποιήθηκε το ίδιο το Twitter, στη σελίδα με το προφίλ που δημιουργήθηκε στα πρώτα βήματα της διαδικασίας.

Οπότε η γενική ιδέα είναι πως αν είμαστε ένας Twitter Developer, για να ξεκινήσουμε να αντλούμε δεδομένα, δε χρειαζόμαστε μόνο έναν λογαριασμό αλλά και ένα App. Μπορούμε φυσικά να έχουμε παράλληλα πολλά Apps με το καθένα να έχει τα δικά του μοναδικά κλειδιά, ωστόσο το Twitter διατηρεί το δικαίωμα να μας εμποδίσει να τραβάμε δεδομένα από πολλαπλές εφαρμογές παράλληλα εάν αυτές έχουν παρόμοιο λόγο άντλησης και ζητάνε παρόμοια δεδομένα, έτσι ώστε να διαφυλάσσεται ως προς το σπαμ και την αλόγιστη παραχώρηση δεδομένων χρηστών του.

Συλλογή δεδομένων και περιορισμοί Twitter

Χρησιμοποιώντας τον crawler, κατάφερα να συλλέξω τα απαραίτητα δεδομένα για την έρευνα που αφορά την παρούσα εργασία. Ο σκοπός ήταν τα δεδομένα να αφορούν ένα διάστημα 3 εβδομάδων, ωστόσο, κάποια στιγμή το Twitter με σταμάτησε από το να αντλώ συνεχόμενα δεδομένα, έως ότου κατέληξε να μη μου επιτρέπει την πρόσβαση σε περαιτέρω πληροφορίες και να σταματήσει τη ροή των δεδομένων βάση των 4 κλειδιών που περιγράφηκαν ανωτέρω.

Δοκίμασα να δημιουργήσω νέα εφαρμογή και να ξαναζητήσω εκ νέου δεδομένα της ημέρα, όμως το μπλοκάρισμα παρέμεινε σε εφαρμογή οπότε και με απέτρεψε.

Αποτέλεσμα αυτού ήταν να υπάρχουν δεδομένα σε ένα span 2 εβδομάδων, τα οποία και χρησιμοποιήθηκαν για την διεκπεραίωση της έρευνας.

Ο Crawler

Για την αυτόματη συλλογή των δεδομένων, δημιούργησα σε γλώσσα προγραμματισμού Python 3.6 έναν crawler, ένα πρόγραμμα δηλαδή που αντλεί πληροφορία μέσα από το διαδίκτυο, βάση εντολών που λαμβάνει από τον προγραμματιστή του.

Η λειτουργία του crawler περιελάμβανε τη χρήση των 4 κλειδιών του Twitter που περιγράφηκαν σε προηγούμενο κεφάλαιο της εργασίας όπως ακριβώς φαίνεται και στο παρακάτω κομμάτι ψευδοκώδικα

```
access_token = "###1"
access_token_secret = "###2"
consumer_key = "###3"
consumer_secret = "###4"
```

όπου στην τιμή του κάθε κλειδιού, υπήρχε το αντίστοιχο κλειδί του Twitter. Για λόγους GDPR αλλά και για λόγους ασφαλείας, τα πραγματικά κλειδιά δεν αποτυπώνονται στη συγκεκριμένη εργασία.

Στη συνέχεια, ο Crawler χρησιμοποιεί τα παραπάνω κλειδιά ως authentication credentials και δημιουργεί ένα API αντικείμενο (object) με σκοπό να καλέσει αργότερα το Twitter API

```
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)
api = tweepy.API(auth, wait_on_rate_limit=True)
```

Έπειτα, το πρόγραμμα, με τη βοήθεια και της csv βιβλιοθήκης, ανοίγει το καθημερινό αρχείο σε mode κατάλληλο για append, για την εισαγωγή δεδομένων στο επιλεγμένο αρχείο, όπου θα αποθηκεύσει τα Tweet της ημέρας, και ακριβώς από κάτω δημιουργεί έναν csv writer ο οποίος θα γράφει τις εγγραφές που θα βρούμε μέσα στο αρχείο.

```
csvFile = open('GamificationDailyData.csv', 'a')
csvWriter = csv.writer(csvFile)
```

Ύστερα, ο crawler ξεκινάει να αντλεί τα δεδομένα με τη χρήση του αντικειμένου API που δημιούργησε παραπάνω, με την εντολή search μέσα στη συνάρτηση Cursor, όπου επιπλέον περνάμε σαν παραμέτρους το hashtag που θέλουμε να αναζητήσουμε (στην προκειμένη περίπτωση το #Gamification) και τη γλώσσα στην οποία επιθυμούμε να είναι γραμμένα τα tweets. Επέλεξα τα αγγλικά για δύο λόγους, ο ένας είναι πως είναι ήθελα έστω και δειγματοληπτικά να μπορώ να ελέγχω την ορθότητα των εγγραφών που αντλώ από το Twitter, και επειδή θα έβρισκα περισσότερα δεδομένα με τη συγκεκριμένη γλώσσα.

```
for tweet in tweepy.Cursor(api.search,
                             q="#Gamification",
                             lang="en",
                             ).items():
```

Τέλος, ο Crawler εκτυπώνει στην οθόνη μερικά στοιχεία από τα tweets που εγγράφει στο αρχείο csv, και στη συνέχεια περνάει στο csv αρχείο τις 3 κολώνες που επέλεξα να έχω σαν δομή για το dataset μου.

```
print(tweet.created_at, tweet.text)
csvWriter.writerow([tweet.user.screen_name, tweet.created_at, tweet.text.encode('utf-8')])
```

Αντίστοιχοι περιορισμοί με αυτούς του Twitter έχει και το Facebook για τα αντίστοιχα Apps. Το Facebook περιορίζει τα Apps να εκτελούν περισσότερα από 200 https calls ανά ώρα, εκτός και αν στην εφαρμογή εμπλέκονται περισσότεροι χρήστες. Τότε ο μέγιστος αριθμός υπολογίζεται από τη σχέση

$$\text{TotalCalls} = 200 * \text{NumberOfUsers}$$

To Dataset

Το Dataset που χρησιμοποιήθηκε στην παρούσα εργασία δημιουργήθηκε με τον τρόπο που περιγράφηκε παραπάνω. Οι συνολικές κολώνες που χρησιμοποιήθηκαν μέσα στο τελικό csv αρχείο με τις συνολικές εγγραφές που αποτέλεσε το dataset ήταν οι εξής 3:

- Username – Αλφαριθμητική μεταβλητή – Μέγεθος ως 4000 χαρακτήρες
- Date – Μεταβλητή Ημερομηνίας – Μέγεθος όσο προβλέπεται από τον SQL Server
- Message – Αλφαριθμητική μεταβλητή – Μέγεθος ως 4000 χαρακτήρες

Παρακάτω θα παρουσιάσω τις μορφές που είχε ανά διαστήματα κατά τη διάρκεια της εργασίας

Η αρχική μορφή ήταν ένα .csv αρχείο, με τρεις στήλες που προέρχονται από τον κώδικα του Crawler. Οι στήλες αυτές ήταν οι Username, Date και Message, όπου Username ήταν το όνομα του χρήστη που έκανε το tweet στην πλατφόρμα, το Date ήταν η ημερομηνία σε μορφή ημέρας της ανάρτησης και τέλος, το Message ήταν το αρχικό κείμενο της ανάρτησης, το οποίο περιελάμβανε όνομα χρήστη εάν το συγκεκριμένο post αποτελούσε retweet ή υπήρχε κάποιο mention, το βασικό σώμα του tweet και ειδικούς χαρακτήρες που μπορεί να υπήρχαν στην πλατφόρμα.

Ύστερα το .csv πέρασε στη βάση δεδομένων του SQL Server Express με τη χρήση του Data Import Wizard, και δημιουργώντας έναν νέο πίνακα για την αποθήκευσή τους. Από εκεί, με τη χρήση της γλώσσας προγραμματισμού SQL, ήταν δυνατή η επεξεργασία και η οργάνωση των δεδομένων σε μορφές καλύτερα κατανοητές και ομαδοποιημένες ώστε να γίνει δυνατή η ανάλυση των δεδομένων.

Για κάθε αρχείο .csv δημιουργήθηκε ξεχωριστός πίνακας ώστε να μπορούν τα δεδομένα να αναλυθούν επί μέρους και όχι μόνο καθολικά, έτσι ώστε να μπορέσει να υπάρξει διαχωρισμός τόσο των χρηστών που υποστηρίζουν το εκάστοτε hashtag, όσο και των δεδομένων για καθαρότητα στην ανάλυση.

Στη συνέχεια, τα δεδομένα από τους δύο πίνακες με τα hashtags ενώθηκαν σε έναν ενιαίο πίνακα που ονομάστηκε FinalData, και ο οποίος θα ήταν η βάση ώστε να έχω τα δεδομένα σε ένα μέρος και να μπορώ να τα επεξεργάζομαι ώστε να ξεκινήσουν να φαίνονται τα δεδομένα με μορφή που με βοηθάει να καταλήξω μέσα από αναλυτική λογική, σε συμπεράσματα για τον τρόπο που προκύπτουν, εμφανίζονται και σχετίζονται.

Στη συνέχεια έγινε ο διαχωρισμός του κάθε hashtag με ένα query ώστε να εμφανίζονται μαζί δίπλα δίπλα στην οθόνη, έτσι ώστε να είναι εμφανές αν κάποιος λογαριασμός έχει εγγραφή και στα δύο hashtags και για να γίνει αργότερα η ομαδοποίηση.

Εδώ έχω διατηρήσει την πρώτη στήλη όπως και στο αρχικό dataset, την ημερομηνία δηλαδή, αλλά έχω ομαδοποιήσει με την ημερομηνία, έτσι ώστε για κάθε ημερομηνία, να έχω τα συνολικά Tweets αλλά και τα επιμέρους για τον κάθε χρήστη.

Η στήλη UsernameClear, περιέχει το όνομα του λογαριασμού, «καθαρισμένο» από κενά κι άλλα σύμβολα, με σκοπό να είναι ακριβής η ομαδοποίηση με αυτό για τις γραφικές αναπαραστάσεις.

Η στήλη NumberOfTweets, περιέχει τα συνολικά Tweets του λογαριασμού ανά ημέρα, χωρίς να λαμβάνει υπόψιν το περιεχόμενο των tweets.

Η στήλη Gamification περιέχει τα συνολικά tweets ανά ημέρα και ανά λογαριασμό στα οποία βρέθηκε η λέξη Gamification.

Και τέλος, η στήλη AugmentedReality περιέχει τα συνολικά tweets ανά ημέρα και ανά λογαριασμό στα οποία βρέθηκε η λέξη AugmentedReality.

Κεφάλαιο 4: Έρευνα

Περίληψη: Στο παρόν κεφάλαιο παρουσιάζεται η έρευνα βάση των δεδομένων με τη χρήση πινάκων και διαγραμμάτων, ανάλυση ευρημάτων και πιθανές αιτίες.

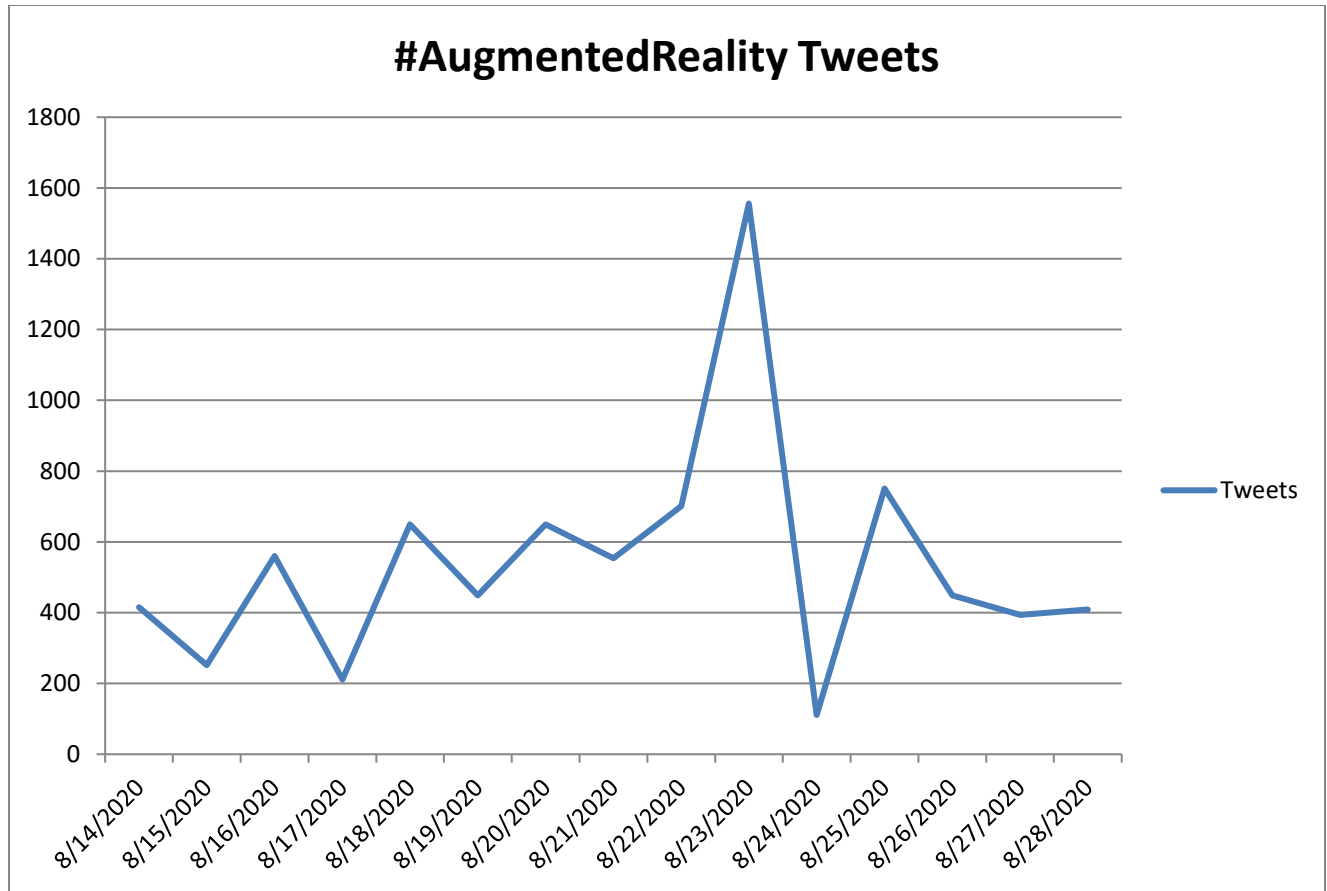
4.1 Έρευνα

Στο πλαίσιο της έρευνας για την παρούσα εργασία, πάρθηκαν σαν παράμετροι σύγκρισης δύο hashtags, τα οποία θεώρησα πως έχουν κάποια συσχέτιση, με σκοπό να δω ποιες άτυπες ομάδες χρηστών “στηρίζουν” τις δύο επιλογές, τόσο σε μεμονωμένο επίπεδο για το κάθε hashtag, όσο και για τον συνδυασμό των δύο ανά ημέρα, και για ένα διάστημα δύο εβδομάδων.

Για λόγους διασφάλισης ουδετερότητας και για λόγους GDPR, οι λογαριασμοί που χρησιμοποιήθηκαν στην έρευνα έχουν ανωνυμοποιηθεί και στο εξής θα αναφέρονται ως `account_xxxx` όπου `xxxx` ένας τετραψήφιος αριθμός.

Αριθμός Αναρτήσεων Ανά Ημέρα

Από τα παρακάτω γραφήματα, γίνεται αμέσως σαφές πως το hashtag `Augmented Reality` έχει μεγαλύτερη χρήση, εμφανίζεται πιο συχνά σε tweets και αποτελεί ένα δημοφιλές θέμα μέσα στον τεχνολογικό κόσμο του σήμερα.



Πίνακας 1: Augmented Reality Tweets Ανά Ημέρα – Πηγή: Συγγραφέας

Μπορεί μάλιστα να παρατηρηθεί μία ενδιαφέρουσα διαφορά μεταξύ του αριθμού των tweets στις 23-08-2020, με τα tweets να ξεπερνάνε τα 1500 σε αριθμό, και την αμέσως επόμενη ημέρα, το ίδιο hashtag να έχει κάτω από 200 posts. Εδώ θα εξετάσουμε πιθανές εξηγήσεις.

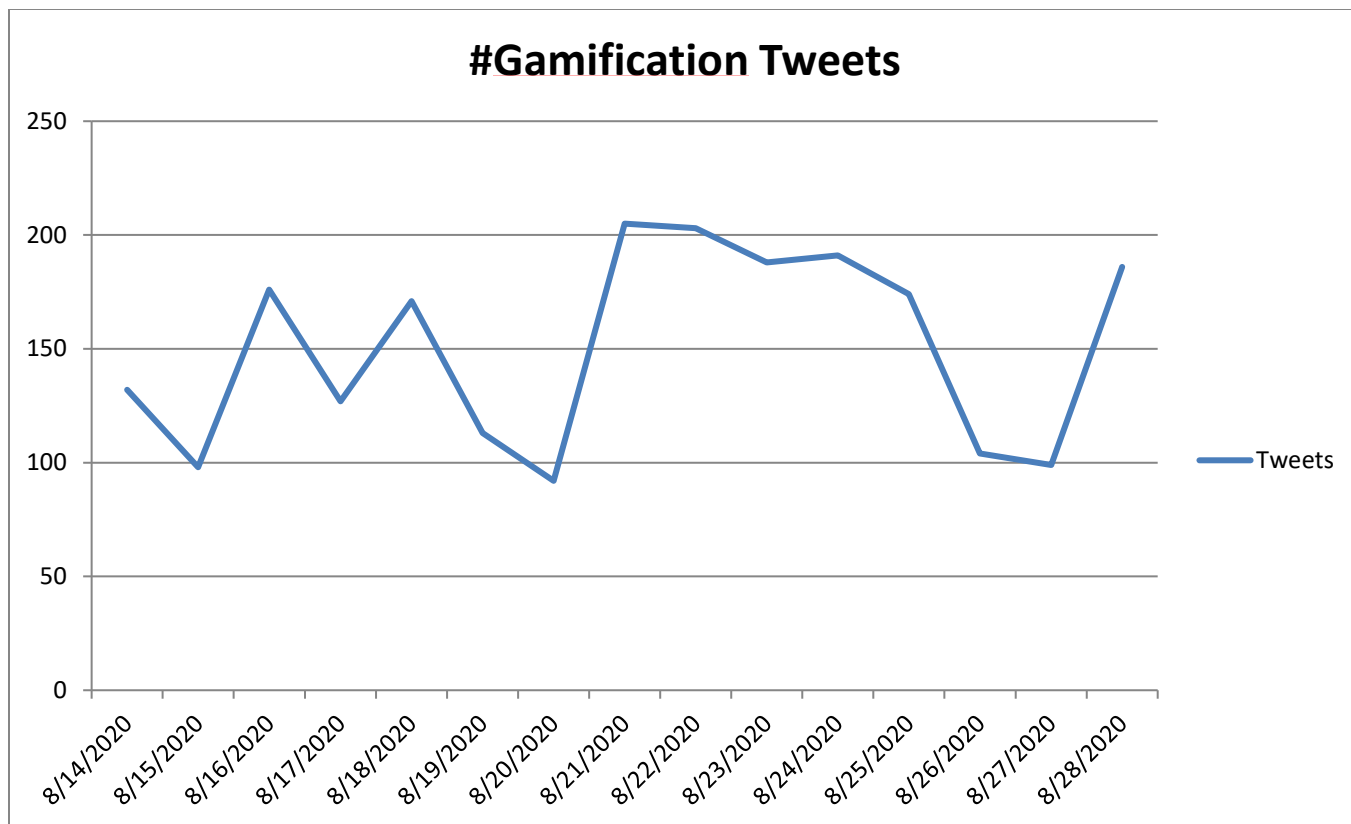
Είναι πιθανόν, η συγκεκριμένη ημέρα, 23-08-2020, να αποτελεί κάποια ημέρα εορτασμού για το συγκεκριμένο θέμα, όπως ας πούμε την πρώτη χρήση του όρου, ή η γέννηση κάποιου μεγάλου επιστήμονα που μπορεί να έχει συνδέσει το όνομά του με το θέμα.

Μία άλλη εξήγηση θα μπορούσε να είναι μία μεγάλη εκδήλωση, ένα conference δηλαδή, όπου βρέθηκαν πολλές μεγάλες εταιρίες, έκθεσε η εκάστοτε εταιρία το έργο της, είτε under development είτε παρουσίαση κάποιου μεγάλου breakthrough, οπότε και θα ακολουθούσε ένας καταιγισμός από tweets και retweets, και έτσι να πέτυχε το συγκεκριμένο hashtag να λάβει τόση προσοχή τη συγκεκριμένη ημέρα.

Μία εναλλακτική εξήγηση, αλλά όχι τόσο πιθανή, ειδικά σε σχέση με την ξαφνική πτώση, θα ήταν τα περισσότερα posts να έχουν γίνει από κάποιο bot, ένα πρόγραμμα δηλαδή, προγραμματισμένο να κάνει retweet posts με συγκεκριμένη θεματολογία είτε και με συγκεκριμένα hashtags, με αποτέλεσμα, εάν αυτό αληθεύει για τις 23-08-2020, να μπορούσε να εξηγήσει τον ασυνήθιστα μεγάλο αριθμό των tweets. Μπορεί δηλαδή κάποιος να ήθελε να τεστάρει εάν το bot του λειτουργεί σωστά ή να δοκιμάσει τις «ανοχές» του Twitter σε ότι αφορά τέτοιες πολιτικές. Δεν είναι απίθανο σενάριο, μιας και μερικοί χρήστες που εμφανίστηκαν στη στατιστική έρευνα που έγινε στην παρούσα εργασία να έχουν υψηλό αριθμό από posts, είχαν ονόματα που περιελάμβαναν μέσα τη λέξη bot, retweeter κ.α.

Όλα τα παραπάνω μένει να εξηγήσουν τι σχέση θα μπορούσαν να έχουν με την απότομη πτώση της επόμενης ημέρας στη χρήση του hashtag, κάτι που αφήνει τα bots εκτός, μιας και αυτό θα σήμαινε, πως για τη γενική κινητικότητα του hashtag ευθύνονται μικρά προγραμματάκια που κάνουν retweet, και όχι άνθρωποι του χώρου που μοιράζονται τη δουλειά τους, εκφράζουν μία σκέψη ή μοιράζονται μία γνώση.

Για το σενάριο του συνεδρίου, οι πιθανότητες δείχνουν να είναι λίγο υψηλότερες για να ισχύει, διότι θα μπορούσε να είναι λογικό, μετά από μία ημέρα έντονης κίνησης, πολλών αναρτήσεων και ενασχόλησης, να υπάρξει και μία μικρή χαλάρωση και απομάκρυνση από το συγκεκριμένο θέμα.



Πίνακας 2: Gamification Tweets Ανά Ημέρα – Πηγή: Συγγραφέας

Από την άλλη πλευρά, το hashtag Gamification δεν αποτελεί τόσο δυνατό πόλο έλξης για τις μεγάλες εταιρίες που συνήθως έχουν τη μερίδα του λέοντος στις αναρτήσεις σχετικά με θέματα τεχνολογίας στα κοινωνικά δίκτυα. Αυτό φαίνεται και από το γεγονός πως τις μέρες που το hashtag είναι πιο κινητικό, ακόμα κι εκείνες τις μέρες, οι φορές που χρησιμοποιήθηκε, είναι πολύ κοντά στη χαμηλότερη απόδοση που είχε στο διάστημα που μελετάμε το hashtag του Augmented Reality.

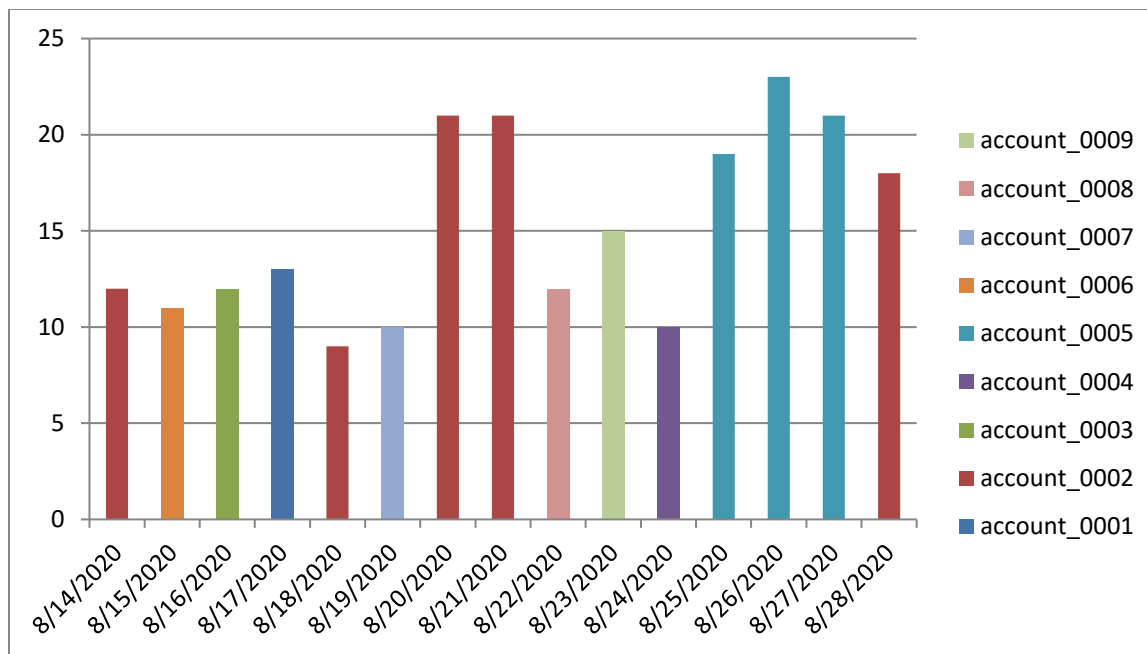
Αυτό πιθανώς οφείλεται όπως είπαμε στο γεγονός ότι σε ότι αφορά τη χρήση του Twitter, είναι σύνηθες τις περισσότερες αναρτήσεις, όπως και τα retweet και τις αναφορές, να τις έχουν οι μεγαλύτερες εταιρίες, έχοντας φήμη, άρα και αρκετούς ακολούθους, ειδικά για ένα θέμα που είναι ανερχόμενο όπως είναι το Augmented Reality.

Παρατηρείται πως η διασπορά του hashtag Gamification είναι σχετικά σταθερή, χωρίς πολλά скаμπανεβάσματα στις φορές που χρησιμοποιείται. Άρα μπορούμε να υποθέσουμε, πως τουλάχιστον για τη

χρονική περίοδο που κάναμε τη μελέτη μας, οι λογαριασμοί που «στηρίζουν» το συγκεκριμένο hashtag είναι σχετικά σταθεροί, αποτελούν δηλαδή μία μικρή άτυπη κοινότητα-υποδίκτυο, με σταθερή παρουσία στο Twitter, και πιθανώς κάποια σχέση μεταξύ τους.

Κορυφαίοι Λογαριασμοί Ανά Ημέρα

Παρακάτω φαίνεται η κίνηση ανά ημέρα για τους δύο πιο δραστήριους λογαριασμούς του κάθε hashtag, όπως επίσης και την εικόνα συγκεκριμένων λογαριασμών οι οποίοι παρουσιάζουν “αγκάθια” στην ενασχόλησή τους με το εκάστοτε hashtag στο Twitter.



Πίνακας 3: Augmented Reality Top Users Ανά Ημέρα – Πηγή: Συγγραφέας

Παραπάνω βλέπουμε το σχετικό πίνακα με τους κορυφαίους λογαριασμούς για το hashtag του Augmented Reality. Παρατηρούμε πως για τις 15 ημέρες που μελετάμε, έχουμε συνολικά 9 διαφορετικούς λογαριασμούς οι οποίοι βρέθηκαν να έχουν τα περισσότερα tweets κάποια

από τις ημέρες, δείχνοντας μας έτσι πως υπάρχουν αρκετοί λογαριασμοί που είναι ενεργοί στο Twitter και ασχολούνται με το συγκεκριμένο αντικείμενο. Βλέπουμε πως ο λογαριασμός που φαίνεται να βρίσκεται πιο συχνά στην κορυφή με τις συνολικές αναρτήσεις ανά ημέρα είναι ο χρήστης account_0002.

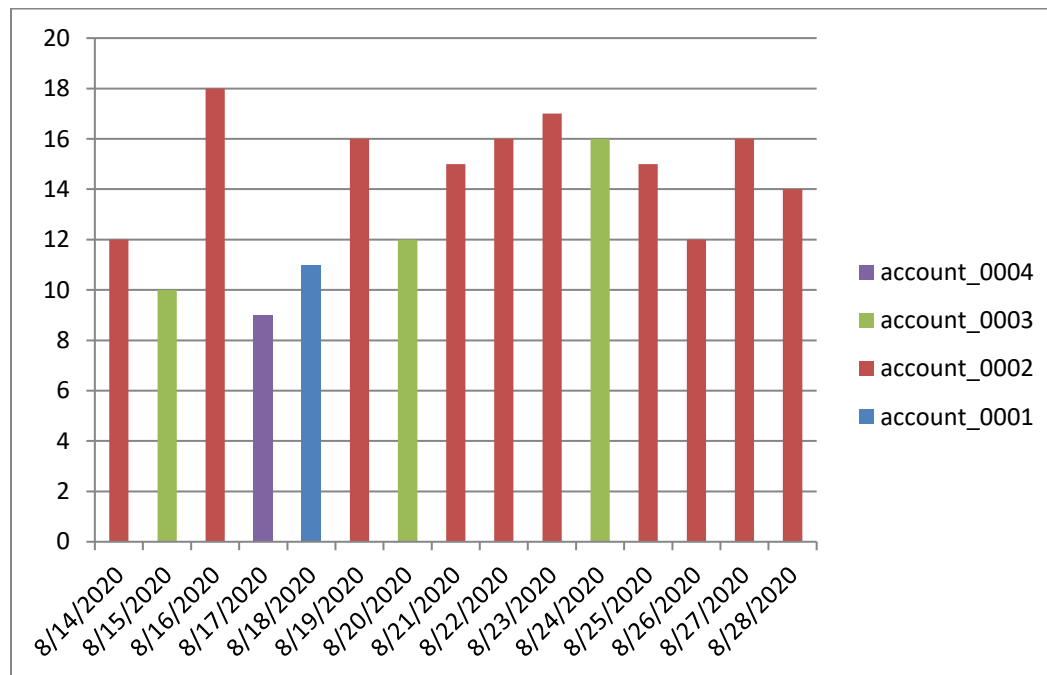
Το όνομα του συγκεκριμένου λογαριασμού περιλαμβάνει τη λέξη research, που μπορεί αμέσως να μας παραπέμψει στο ότι ο συγκεκριμένος λογαριασμός είναι πιθανό να ασχολείται είτε με έρευνα σαν προσωπική ενασχόληση, είτε να είναι διαμοιραστής πληροφοριών σχετικά με έρευνες που γίνονται πάνω στο συγκεκριμένο αντικείμενο.

Η παραπάνω παρατήρηση είναι άλλο ένα δείγμα στην ήδη υπάρχουσα κατάσταση που υπάρχει γύρω από το Augmented Reality, ότι δηλαδή είναι ένα πεδίο υπό ανάπτυξη, κυρίως ερευνητικό ακόμα, άρα και είναι λογικό ένας λογαριασμός έρευνας υπό μία έννοια, να είναι ενεργός με το συγκεκριμένο πεδίο.

Γενικότερα πάντως, το γεγονός πως φαίνονται διάφοροι λογαριασμοί να βρίσκονται στην πρώτη θέση με τις περισσότερες αναρτήσεις ανά ημέρα, χωρίς πολλές φορές να ξεπερνάνε τις 15 αναρτήσεις, μπορεί να μας οδηγήσει στην υπόθεση πως οι λογαριασμοί που ασχολούνται με το Augmented Reality, δεν είναι λογαριασμοί που είναι σταθερά ενεργοί στις αναρτήσεις τους, παρουσιάζουν δηλαδή «καρφιά» στη χρήση του. Μπορεί δηλαδή τη μία ημέρα να έχουν 15 αναρτήσεις και την άλλη να μην πραγματοποιούν καν 5 posts όλη μέρα.

Από όλα τα παραπάνω, μπορούμε να υποθέσουμε πως ο λόγος που παρουσιάζονται αυτές οι μεγάλες διακυμάνσεις στις αναφορές από τους συγκεκριμένους λογαριασμούς, είναι το γεγονός πως οι συγκεκριμένοι χρήστες, είτε βρίσκονται σε κάποια εταιρεία, είτε δουλεύουν ανεξάρτητα, χρησιμοποιούν ανά διαστήματα το Twitter για να κρατήσουν τους ακολούθους τους ενήμερους σχετικά με την τελευταία τους δουλειά. Είναι

δηλαδή ένα στοιχείο, πως εργάζονται πάνω σε κάτι συγκεκριμένο, και θέλουν να μοιραστούν την πρόοδο τους, τόσο για θέματα ενημέρωσης, όσο και για δείγμα ενεργής ενασχόλησης με το εκάστοτε αντικείμενο, ότι δηλαδή η εταιρία-χρήστης είναι ενεργός και με συνεχή πρόοδο στο συγκεκριμένο τεχνολογικό πεδίο.



Πίνακας 4: Gamification Top Users Ανά Ημέρα – Πηγή: Συγγραφέας

Για το hashtag Gamification από την άλλη πλευρά, βλέπουμε έναν λογαριασμό να έχει κυριαρχήσει στις αναρτήσεις ανά ημέρα, να βρίσκεται σχεδόν πάντα μπροστά από άλλους λογαριασμούς του αντίστοιχου πεδίου, και να βρίσκεται πρώτος στο εντυπωσιακό ποσοστό του 66%, αφού έχει τις περισσότερες αναρτήσεις τις 10 από τις 15 συνολικά ημέρες που μελετάμε.

Ο συγκεκριμένος λογαριασμός είναι ο account_0002, με μεγαλύτερο αριθμό ανά ημέρα να είναι οι 19 αναρτήσεις. Οι υπόλοιποι 3 λογαριασμοί που παρουσιάζονται παραπάνω, φαίνεται σαν να βρήκαν ένα «κενό» και κατάφεραν να περάσουν πρώτοι στις αναρτήσεις της ημέρας. Είναι πιθανό λοιπόν, ο συνήθης κορυφαίος λογαριασμός είτε να ήταν σχετικά ανενεργός εκείνες τις μέρες, είτε όντως να είχε απλώς λιγότερες αναρτήσεις. Αρχικά,

αυτό μπορεί να μας οδηγήσει με ασφάλεια στο συμπέρασμα πως ο συγκεκριμένος λογαριασμός δεν διαχειρίζεται από κάποιο bot, αλλά από έναν ή περισσότερους ανθρώπους, και είναι πιθανόν να είναι περισσότερο διαμοιραστή και επαναπροωθητής πληροφορίας, παρά παράγοντας αυθεντικής πληροφορίας.

Ακόμα και ο αριθμός των διαφορετικών λογαριασμών σε αυτό το πεδίο είναι σαφώς μικρότερος από τον αριθμό των λογαριασμών στο hashtag Augmented Reality, αφού για τους κορυφαίους ανά ημέρα, έχουμε μόλις 4 διαφορετικούς λογαριασμούς σε 15 ημέρες.

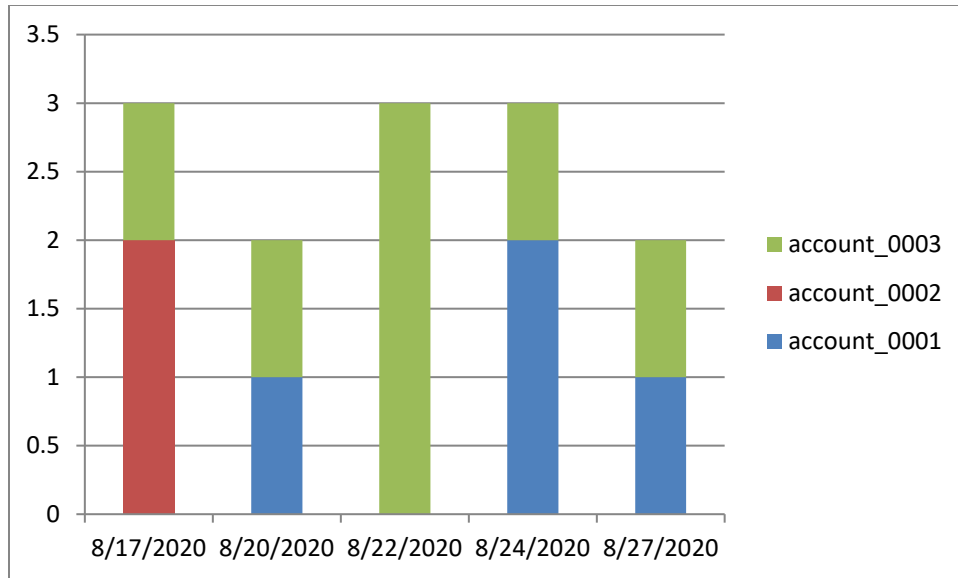
Το παραπάνω μας δείχνει πως είτε υπάρχει ένας ισχυρός πυρήνας ως προς τη χρήση του συγκεκριμένου hashtag, και άρα υπάρχουν συγκεκριμένοι λογαριασμοί οι οποίοι στηρίζουν το Gamification, είτε πρόκειται για λογαριασμούς οι οποίοι ασχολούνται αρκετά με την ενημέρωση για θέματα Gamification, άρα κάνουν retweet πολλά posts με το σχετικό θέμα, είτε ακόμα πρόκειται για λογαριασμούς οι οποίοι θέλουν να προωθήσουν ή να διαφημίσουν τη δουλειά τους.

Είναι συνήθης η χρήση του Gamification σε εκπαιδευτικά προγράμματα, τόσο σε σχολεία και άλλα εκπαιδευτικά ιδρύματα, όσο και σε εταιρίες όπου ενσωματώνονται στοιχεία Gamification ως μέρος της εκπαίδευσης προσωπικού. Για αυτό το λόγο θα μπορούσε κάποιος να χρησιμοποιήσει τις αναρτήσεις του Twitter και σαν εργαλείο διαφήμισης για τη δουλειά του, τα προϊόντα του κα, έτσι ώστε να τα προωθήσει προς πιθανούς αγοραστές (εταιρίες, ιδρύματα κτλ).

Χρήση Και Των Δύο Hashtags

Τι γίνεται όμως με τους χρήστες που χρησιμοποίησαν και τα δύο hashtags ;

Το παρακάτω γράφημα μας δίνει μία πλήρη εικόνα της κατάστασης. Μόλις τρεις χρήστες μέσα στο διάστημα που μελετήσαμε έκαναν χρήση και των δύο hashtags, και μάλιστα τις περισσότερες μέρες δεν είχαν ούτε μία ανάρτηση!



Πίνακας 5: Both Hashtags Top Users Ανά Ημέρα – Πηγή: Συγγραφέας

Σε ότι αφορά το παραπάνω δεδομένο, οι υποθέσεις που θα μπορούσαμε να κάνουμε ως προς το γιατί τα δύο hashtags, με όση δυναμική έχει το καθένα, δεν δείχνουν να συνυπάρχουν μέσα στα tweets των πιο ενεργών χρηστών του κάθε τομέα, αλλά και γενικά δε φαίνεται να βρίσκονται συχνά μαζί σε ένα tweet, παρά το γεγονός πως κάποιος θα υπέθετε πως είναι ταιριαστά θέματα, είναι πως είτε ο εκάστοτε χρήστης, έχοντας στο μυαλό του και έχοντας ετοιμάσει μία ανάρτηση η οποία θα δημοσιευθεί, είναι διστακτικός ως προς το τι απήχηση θα έχει προς τον υπόλοιπο τεχνολογικό κόσμο η χρήση των δύο hashtags μαζί, αν θα ήθελε δηλαδή να χρησιμοποιήσει πολλαπλά hashtags συμπεριλαμβανομένου και των δύο hashtags που μελετάμε στην παρούσα εργασία, ώστε να εκθέσει την ανάρτησή του σε όσο το δυνατόν περισσότερο κόσμο. Αυτό θα είχε σαν αποτέλεσμα πιθανώς να κατηγορηθεί για αναφορά σε άσχετο θέμα, με μόνο στόχο την ανάδειξη της ανάρτησης.

Ένας άλλος πιθανός λόγος, είναι πως λόγω εποχής που έγινε η καταγραφή των δεδομένων (14 Αυγούστου 2020 με 28 Αυγούστου 2020), πιθανώς όλος ο επιχειρησιακός κόσμος να ήταν λιγότερο ενεργός εκείνη την περίοδο, λόγω περιόδου διακοπών για παράδειγμα. Ωστόσο η συγκεκριμένη υπόθεση δεν μπορεί να στηριχτεί βάση των δεδομένων για

την ξεχωριστή χρήση των δύο hashtags, αφού είδαμε το Augmented Reality να σημειώνει πολύ μεγάλο αριθμό σε κάποιες μέρες, και ειδικά την ημερομηνία που αναλύσαμε παραπάνω.

Άλλη μία πιθανή εξήγηση για αυτή την απουσία χρήσης των δύο hashtag μαζί, θα μπορούσε να είναι η σχέση που έχει το Gamification με τις μεγάλες εταιρίες, οι οποίες ως επί το πλείστον, είναι οι εκφραστές του ανερχόμενου Augmented Reality, μιας και είναι κάτι το οποίο χρειάζεται μεγάλο κεφάλαιο ώστε να μελετηθεί και να αναπτυχθεί. Το Gamification από την άλλη πλευρά, έως τώρα, έχει δείξει να έχει βρει μία θέση σε άλλου είδους τεχνολογίες, κι άλλου είδους εφαρμογές, κυρίως στην εκπαίδευση, τόσο σε επίπεδο ιδρυμάτων όσο και σε επίπεδων εταιριών.

Βλέπουμε το Gamification να πρωταγωνιστεί σε τομείς όπως η εκπαίδευση, όπου με πολλά προγράμματα και δραστηριότητες να δανείζονται Gamification στοιχεία, καταφέρνουν οι εκπαιδευτικοί να κάνουν τα μαθήματά τους πιο ευέλικτα, και πολλές φορές πιο ευχάριστα για τους μαθητές, με κατάληξη να είναι και πιο αποτελεσματικά. Επίσης, βλέπουμε μια πολύ έντονη χρήση Gamification στοιχείων σε υπηρεσίες που απευθύνονται σε πολύ κόσμο που δεν αφορούν τη διασκέδαση, όπως, το LevelUp System που ξεκίνησε πρόσφατα το e-food.gr, μία σελίδα που διαχειρίζεται την ένωση καταστημάτων εστίασης με τους πελάτες της και παραγγελίες online, όπου με κάθε παραγγελία ο εκάστοτε πελάτης μαζεύει Exp Points, έχοντας σαν αποτέλεσμα να δίνουν ένα επιπλέον κίνητρο στους πελάτες να χρησιμοποιούν τη συγκεκριμένη υπηρεσία, αφού ανεβαίνοντας level, ο πελάτης μπαίνει στην κλήρωση όπου μπορεί να κερδίσει μία ποικιλία από βραβεία και δώρα. Παρόμοια τακτική ακολουθεί και το Booking.com.

Βλέπουμε λοιπόν πως το Gamification είναι κάτι που έχει ήδη βρει χρήσεις στην ανοικτή αγορά και μπορεί να απευθυνθεί στο ευρύ κοινό, το οποίο δεν είναι πιθανό να χρησιμοποιήσει μία πλατφόρμα όπως το Twitter και να χρησιμοποιήσει το hashtag #Gamification σε κάποια του ανάρτηση.

Από την άλλη, το Augmented Reality είναι μία τεχνολογία αναπτυσσόμενη, πολλά υποσχόμενη, που όμως δεν έχει βρει ακόμα τις χρήσεις που συζητήσαμε για το Gamification, ώστε να μπει στην αγορά και στο ευρύ κοινό και να καταξιωθεί σαν μέσο διαφοροποίησης, τουλάχιστον

σαν καταναλωτικό μέσο, με ελάχιστες εξαιρέσεις που είναι λίγο πολύ γνωστές σε όλους μας.

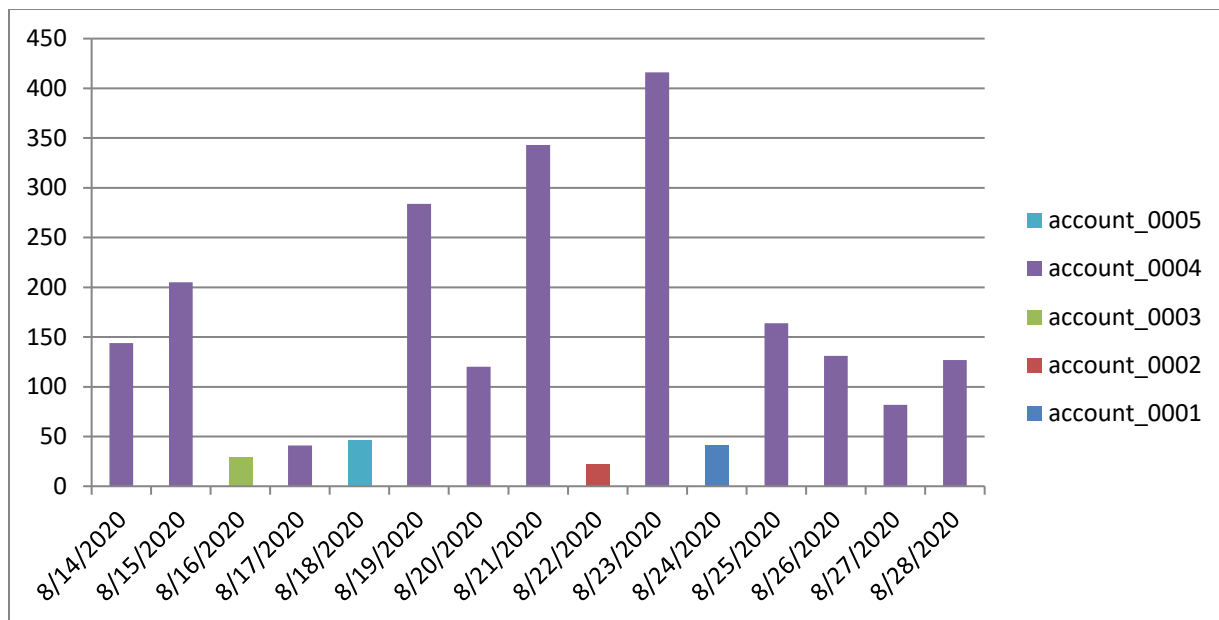
Βλέπουμε την κυριότερη χρήση του Augmented Reality προς το παρόν να είναι στη διαφήμιση, αλλά η λέξη κλειδί πάντα είναι η λέξη μέλλον. Αν παρατηρήσουμε, όταν κάποιος χρησιμοποιεί τις λέξεις Augmented Reality σε μία πρόταση, συνήθως είναι για να εξηγήσει γιατί είναι το μέλλον. Το μέλλον στη διαφήμιση, στο gaming, στις αγορές, στις διεπαφές κτλ.

Οπότε και η κοινή χρήση τους είναι κάτι που μπορεί να μην είναι πολύ πιθανό προς το παρόν, γιατί δύσκολα μία μεγάλη εταιρία, που είναι πιο πιθανό να ασχολείται ενεργά με το Augmented Reality, θα εντάξει κάπου ανάμεσα και το Gamification. Είναι πιθανό να χρησιμοποιεί το Gamification σε άλλους τομείς της εταιρίας, αλλά δύσκολο να ανεβάσει κάποια ανάρτηση στο Twitter που να συμπεριλαμβάνει και τα δύο.

Περισσότερα Retweets Ανά Ημέρα

Παρακάτω θα αναλύσω τους λογαριασμούς που είχαν τα περισσότερα retweets ανά ημέρα για το χρονικό διάστημα που μελετήσαμε, δηλαδή πόσες φορές έγινε retweet κάποια ανάρτησή τους.

Θα ξεκινήσω με το διάγραμμα που παρουσιάζει τους κορυφαίους retweeted λογαριασμούς για το hashtag Augmented Reality.



Πίνακας 6: Augmented Reality Top Users Retweets – Πηγή: Συγγραφέας

Όπως μπορούμε να παρατηρήσουμε από το παραπάνω διάγραμμα, υπάρχει εδώ ένας λογαριασμός ο οποίος είναι εξαιρετικά πιο «αγαπητός» στον κόσμο του Twitter Augmented Reality, με τεράστια διαφορά από τους υπόλοιπους λογαριασμούς. Είναι χαρακτηριστικό πως τις ημέρες που δεν έγινε retweet ο συγκεκριμένος λογαριασμός, την πρώτη θέση κατέλαβαν λογαριασμοί με λιγότερα από 50 retweets, ενώ ο λογαριασμός account_0004, έφτασε κάποια στιγμή να έχει πάνω από 400 retweets.

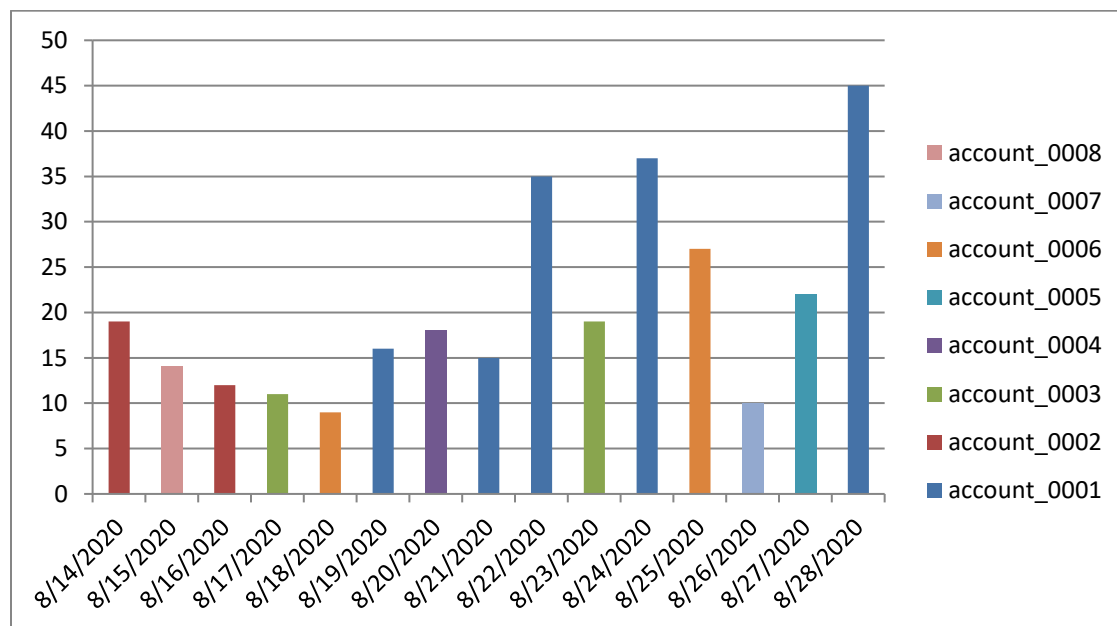
Το ενδιαφέρον σημείο του παραπάνω πίνακα, είναι το γεγονός πως την ημέρα που ο λογαριασμός account_0004 είχε τα περισσότερα retweets, είναι η ίδια μέρα μέσα στο διάστημα που μελετάμε, που το hashtag Augmented Reality είχε την περισσότερη κινητικότητα μέσα στο διάστημα. Θα μπορούσαμε λοιπόν να κάνουμε μία υπόθεση πως ίσως τα δύο γεγονότα συνδέονται μεταξύ τους, μιας και συμπίπτει η μέρα που σημείωσαν τον μεγαλύτερο αριθμό αναρτήσεών τους.

Επίσης, Παρατηρούμε πως την ακριβώς επόμενη ημέρα υπάρχει μία πτώση στις αναφορές του Augmented Reality, κάτι που στον πίνακα που είδαμε στην προηγούμενη ενότητα με τις χρήσεις ανά ημέρα φαινόταν ξεκάθαρα, και το σημαντικό στατιστικό εδώ, είναι πως έχει αλλάξει ο κορυφαίος χρήστης. Ωστόσο αυτό το γεγονός έχει επαναληφθεί και ίσως να

μην είναι ο κύριος παράγοντας που υπάρχει αυτή η απότομη άνοδος και πτώση του hashtag.

Θα μπορούσαμε λ.χ. να υποθέσουμε πως ο συγκεκριμένος λογαριασμός αποτελούσε ίσως μέρος του συνεδρίου, εάν πάμε με αυτό το σενάριο, και τις προηγούμενες ημέρες είχε αρκετές αναρτήσεις με το τι θα ακολουθούσε τη συγκεκριμένη ημέρα, μιας και βλέπουμε τα συνολικά retweets του να φθίνουν όσο περνάμε την ημερομηνία 23-08-2020.

Παρακάτω βλέπουμε τον αντίστοιχο πίνακα για τους κορυφαίους retweeted λογαριασμούς για το hashtag Gamification



Πίνακας 7: Gamification Top Users Retweets – Πηγή: Συγγραφέας

Παρατηρούμε πως σε αντίθεση με τον πίνακα των συνολικών αναρτήσεων ανά ημέρα, όπου οι λογαριασμοί ήταν μόλις 4, εδώ έχουμε ποικίλους λογαριασμούς που έγιναν retweet, συγκεκριμένα έχουμε 8 διαφορετικούς λογαριασμούς, την ίδια στιγμή που το ίδιο στατιστικό για το Augmented Reality εμφανίζεται να έχει την ακριβώς αντίθετη τάση, έχει δηλαδή ποικίλους λογαριασμούς στις αναρτήσεις ανά ημέρα και περιορισμένους λογαριασμούς που έγιναν retweet.

Επιπλέον, βλέπουμε ότι 3 από τους λογαριασμούς της λίστας με τους κορυφαίους, περιλαμβάνουν μέσα τη λέξη learn ή τη λέξη school, κάτι που

αμέσως μας παραπέμπει σε εκπαιδευτικές τεχνικές και στρατηγικές που αφορούν το Gamification. Άλλωστε, είναι πολύ συνηθής τακτική πλέον να χρησιμοποιούνται τεχνικές Gamification στην εκπαίδευση, είτε προσωπικού σε μία εταιρία είτε σε εκπαιδευτικά ιδρύματα, οπότε και δε μας κάνει εντύπωση το ότι βρίσκουμε δύο τέτοιους λογαριασμούς στους κορυφαίους που έγιναν retweeted.

Από τους δύο παραπάνω πίνακες, μπορούμε να κάνουμε δύο υποθέσεις. Η πρώτη υπόθεση είναι πως το αυθεντικό περιεχόμενο για το κάθε hashtag, διαφέρει ως προς το πόσο θα γίνει retweet, βλέποντας λιγότερους λογαριασμούς στο Augmented Reality να φτάνουν υψηλότερα νούμερα συνολικών αναρτήσεων, και η δεύτερη υπόθεση, είναι πως οι λογαριασμοί που ασχολούνται με το Gamification δείχνουν να δημιουργούν περισσότερο περιεχόμενο από τους λογαριασμούς του Augmented Reality. Αυτό μπορούμε να το υποθέσουμε από το γεγονός πως για το Gamification έχουμε 8 διαφορετικούς κορυφαίους λογαριασμούς, άρα και μεγαλύτερο όγκο αυθεντικής πληροφορίας, όπως λ.χ. προγράμματα εκπαίδευσης που διαφημίζονται και αναπαράγονται από κοντινούς λογαριασμούς, εργαλεία εκπαίδευσης κτλ.

Η παραπάνω παρατήρηση μας οδηγεί στην υπόθεση πως οι λογαριασμοί που ασχολούνται με το Augmented Reality, δείχνουν να έχουν μία τάση περισσότερο να αναπαράγουν αναρτήσεις που έχουν κάποιο ενδιαφέρον, σαν ένα δείγμα ενασχόλησης με το συγκεκριμένο τεχνολογικό πεδίο, και όχι να αναρτούν αυθεντικές, δικές τους αναρτήσεις με νέο περιεχόμενο. Και από την άλλη, έχουμε το Gamification, όπου οι χρήστες του φαίνεται να έχουν την τάση να αναρτούν δικό τους περιεχόμενο συχνότερα από τους χρήστες του Augmented Reality, πιθανώς και για εμπορικούς λόγους.

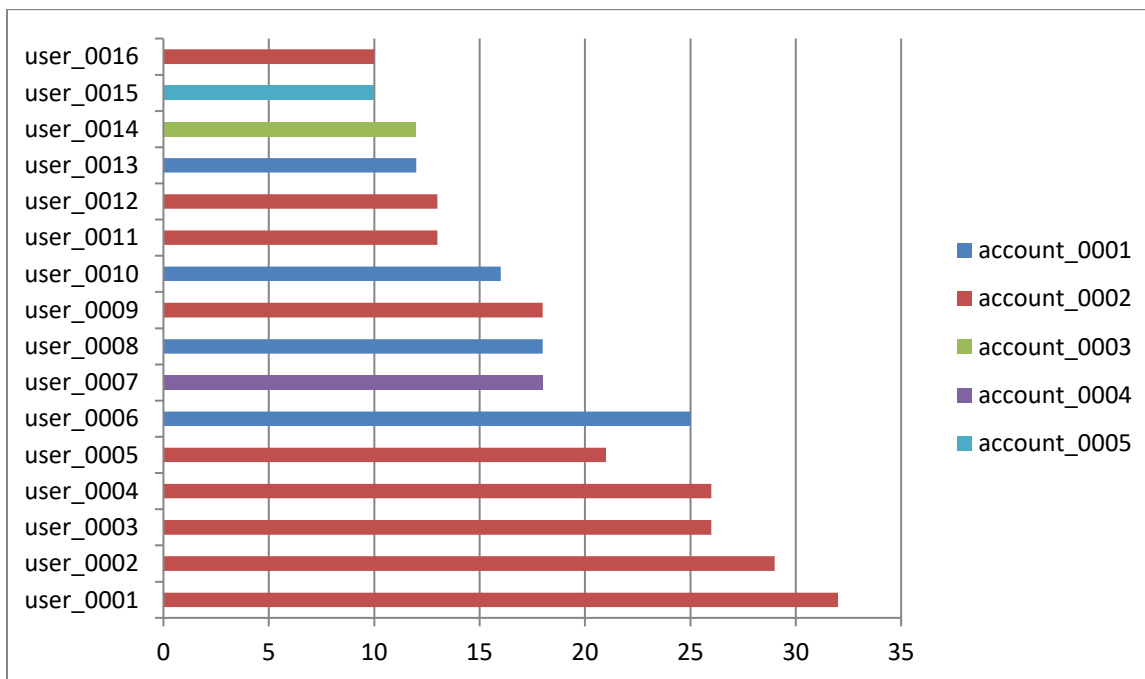
Η παραπάνω παρατήρηση μπορεί να οφείλεται σε διάφορους παράγοντες. Για παράδειγμα, είναι πιθανόν λόγω της τάσης που υπάρχει τελευταία με το Augmented Reality, να υπάρχουν και οι χρήστες που μπορεί να μην έχουν άμεση σχέση με το αντικείμενο, αλλά λόγω τάσης να το ακολουθούν και να κάνουν retweets, χωρίς ποτέ να ανεβάζουν δικό τους περιεχόμενο. Μία άλλη υπόθεση είναι πως λόγω του ότι το Augmented Reality είναι κατά κύριο λόγο ερευνητικό ακόμα, δεν υπάρχουν πολλά πράγματα να αναρτήσει κανείς, διότι στο Twitter δε θα αναρτήσει κάποιος

ένα paper, αλλά κάτι πιο χειροπιαστό, οπότε και δε βλέπουμε τόσες αναρτήσεις με αυτό το θέμα, αλλά επειδή υπάρχει πλήθος κόσμου που ακολουθεί το πεδίο, βλέπουμε και τέτοιους υψηλούς αριθμούς αναρτήσεων.

Σαν επιπρόσθετο στοιχείο στην παραπάνω παρατήρηση, έρχεται να προστεθεί η τάση που υπάρχει, οι χρήστες Gamification να χρησιμοποιούν πολλές φορές το Twitter για θέματα διαφήμισης κάποιου προσωπικού προϊόντος ή πιθανώς κάποιου που γνωρίζουν, κάτι δηλαδή, που κάποιος ακόλουθος θα μπορεί να κάνει κάτι περισσότερο από το να το παρατηρήσει απλώς.

Λογαριασμοί που έκαναν retweet

Παρακάτω θα παρουσιαστούν οι κορυφαίοι χρήστες που έκαναν retweet κάποια ανάρτηση άλλου χρήστη, ξεκινώντας με το hashtag του Augmented Reality

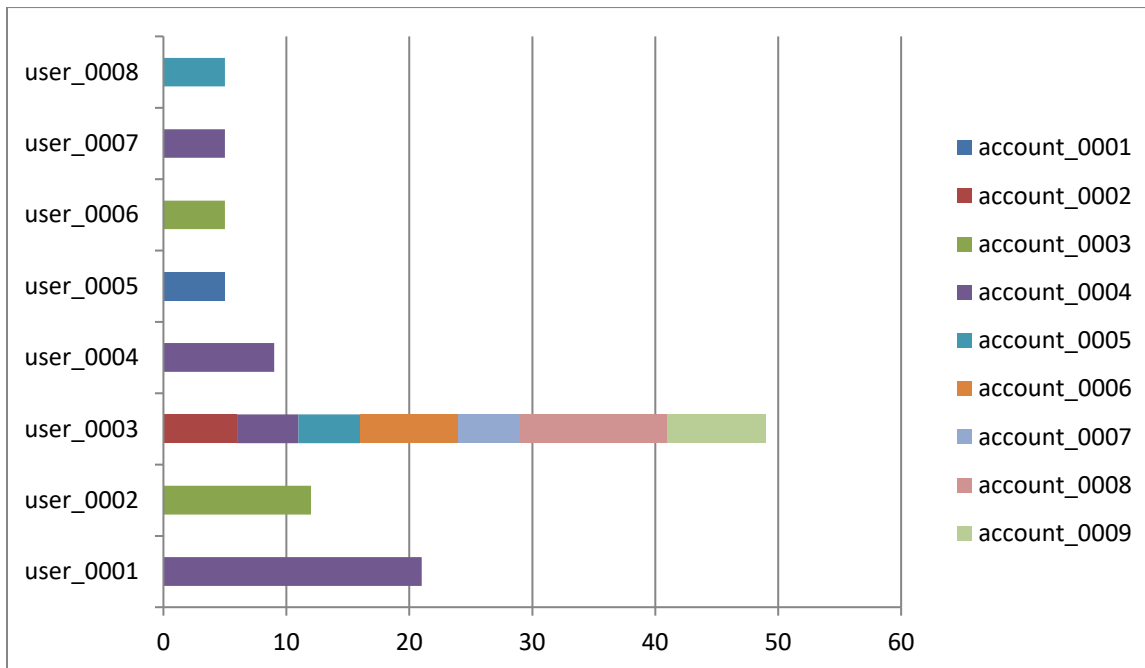


Πίνακας 8: Augmented Reality Top Users Retweeted – Πηγή: Συγγραφέας

Όπως βλέπουμε από το παραπάνω γράφημα, συνολικά, ο χρήστης που γινόταν συχνότερα retweet από τους παραπάνω ήταν ο χρήστης

account_0002, ο οποίος και βρισκόταν συχνότερα από όλους στην κορυφή της λίστας με τις περισσότερες αναρτήσεις ανά ημέρα, όμως το σημαντικό εδώ, είναι το γεγονός πως από τα ονόματα των χρηστών που τον έκαναν περισσότερο retweet, προκύπτει πως μεγάλο κομμάτι αυτών αποτελούν bots. Οι 5 τουλάχιστον χρήστες έχουν ξεκάθαρα το όνομα bot μέσα στο όνομα λογαριασμού τους, πράγμα που δεν αφήνει αμφιβολία, πως για το μεγαλύτερο κομμάτι των retweet του συγκεκριμένου χρήστη, ευθύνονται bots.

Το αντίστοιχο γράφημα για το Gamification, φαίνεται να μας δίνει μία τελείως διαφορετική εικόνα.



Πίνακας 9: Gamification Top Users Retweeted – Πηγή: Συγγραφέας

Από το παραπάνω γράφημα που αφορά το hashtag του Gamification, παρατηρούμε πως υπάρχει ένας χρήστης ο οποίος έχει κάνει retweet σε πολλούς διαφορετικούς χρήστες, κάτι το οποίο μας βάζει σε θέση να κάνουμε δύο υποθέσεις.

Η πρώτη υπόθεση είναι πως ο συγκεκριμένος χρήστης θα μπορούσε να είναι bot, κάτι το οποίο θα δικαιολογούσε τον αρκετά μεγάλο αριθμό από retweets, ωστόσο, επειδή τα bots του Augmented Reality δουλεύουν σαν

να είναι στοχευόμενα, δηλαδή επικεντρώνονται σε ένα λογαριασμό, ενώ ο συγκεκριμένος χρήστης αναφέρεται σε πολλούς λογαριασμούς, θα μπορούσε να μην είναι bot.

Η δεύτερη υπόθεση είναι πως ο παραπάνω ισχυρισμός θα μπορούσε να αληθεύει, ωστόσο το Gamification δεν έχει αρκετά μεγάλους αριθμούς για να μπορούμε να πούμε με μεγαλύτερη βεβαιότητα πως ο παραπάνω αναφερόμενος λογαριασμός δεν είναι bot, άρα και δεν μπορούμε να είμαστε σίγουροι επειδή οι αριθμοί μας ενδέχεται να επηρεάσουν την κρίση μας.

Κεφάλαιο 5: Επίλογος

5.1 Γενικές Παρατηρήσεις Και Συμπεράσματα

Οι γενικές παρατηρήσεις και τα συμπεράσματα της παρούσας διπλωματικής εργασίας θα παρουσιαστούν παρακάτω μαζί με μερικές υποθέσεις για την αξιολόγηση των αποτελεσμάτων.

Το πρώτο συμπέρασμα που προκύπτει από την έρευνα, είναι πως το hashtag Gamification, παρά το γεγονός πως έχει αρκετά λιγότερες αναρτήσεις ανά ημέρα σε σχέση με το Augmented Reality, φαίνεται να έχει περισσότερους πραγματικούς χρήστες να το στηρίζουν σε καθημερινή βάση με αναρτήσεις.

Παρατηρείται πως τα δύο hashtags, παρά το γεγονός πως φαίνονται αρκετά κοντινά, δε συναντούνται πολύ συχνά μαζί, είτε επειδή στις πρακτικές τους μορφές έχουν διαφορετικά πεδία εφαρμογής, είτε επειδή δεν έχει έρθει ακόμα η σωστή χρονική στιγμή να συμβαδίζουν. Είναι κάτι που ξέρουμε πως συμβαίνει εν μέρει και στο μέλλον θα συμβαίνει πιο έντονα, ωστόσο τη χρονική περίοδο της συγγραφής της συγκεκριμένης εργασίας, και με δεδομένο πως η περίοδος έρευνας αποτέλεσε τις μισές μέρες του μήνα Αύγουστου, του λιγότερου ενεργού μήνα επαγγελματικά, μπορεί απλώς να είναι σύμπτωση.

Επίσης, παρατηρείται πως κανένας από τους δυνατούς λογαριασμούς των δύο hashtags δεν ανήκει στη λίστα με τους λογαριασμούς που χρησιμοποίησαν και τα δύο hashtags σε κάποια ανάρτησή τους, πράγμα που οδηγεί στο συμπέρασμα πως δεν υπάρχει κάποιος μεγάλος εκφραστής της κοινής χρήσης των δύο πεδίων. Πάλι υπάρχει το ενδεχόμενο αυτό να προκύπτει και λόγω περιόδου έρευνας.

Σαν επόμενο συμπέρασμα, και βάση των ονομάτων των δυνατότερων λογαριασμών για το hashtag του Gamification, οι οποίοι περιείχαν τη λέξη learn ή school, προκύπτει πως το hashtag Gamification έχει σαν πιο ισχυρούς λογαριασμούς, χρήστες που φαίνεται πως ασχολούνται με εκπαίδευση και e-learning.

Η έρευνα της συγκεκριμένης εργασίας έγινε σε προσωπικό υπολογιστή, κάτι που σημαίνει πως η διαθέσιμη υπολογιστική δύναμη δεν ήταν μεγάλη. Τα δεδομένα που συλλέχτηκαν δεν ήταν μεγάλα σε αριθμό, και ο Crawler που ήταν το βασικό στοιχείο του συστήματος, χρειαζόταν μερικά λεπτά καθημερινά για την άντληση των εγγραφών. Θα μπορούσε λοιπόν η συγκεκριμένη έρευνα να κλιμακωθεί ώστε να περιλαμβάνει μεγαλύτερο αριθμό εγγραφών (πάντα με την άδεια του Twitter), ώστε να είναι περισσότεροι οι λογαριασμοί που θα εξεταστούν, ενδεχομένως σε μεγαλύτερο χρονικό διάστημα, και τα συμπεράσματα να είναι πιο καθαρά.

Σε ότι αφορά τις τεχνολογίες που χρησιμοποιήθηκαν, φαίνεται πως η Python και η βιβλιοθήκη Tweepy, ήταν τα κατάλληλα εργαλεία για την έρευνα, δεδομένης και της έκτασης αυτής. Αν η κλίμακα της έρευνας ήταν μεγαλύτερη, η Python είναι μία γλώσσα που θα μπορούσε να υποστηρίξει μία μεγαλύτερη κλίμακα, όχι ωστόσο με τη βοήθεια του PyCharm. Ενδεχομένως με μία πλατφόρμα ειδικά σχεδιασμένη για τέτοιες εργασίες.

Κεφάλαιο 6: Βιβλιογραφία

6.1 Βιβλιογραφικές Πηγές

- <https://developgreece.com/gamification/>
- <https://el.wikipedia.org/wiki/>
- <https://softweb.gr/el/community/blog/augmented-reality>
- <https://twitter.com/>
- <http://www.neoludica.eu/>
- <https://www.linkedin.com/>
- <https://www.gyrus.com/>
- <http://www.gamificationschoolhouse.com/>
- <https://www.adobe.com/insights/5-realworld-examples-of-augmented-reality-innovation.html>
- <https://www.apple.com/augmented-reality/>
- <https://www.plus.edu.gr/articles/ti-einai-to-gamification-kai-pos-efarmozetai/>
- <https://accelerole.com/10-tips-gamification-techniques/>
- <https://potion.social/en/blog/10-amazingly-successful-examples-of-gamification/>
- <https://thrive.kaiserpermanente.org/>
- <http://www.aranjomarketingsolutions.org/>
- <https://catchar.io/>
- <https://flamingofilter.co/>
- <https://babin-business-consulting.com/blog>
- <https://memberpress.com/using-social-media-bots-the-good-and-bad/>
- https://www.togaware.com/datamining/survivor/CSV_Option.html

- http://datamining.togaware.com/survivor/CSV_Data.html
- <https://www.toptal.com/python/twitter-data-mining-using-python>
- <https://developers.facebook.com/docs/graph-api/overview/rate-limiting/>
- <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7786908>
- <https://ieeexplore.ieee.org/abstract/document/9248469>
- <https://www.nature.com/articles/s42005-020-0340-4>
- <https://journals.sagepub.com/doi/pdf/10.1177/2056305117691545>
- <https://www.pewresearch.org/internet/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/>
- https://ieeexplore.ieee.org/abstract/document/5599247?casa_token=CwEVEoHNAN8AAAAA:tWkFd1uPkMRP8l2QjOb16E7NqIwMGRV5szZYGWoagZc2Q5D0hOHqD9qtXbEb4YKnO910abligQ
- <https://www.cogentoa.com/article/10.1080/23311983.2016.1171458>
- <https://www.jmir.org/2020/5/e19458/>
- <http://ece-research.unm.edu/ifis/papers/community-moore.pdf>
- <https://www.cogitatiopress.com/mediaandcommunication/article/view/2847>
- <https://dl.acm.org/doi/abs/10.1145/3320269.3384770>
- https://link.springer.com/chapter/10.1007/978-3-030-58219-7_13