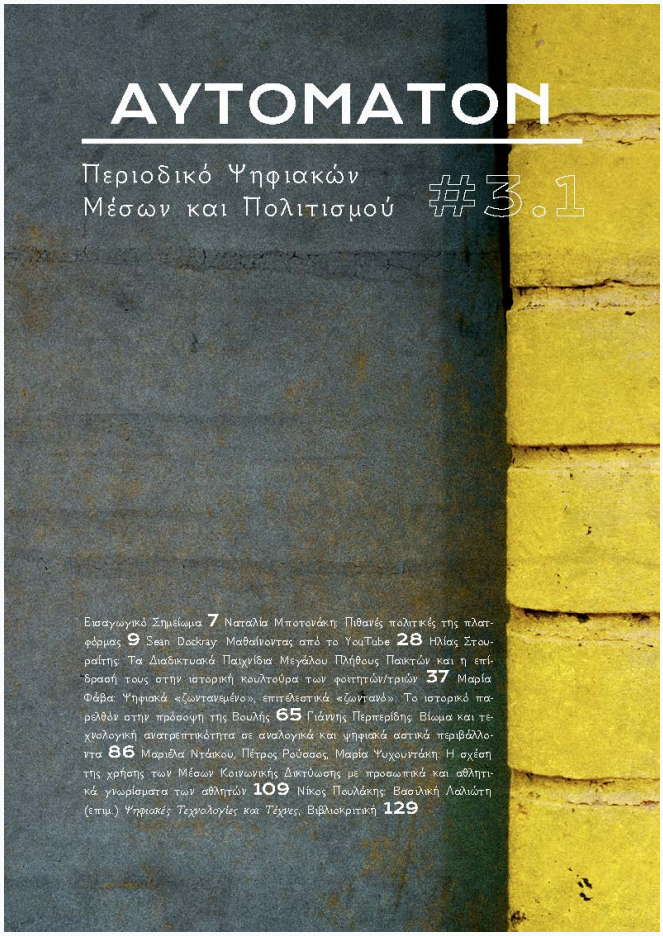
ePublishing

ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ
NATIONAL DOCUMENTATION CENTRE

Automaton: Journal of Digital Media and Culture

Vol 3, No 1 (2024)

Αυτόματον 3.1



Learning from YouTube

Sean Dockray

doi: 10.12681/automaton.38844

Copyright © 2024

This work is licensed under a Creative Commons Attribution 4.0.

To cite this article:

Dockray, S. (2024). Learning from YouTube. *Automaton: Journal of Digital Media and Culture*, *3*(1), 28–36. https://doi.org/10.12681/automaton.38844

https://epublishing.ekt.gr | **e-Publisher:** EKT | **Downloaded at:** 26/01/2026 19:30:57

ΑΥΤ
ΟΜΑ
ΤΟΝ

3(1): 28–36

Μαθαίνοντας από το YouTube*

Sean Dockray**

Περίληψη

Στο *Μαθαίνοντας από το YouTube*, ο Sean Dockray χρησιμοποιεί ένα επεξηγηματικό βίντεο στο YouTube για να παρουσιάσει το Audio Set, ένα πειραματικό πρόγραμμα που ξεκίνησε από τη «Sound Understanding Team» της Google. Το Audio Set είναι μια συλλογή μεγαλύτερη των δύο εκατομμυρίων μικρών ηχητικών αποσπασμάτων που εξάγονται από ανεξάρτητα βίντεο του YouTube και χρησιμοποιούνται ως δεδομένα για την εκπαίδευση των «συστημάτων βαθιάς μάθησης» της Google, με σκοπό να *επισημάνουν* εκατοντάδες ή χιλιάδες διαφορετικά ηχητικά συμβάντα κατ' αντιστοιχία με ηχογραφήσεις πραγματικού πεδίου, όπως ακριβώς οι ακροατές μπορούν να αναγνωρίσουν και να συσχετίσουν τους ήχους που ακούνε.

Λέξεις κλειδιά: YouTube, μηχανική μάθηση, Τεχνητή Νοημοσύνη, αλγόριθμοι ακρόασης, κοινωνία του ελέγχου

* Δημοσιεύτηκε στο J. Parker και J. Stern (επιμ.). 2019. *Eavesdropping: A Reader*. Μελβούρνη: Liquid Architecture, σσ. 100-113. Μετάφραση Γιάννης Πεδιώτης.

** Senior Lecturer, Department of Fine Art, Monash University, Melbourne, Sean.Dockray@monash.edu

Learning from YouTube

Sean Dockray*

Abstract

In *Learning from YouTube*, Dockray uses the form of an explanatory YouTube video to introduce the reader to Audio Set, an experimental project instigated by Google's 'Sound Understanding Team'. Audio Set is a collection of more than two million short sound clips extracted from YouTube videos that are used as data for training Google's "deep learning systems" to label hundreds or thousands of different sound events in real-world recordings just as human listeners can recognize and relate the sounds they hear.

Keywords: YouTube, machine-learning, Artificial Intelligence, listening algorithms, society of control.

* Senior Lecturer, Department of Fine Art, Monash University, Melbourne, Sean.Dockray@monash.edu.

Εκει που χάζευα στο YouTube, έπεσα πάνω σε ένα βίντεο που είχε τον τίτλο *Why Audio Analytics?*[1] Ήταν μια διαφήμιση της εταιρείας «Louroe Electronics» για ένα προϊόν που έμοιαζε με ανεστραμμένο ανιχνευτή καπνού, ικανό να αναλύει και να ανιχνεύει ήχους «μέσω προηγμένων αλγορίθμων». Το βίντεο παρουσίαζε διάφορα σενάρια με τα είδη των ήχων που η συσκευή μπορούσε να αναγνωρίσει: σπάσιμο τζαμιού τη νύχτα σε εκθεσιακό χώρο αντιπροσωπείας αυτοκινήτων, πυροβολισμό σε διάδρομο σχολείου, επιθετική συμπεριφορά σε δημόσιο χώρο. Ανεξάρτητα από τις διαφορές στα εν λόγω σενάρια –δεν σχετίζεται άραγε η επιθετική συμπεριφορά με την ερμηνεία και την κατανόηση του πλαισίου, παρά από το αν ακούγεται ή όχι ένας πυροβολισμός;– όλα τους είναι παραδείγματα *μηχανικής ακρόασης* και σηματοδοτούν τόσο μια διαφοροποίηση όσο και μια διεύρυνση της *αναγνώρισης ομιλίας*, του είδους που διαθέτουν τα Siri, Alexa και Google Assistant.

Στη μετατροπή ομιλίας-σε-κείμενο, ένα ηχητικό σήμα που περιέχει ανθρώπινη ομιλία μετατρέπεται σε κειμενική αναπαράσταση των λέξεων που εκφέρονται. Στην πιο γενική «αναγνώριση ήχου συμβάντος», όμως, *όλοι οι ήχοι* αντιστοιχίζονται σε περιγραφικές κατηγορίες. Για να δώσω ένα παράδειγμα για το τι σημαίνει κάτι τέτοιο, η Google έχει δημιουργήσει μια *οντολογία*[2] που ορίζει έναν εννοιολογικό χώρο 632 πιθανών κατηγοριών ήχου. Δεκατρείς από τις κατηγορίες αυτές αφορούν ήχους «ανθρώπινης φωνής», όπως «αναστεναγμός» ή «κλάμα, βογκητό». Μόνο μια από αυτές τις κατηγορίες αφορά την «ομιλία».

Παρά την ευρύτητα αυτού του φάσματος ήχων, οι πρώτες εμπορικές εφαρμογές της μηχανικής ακρόασης τείνουν να αφορούν τομείς όπως η ασφάλεια και η επιτήρηση. Η «Audio Analytic», που διατηρεί το δικό της ιδιόκτητο σύνολο ηχητικών δεδομένων με το όνομα Alexandria, αναπτύσσει ένα λογισμικό που εφαρμόζεται σε έξυπνες οικιακές συσκευές με σκοπό την ακρόαση συναγερμών ή διαρρήξεων. Η εταιρεία «Shooter Detection Systems» παρέχει τεχνολογία για την έγκαιρη ανίχνευση καταστάσεων ενεργού πυρός (active shooter situations), με το διαφημιστικό υλικό να ισχυρίζεται ότι το εν λόγω Guardian System «απομακρύνει τον "ανθρώπινο παράγοντα", ώστε να μην αφήνεται τίποτα στην ερμηνεία και να αποφεύγονται καθυστερήσεις που στοιχίζουν ακριβά».[3] Σε ανάλυση της τεχνολογίας αναγνώρισης προσώπου μετά την 11η Σεπτεμβρίου του 2001, η Γουέντι Τσαν (Wendy Hui Kyong Chun) παραθέτει ένα άλλο παράδειγμα χρήσης του ψηφιακού για την αποφυγή του ανθρώπινου παράγοντα. Η τεχνολογία «διορθώνει την οπτική υποκειμενική προκατάληψη, παρακάμπτοντας απάνθρωπα τον ορθολογισμό και την εξαγωγή συμπερασμάτων» (Chun 2006: 263), επιτυγχάνοντας την αναγνώριση τρομοκρατών μέσω συσχέτισης μοτίβων με δεδομέ-

1 'Why Audio Analytics?', YouTube, www.youtube.com/watch?v=fxg6ZfkgpM8, τελευταία πρόσβαση Φεβρουάριος 2024.

2 «Οντολογία» είναι ο όρος που χρησιμοποιούν οι μηχανικοί της Google, αντλώντας από μια πλούσια ιστορία χρήσης του όρου στην πληροφορική, με σκοπό να καθιερώσουν ένα «σταθερό, οργανωμένο λεξιλόγιο» για τη μοντελοποίηση κάποιας πτυχής του κόσμου. «Ontology (Information Science)», Wikipedia, https://en.wikipedia.org/wiki/Ontology_(information_science), τελευταία πρόσβαση Φεβρουάριος 2024.

3 «The Guardian Indoor Active Shooter Detection System», *Shooter Detection Systems*, http://shooterdetectionsystems.com/products/guardian/, τελευταία πρόσβαση Φεβρουάριος 2024.

να από κάμερες. Η Τσαν παραθέτει επίσης και το σχετικό διαφημιστικό υλικό: «Δεν υπάρχει πιθανότητα ανθρώπινου σφάλματος ή "φυλετικού χαρακτηρισμού", επειδή δεν απαιτείται ένας ανθρώπινος χειριστής για να επικεντρωθεί σε ένα συγκεκριμένο άτομο. Η κάμερα τα κάνει όλα αυτόματα» (ό.π. 262).[4]

Μέρος του επιχειρήματος που παρουσιάζει η «Louroe Electronics» στο βίντεό της είναι ότι με την ανίχνευση της επιθετικής συμπεριφοράς στο περιβάλλον μπορούμε να παρέμβουμε πριν καν ένα έγκλημα συμβεί. Αυτό το επιχείρημα, σε συνδυασμό με τις υποσχέσεις προγνωστικής ικανότητας της Τεχνητής Νοημοσύνης, υποδηλώνει ότι είναι δυνατόν να αλλάξουμε το μέλλον. Πέρα από την απλή αποτροπή ενός συγκεκριμένου συμβάντος, σε ένα τέτοιο προληπτικό σύστημα, το άτομο δεν θα αισθάνεται ποτέ άμεσα περιορισμένο, αλλά θα καθοδηγείται σε πλαίσια όπου η ανεπιθύμητη συμπεριφορά θα είναι ελάχιστα πιθανή (βλ. Rouvroy & Berns 2013). Με την ανίχνευση της επιθετικής συμπεριφοράς και του σπασίματος τζαμιών, η «Louroe Electronics» ξαναφέρνει στο νου τη θεωρία αστυνόμευσης των «σπασμένων παραθύρων», η οποία εξετάζει το πώς να «εντοπίζονται οι γειτονιές στο σημείο καμπής... όπου ένα παράθυρο αναμένεται ανά πάσα στιγμή να σπάσει και πρέπει να επισκευαστεί άμεσα αν δεν θέλουμε να σπάσουν και όλα τα άλλα». Την εποχή που εμφανίστηκε η θεωρία αυτή, το 1982, η αστυνομία δεν είχε «τρόπους συστηματικού εντοπισμού τέτοιων περιοχών»,[5] περιορισμός που υπερβαίνεται με την αλγοριθμική επιτήρηση. Η «Palantir Technologies» –ιδρυθείσα το 2004 από τον ελευθεριστή (libertarian) Πίτερ Τιλ (Peter Thiel) της Silicon Valley– είναι μια εταιρεία εξόρυξης δεδομένων στον τομέα των μυστικών πληροφοριών και της εθνικής ασφάλειας, που πρόσφατα διαπιστώθηκε ότι χρησιμοποιήθηκε από το αστυνομικό τμήμα της Νέας Ορλεάνης ως πεδίο δοκιμών για την προγνωστική αστυνόμευση. Η λογική της προληπτικής δράσης που εισηγείται η θεωρία των «σπασμένων παραθύρων» –η οποία υποθέτει ότι η αστική αταξία καλλιεργεί την πραγματική, σοβαρή εγκληματικότητα– ενισχύεται από αυτούς τους αλγόριθμους, σκοπός των οποίων δεν είναι άλλος από το να παρέμβουν «πριν ένα περιστατικό γενικευτεί σε βίαιη έκρηξη»[6], εγείροντας έτσι έναν φόβο και μια ανησυχία που τείνει να στρέφεται σε συγκεκριμένες ομάδες ανθρώπων, όπως «οι ζητιάνοι, οι μεθύστακες, οι τοξικομανείς, οι ατίθασοι έφηβοι, οι πόρνες, οι χασομέρηδες, οι ψυχικά διαταραγμένοι».[7]

Αυτό το αποκρινόμενο περιβάλλον ασφαλείας των συσκευών αισθητηριακής επιτήρησης προεικονίζεται στη φανταστική ηλεκτρονική κάρτα πρόσβασης του Φελίξ Γκουαταρί

4 Αν και δεν κατάφερα να βρω την ίδια πηγή από το *New York Times Magazine* που αναφέρει η Τσαν, εντόπισα την ίδια φράση σε άρθρο του CBS News, «Facial Recognition Technology May Screen for Terrorists», www.cbsnews.com/news/facial-recognition-technology-may-screen-for-terrorists/, τελευταία πρόσβαση Φεβρουάριος 2024.

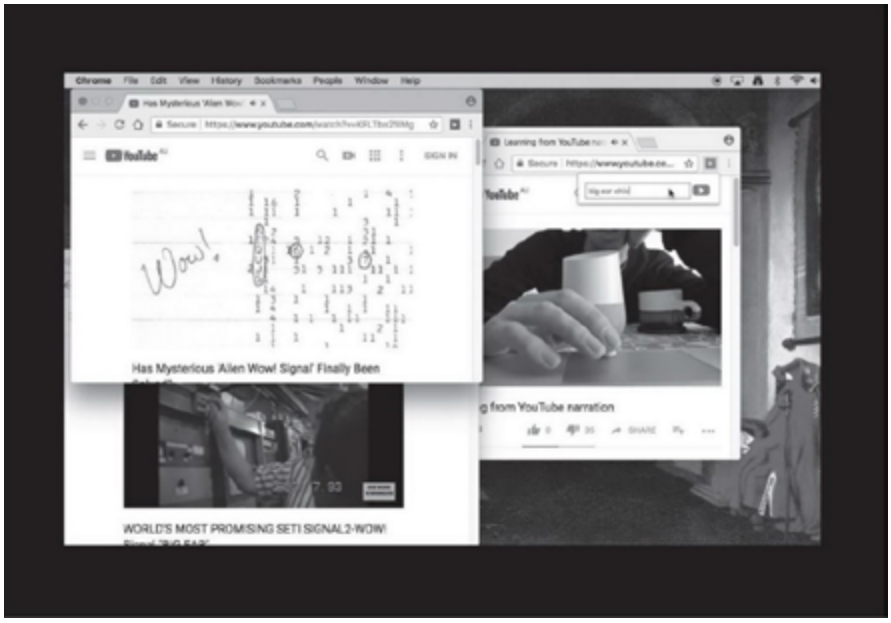
5 Βλ. *The Atlantic*, www.theatlantic.com/magazine/archive/1982/03/broken-windows/304465/, τελευταία πρόσβαση Φεβρουάριος 2024.

6 Βλ. 'Why Audio Analytics?', YouTube, www.youtube.com/watch?v=fxg6ZfkgpM8, τελευταία πρόσβαση Φεβρουάριος 2024.

7 Βλ. *The Atlantic*, www.theatlantic.com/magazine/archive/1982/03/broken-windows/304465/, τελευταία πρόσβαση Φεβρουάριος 2024.

(Felix Guattari), την οποία περιγράφει ο Ζιλ Ντελέζ (Gilles Deleuze) στο περίφημο *Υστερόγραφο για τις κοινωνίες του ελέγχου*. Ο μηχανισμός ελέγχου –η κάρτα– θα μπορεί να παρακολουθεί τη θέση του/της κατόχου της, εντοπίζοντας το άτομο στον χώρο και τον χρόνο, επιτρέποντας ή απαγορεύοντας την πρόσβασή του κάπου σύμφωνα με ένα σύνολο κανόνων, οι οποίοι και αυτοί δυνητικά αλλάζουν σε πραγματικό χρόνο. Ο Γουίλιαμ Μπάροουζ (William Burroughs), στον οποίο ο Ντελέζ (1998: 17) πιστώνει ότι «κατονόμασε το τέρας» του ελέγχου (Deleuze 1995: 179), παρατήρησε πως μια αίσθηση ελεύθερης βούλησης ήταν *απαραίτητη* προκειμένου να είναι ο έλεγχος αποτελεσματικός. Όταν «οι εργαζόμενοι έχουν γίνει σαν συσκευές κασετόφωνου» (Burroughs 1978: 38), τότε απλώς *χρησιμοποιούνται*, δεν ελέγχονται. Για τον Μπάροουζ, ο έλεγχος απαιτεί κάποιου είδους ατέλεια, ή ένα κενό μεταξύ του επιτηρητή και του επιτηρούμενου, κάτι που δεν συνιστά άμεση εκτέλεση των επιθυμιών του επιτηρητή, αλλά *περίπου*.

Όταν ο Βιλέμ Φλούσερ (Vilém Flusser) κάνει λόγο για έλεγχο, ωστόσο, δεν εννοεί έναν τρόπο λειτουργίας της εξουσίας, αλλά κάτι που έχει *χαθεί*, δηλαδή την ελευθερία: «Η κρίση της εξουσίας δεν έχει οδηγήσει σε χειραφέτηση της κοινωνίας, αλλά, καθώς επιτρέπει μια φαινομενική ελευθερία επιλογής, έχει οδηγήσει σε έναν ολοκληρωτισμό της κυβερνητικής που είναι μηχανικά προγραμματισμένος» (Flusser 2013: 86). Στη μεταβιομηχανική κοινωνία του Φλούσερ, οι άνθρωποι δεν εργάζονται, αλλά απασχολούνται. Η εργασία αφήνεται στις αυτόματες μηχανές που μετασχηματίζουν τον υλικό κόσμο σε αντικείμενα μαζικής παραγωγής. Το να είσαι απασχολούμενος σημαίνει να είσαι ένας *υπάλληλος* και *προγραμματιστής* που εμπλέκεται στην επεξεργασία συμβόλων, όπως ένας υπάλληλος γραφείου που στέλνει υπομνήματα και συμπληρώνει λογιστικά φύλλα. Σε επιφανειακό επίπεδο, οι προγραμματιστές είναι αυτοί που γράφουν τα προγράμματα και οι υπάλληλοι αυτοί που τα χρη-



Εικόνα 1: Sean Dockray *Learning from YouTube*, performance, 2018

σιμοποιούν. Αλλά αν κοιτάξουμε βαθύτερα, τα δύο αυτά συγχωνεύονται. Οι προγραμματιστές υπολογιστών προγραμματίζουν πατώντας κουμπιά για να χειριστούν σύμβολα. Κάθε επιλογή που γίνεται με κάθε πληκτρολόγηση, όμως, είναι μια επιλογή που γίνεται μέσα σε ένα άλλο πρόγραμμα, ένα *μεταπρόγραμμα*. «Και αυτή η παλινδρομική κίνηση από μέτα- σε μέτα-, από τον έναν προγραμματιστή στον άλλον και πάει λέγοντας, αποδεικνύεται ότι είναι άπειρη» (1999: 93). Σε κάθε επίπεδο, οι προγραμματιστές είναι συγχρόνως υπάλληλοι, και αντιστρόφως. Αυτό το παράδοξο βρίσκει την επιτομή του στις πλατφόρμες των κοινωνικών μέσων, οι οποίες είναι σκηνές μαζικής εξατομικευμένης αυτοέκφρασης και, την ίδια στιγμή, εξαιρετικά σεναριογραφημένα εθιστικά πλαίσια που επιβάλλουν τη συμμετοχή.

Οι πλατφόρμες των κοινωνικών μέσων λειτουργούν τόσο ως χώροι ελέγχου όσο και ως μηχανές συγκέντρωσης δεδομένων που δύναται να χρησιμοποιηθούν για μελλοντικές μορφές ελέγχου. Το 2017, μέλη της ομάδας Sound and Video Understanding[8] της «Google» ανακοίνωσαν το *Audio Set*, ένα σύνολο δεδομένων από δύο εκατομμύρια βίντεο του YouTube, που φιλοδοξούσε να «ενθαρρύνει ουσιαστικά» (Gemmeke κ.ά. 2017: 776) την ανάπτυξη των αλγορίθμων μηχανικής ακρόασης. Η εν λόγω ανακοίνωση δεν συνοδεύτηκε από πολλές τυμπανοκρουσίες, διότι, αντί για τα μέσα ενημέρωσης, δημοσιεύθηκε στο ερευνητικό ιστολόγιο της «Google» και σε ένα ακαδημαϊκό άρθρο. Ήταν –και εξακολουθεί να είναι, τη στιγμή που γράφεται το κείμενο– μια εσωτερική εξέλιξη, που ενδιαφέρει κυρίως τους προγραμματιστές και τους λάτρεις της μηχανικής μάθησης. Και το σημαντικότερο είναι ότι *προηγείται* χρονικά οποιασδήποτε συγκεκριμένης εφαρμογής τεχνητής νοημοσύνης ή μηχανικής μάθησης στη βάση αυτού του συνόλου δεδομένων, είτε πρόκειται για την αυτοματοποίηση της οικίας, την παρακολούθηση του χώρου εργασίας, είτε για την αυτοματοποιημένη αστυνόμευση. Γεγονός που κάνει το *Audio Set* να φαντάζει προ-πολιτικό, απαλλαγμένο από την ανισότητα και την προκατάληψη, που φαίνεται να προκύπτουν μόνο σε περίπτωση που διαπιστωθεί ότι ένα αυτοματοποιημένο σύστημα έχει, για παράδειγμα, στοχοποιήσει κάποιον μαύρο.

Τα βίντεο στο *Audio Set* έχουν επιλεγεί τυχαία, οπότε είναι απίθανο να γνωρίζουν οι χρήστες που τα ανέβασαν ότι το περιεχόμενό τους χρησιμοποιείται με τέτοιο τρόπο. Είναι, ωστόσο, εξίσου απίθανο να τους ενδιαφέρει. Ποιος/α μπορεί να γνωρίζει με πόσους τρόπους έχει αναλυθεί ένα βίντεο προκειμένου να χρησιμοποιηθεί για προτάσεις και διαφημίσεις; Στο YouTube, τα βίντεο ζουν διπλή ζωή, αποτελούν ταυτόχρονα ψυχαγωγία για το ανθρώπινο κοινό *και* δεδομένα για το αλγοριθμικό κοινό, ενώ η συνεχής εφεύρεση νέων αλγορίθμων που είναι ικανοί να τα παρακολουθούν με νέους τρόπους καθιστά τα παλιά βίντεο πάλι καινούργια. Είναι γεγονός, άλλωστε, ότι για την ύπαρξη κοιτασμάτων ουρανίου στα βουνά γνωρίζαμε αιώνες προτού αρχίσουμε να τα εξορύσσουμε για την παραγωγή ράδιου. Τα δεδομένα θα συγκεντρώνονται σε φάρμες διακομιστών για χρόνια προτού αξιοποιηθούν με κερδοφόρο τρόπο.

Η μάζα των βίντεο από το YouTube στο *Audio Set* μπορεί να συγκριθεί με το κροπαρισμένο εξώφυλλο του μοντέλου του *Playboy*, Λένα Σόντερμπεργκ (Lena Söderberg), το οποίο χρησιμοποιήθηκε ως δοκιμαστική εικόνα για την έρευνα της ψηφιακής συμπίεσης εικό-

8 Μέρος της Google Machine-Perception Team.

Εικόνα 2: Κέντρο δεδομένων της Microsoft στο Δουβλίνο. Εικόνα: Google Maps

νων και έκτοτε έχει καθιερωθεί ως βιομηχανικό πρότυπο για τη δοκιμή αλγορίθμων απεικόνισης από τη φόρμα του JPEG έως τα εφέ του Photoshop. Στην εποχή της μηχανικής μάθησης, η δοκιμαστική εικόνα αποτελείται από τεράστια σύνολα δεδομένων. Προς το τέλος του 2016, η «Google» ανακοίνωσε το YouTube-8M, ένα σύνολο δεδομένων οκτώ εκατομμυρίων κατηγοριοποιημένων βίντεο από το YouTube (του οποίου το εν λόγω *Audio Set* αποτελεί υποσύνολο), με σκοπό να επιταχυνθεί η πρόοδος στη μηχανική μάθηση και τη μηχανική αντίληψη.[9] Όχι πολύ αργότερα, ο Διευθύνων Σύμβουλος της «Google», Σούνταρ Πιτσάι (Sundar Pichai), άλλαξε την εταιρική στρατηγική θέτοντας την «Τεχνητή Νοημοσύνη πρώτα». Ξαφνικά, η απόφαση της «Google» να αποκτήσει το YouTube το 2006 φάνηκε να αφορά λιγότερο το ανθρώπινο κοινό και περισσότερο το αλγοριθμικό. Ήταν εκείνη τη στιγμή που οι δημιουργοί των βίντεο αυτοματοποιήθηκαν αναδρομικά, έγιναν συμμέτοχοι σε έναν πληθοπορισμό (crowdsource) χωρίς καν να το αντιληφθούν, μετατράπηκαν σε μνήμες ενός αλ-

9 Βλ. «Announcing YouTube-8M: A Large and Diverse Labeled Video Dataset for Video Understanding Research», *Google Research Blog*, https://research.googleblog.com/2016/09/announcing-youtube-8m-large-and-diverse.html, τελευταία πρόσβαση Φεβρουάριος 2024.

γορίθμου με άγνωστη πολιτική. Με τις εταιρικές της μηχανορραφίες η «Google» ανασχεδιάζει το παρελθόν, και με τις προβλεπτικές ικανότητες της Τεχνητής Νοημοσύνης, το μέλλον.

Το γεγονός ότι τα διάφορα είδη αυτοματοποίησης βασίζονται σε τέτοιο βαθμό στη συλλογή δεδομένων και την εκπαίδευση νευρωνικών δικτύων προεικονίζει τον τρόπο με τον οποίο η Τεχνητή Νοημοσύνη θα αυτοματοποιεί ορισμένες εργασίες, συμπεριλαμβανομένης της αστυνομικής επιτήρησης. Η θεωρία των «σπασμένων παραθύρων» εισάχθηκε σε μια περίοδο περικοπών στις αστυνομικές δυνάμεις των ΗΠΑ και θα πρέπει, εν μέρει, να γίνει αντιληπτή ως στρατηγική για την αναδιοργάνωση της αστυνόμευσης όταν οι πόροι δεν θα επιτρέπουν τις πεζές περιπολίες. Οι κάμερες της δικτυωμένης παρακολούθησης επιτρέπουν σε λίγα άτομα να επιβλέπουν πολλές διαφορετικές τοποθεσίες από απόσταση. Αλλά η ανθρώπινη εργασία θα μπορούσε να μειωθεί ακόμα περισσότερο: κάθε εικόνα υπακούει στους περιορισμούς του ζουμ, της εστίασης και του προσανατολισμού μιας συγκεκριμένης κάμερας, και κάθε εικόνα εξαρτάται από έναν χειριστή για να δει τι θα εμφανίζει, ενώ ένα πανκατευθυντικό μικρόφωνο καλύπτει ένα πολύ μεγαλύτερο εύρος, συμπεριλαμβανομένων των χώρων εκτός του πλαισίου της εικόνας. Επιπλέον, είναι περιττό, ακόμα και αδύνατο, για έναν ανθρώπινο χειριστή να ακούει το σύνολο του ήχου, οπότε ο ήχος παρακολουθείται από αλγόριθμο. Όχι μόνο αυτό προάγει τη μετατροπή του σώματος του αστυνομικού σε ηλεκτρονικό σύστημα και κώδικα –με τον ίδιο τρόπο που ο τροχονόμος του 19ου αιώνα έχει εξαφανιστεί μέσα στα συστήματα ελέγχου της κυκλοφορίας του 21ου αιώνα (βλ. Dockray, Rowell, & Whitton 2005)–, αλλά δίνει τη δυνατότητα για μια αστυνόμευση ακριβείας, που αποτρέπει τις ενεργοβόρες ποινικές έρευνες και νομικές διαβουλεύσεις, εφόσον προλαμβάνει τις εγκληματικές πράξεις.

Η τάση για αυτοματοποίηση και έλεγχο δεν περιορίζεται αποκλειστικά στην αστυνόμευση, ακόμα κι αν τα αποτελέσματα είναι πιο συχνά ορατά εκεί. Η «Google» πρόσφατα δοκίμασε ένα νευρωνικό δίκτυο για να προβλέψει πότε θα πεθάνει ένας νοσηλευόμενος ασθενής, υποτίθεται με ποσοστό ακριβείας 95%. Θα ταξινομούν οι εταιρείες δεδομένων τους αρρώστους και ασθενείς πριν τους δει το ιατρικό προσωπικό; Ίσως να μην υπάρχουν πια γιατροί και νοσοκόμες παρά μόνο τεχνικοί. Η «Microsoft» πιστεύει ότι οι ενδείξεις από τις αναζητήσεις στο διαδίκτυο μπορούν να χρησιμοποιηθούν για την πρόβλεψη του καρκίνου, και η «Target» διαπιστώνει εγκυμοσύνες από το ιστορικό αγορών. Είναι άλλο πράγμα να γίνονται τέτοιες προβλέψεις ως εάν ο στόχος τους είναι «να βγουν σωστές», και άλλο να χρησιμοποιούνται για την ανακατασκευή ολόκληρων εξατομικευμένων κόσμων, όπως συμβαίνει με τους αλγόριθμους προτάσεων του YouTube ή τις στοχευμένες διαφημίσεις του Facebook. Οι τοίχοι που στενεύουν γύρω μας δεν έχουν τη μορφή κελιού φυλακής, αλλά διαμορφώνονται στο σχήμα του σώματός μας. Αν μας έχουν κατασχέσει το μέλλον, είναι προς την τροχιά που ούτως ή άλλως θα είχαμε πάρει. Ο Φλούσερ (1999: 48) έγραψε ότι «το ανθρώπινο ον μπορεί να θέλει μόνο ό,τι το ρομπότ μπορεί να κάνει», περιγράφοντας έτσι ένα μέλλον στο οποίο οι μηχανές δεν γίνονται ακριβώς πιο αισθανόμενες, αλλά όπου αυτή η αισθητικότητα (sentience) γίνεται πιο μηχανική. Φυσικά, μια τέτοια διαδικασία δεν ολοκληρώνεται ποτέ. Δεν φτάνει στο πέρας της με την παραγωγή ενός νευρωνικού δικτύου και κάποιων προβλέψεων. Απεναντίας, επιτρέπει την περαιτέρω εξαγωγή δεδομένων και, προφανώς, την περαιτέρω εκπαίδευση της Τεχνητής Νοημοσύνης, προάγοντας περαιτέρω την παλινδρομική κίνηση από μέτα- σε μέτα- σε μέτα-.

Αναφορές

Burroughs, S. William. 1978. «The Limits of Control». Στο *Semiotext(e): Schizo-Culture*, 3:2. 38-43.

Chun, Wendy. 2006. *Control and Freedom: Power and Paranoia in the Age of Fiber Optics*. Μασαχουσέτη: MIT Press.

Deleuze, Gilles. 1995. «Postscript on control societies». Στο *Negotiations, 1972-1990*, 177-182. Νέα Υόρκη: Columbia University Press.

Deleuze, Gilles. 1998. «Having an idea in cinema». Στο *Deleuze & Guattari: New Mappings in Politics, Philosophy, and Culture*, 14-22. Μινεάπολη: University of Minnesota Press.

Dockray, Sean, Steve Rowell, και Fiona Whitton. 2018. «Blocking all lanes». *Cabinet* 17: Laughter. http://cabinetmagazine.org/issues/17/blocking.php, τελευταία πρόσβαση Φεβρουάριος 2024.

Flusser, Vilém. 1999. *The Shape of Things: A Philosophy of Design*. Λονδίνο: Reaktion.

Flusser, Vilém. 2013. *Post-History*. Μινεάπολη: Univocal Publishing.

Gemmeke F. Jort κ.α. 2017. «Audio set: An ontology and human-labeled dataset for audio events» *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 776-780.

Rouvroy, Antoinette, και Thomas Berns. 2013. «Gouvernementalité algorithmique et perspectives d' émancipation: Le disparate comme condition d' individuation par la relation?», *Réseaux* 177: 163-196.