

The Frontiers of Big Data

Kenneth Cukier

Data Editor, The Economist

Co-author “Big Data: A revolution that will transform how we live, work and think”

OECD Global Forum on the Knowledge Economy, Tokyo, October 2, 2014

more...

more is not more...

more is not more...

more is new

more is not more...

more is new

more is better

more is not more...

more is new

more is better

more is different





In apic epistolae sancti iheronimi ad paulinum presbiterum de omnibus diuine historie libris capituli primi.

Hic autem ambrosius tua michi munuscula precor deus sit et suauissimas leas. Non a principio amicitiam habet, plura et iam sed et veteris amicitiae noua precor. Et ita cum illa necessitas est et spiritus glauco copulata. Non non ualens in familiaritate non pueri a canem corpus non libola et palpas adulari sed dei timor et diuinitas temperatam studia conuolant. Legimus in ueteris historis quosdam sustulisse pueros a nouos adhibere pios. Maria et alii. ut eos quos et libris nouauerunt. coramque uidetur. Sicur piagoras manuphicos uero sic plato egypti et archimedeo carum. eandemque oram rrasit. que quosdam magna grecia diebus. laboriosissime peragunt. et ut qui aduerso mte erant. et potius amplexus doctrinas ad aduic gignantia psonabatur. hanc pagnus atque disciplus. malis aliena ueritate discere. Non sua spudat in gerat. Denique cum leas quasi roto orbe fugiat. roto psequitur. capere a piratio et uenidano. oratio curdi. lino paruit. dudum capinus uindum et seruis. Tame quia plus maior emere se hinc. ad orum lunt. ludo eloquentie fonte manant. et uisum hispanie galliarumque hinc. quosdam uenisse uobis legimus. et quos ad complacitum sui roma non recare. um? hois fama pduerit. Et habuit illa eras in audiam omnes seculis. et eandemque miram. ut uerbi rana

ingredere. aliud quia uerben queam. Hippoloni? hinc ille magis ut uulgus loquitur. hinc plus ut piagora eradunt. iuraue plas. pstant caucali. albanos. scythas. mallagras. opulenti summa indie regna pertrauit. et ad grecorum lanilino phison ampne admissio puenit ad braguanas. ut hircam in throno sedent auro et de caucali fonte porantou. inter pauco discipulos. de natura. et moab. ac de anlu diei et sibi audire docerant. Hinc y elamitas. babilonios. thaldis. os. uidos. althios. parthos. hircos. pharicos. arabos. palestinos. caisus ad allea uerba. paxerit ad ethiopi. ut gignosophistas et famosissimam solis mensam uideret in sabulo. Inuenit ille uir ubique et discere. et semp proficere. semp se melior fieri. Scripsit super hoc plenissime octo uoluminibus. phisosteanis.

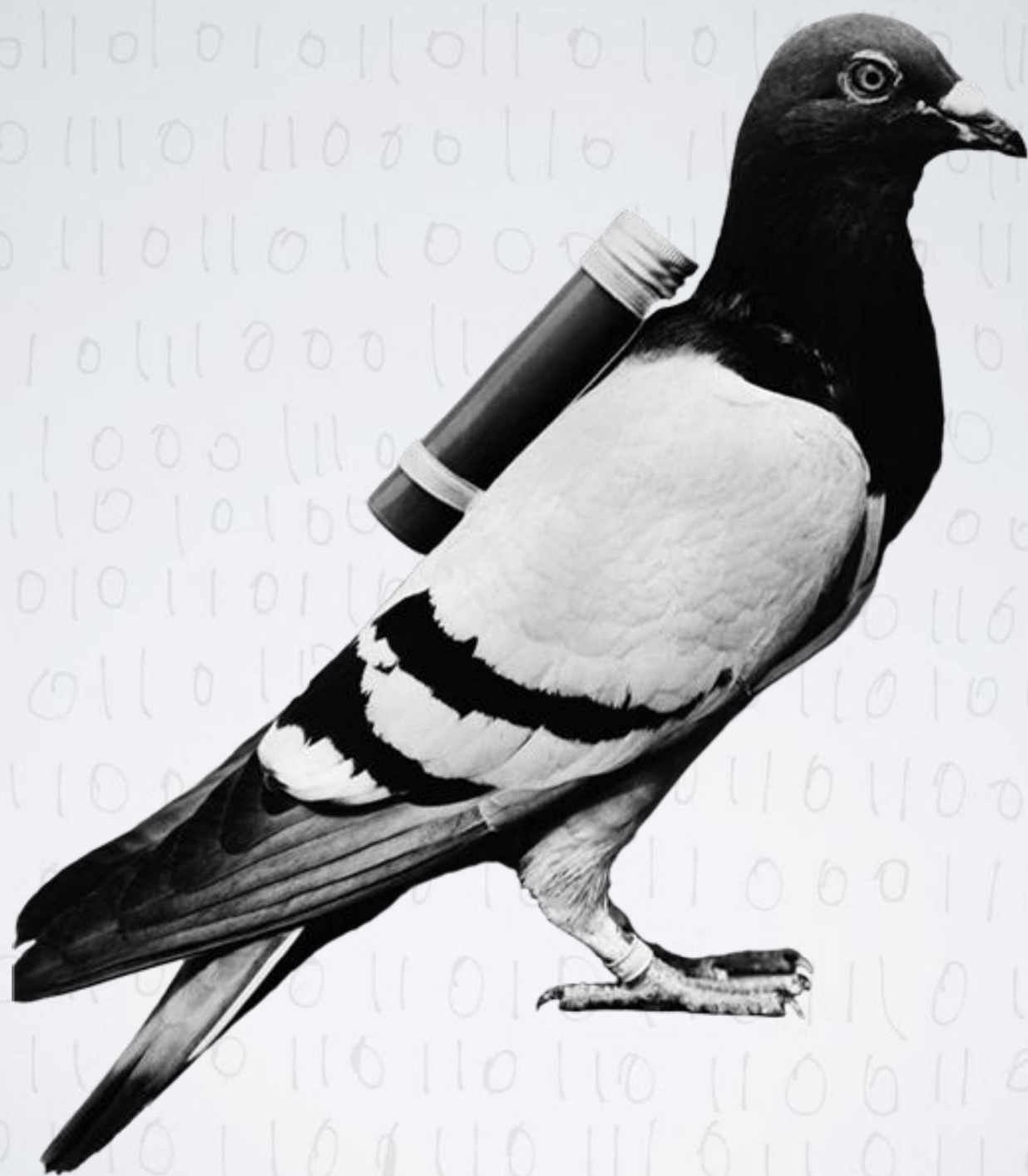
Quid loquar de secti hominibus. cum aphro paulus. uas electum. et magister gentium. qui de consuetudine tam in se hospitis loquebat. dicit. In repetentiis quibus eius qui in me loquitur spiritus. Post demascei arabias sustulit. alcedre iherosolimam ut uideret petru et maste apud cum diebus quindam. Hinc cum iustis ebdomadis et o gdo. adio. hinc? gentium predicato. instructus erant. Rursumque post annos quatuordecim assumpto barnaba et oro. expositum cum aphro ewagelium. ne forte in uacuum curaret aut curaret. Habet ueritas quod lacrimis emergit. uir uocatus adus. et in auro discipuli et auditoris ore mansula. foras sonat. Et hinc et ceteris. cum rodi gularat. et legatur

more is not more...

more is new

more is better

more is different





more is not more...

more is new

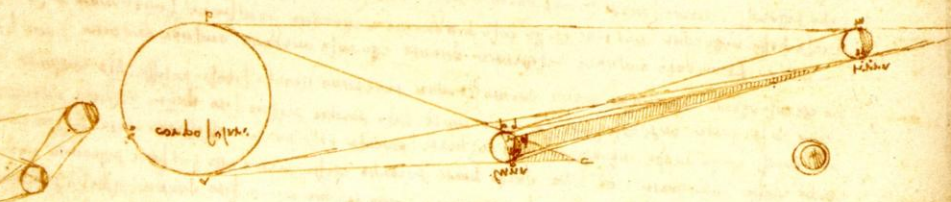
more is better

more is different

Handwritten text in Italian, likely a preface or introduction to a treatise on optics or geometry. The text is dense and written in a cursive script.

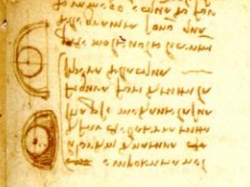


Handwritten text in Italian, continuing the treatise. It includes some mathematical or optical principles.



Handwritten text in Italian, discussing optical phenomena. The text is written in a cursive script.

Handwritten text in Italian, continuing the treatise. It includes some mathematical or optical principles.



Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.



Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.



Handwritten text in Italian, likely a caption or description for the diagram above.



Handwritten text in Italian, likely a caption or description for the diagram above.

Handwritten text in Italian, likely a preface or introduction to a treatise on optics or geometry. The text is dense and written in a cursive script.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

Handwritten text in Italian, likely a caption or description for the diagrams above.

口02千く
又5中ケケ9カ5カ5
甘ケ千号カ4口02千く

又中9廿井8廿2
千71080号カ477
千710

66の中05千く井ア
ケカ05ケのケ3廿7口
ム又廿ケ4のケカ05ケのケ
4井7

の5く井7口9106
井0ノ千1廿ゾのび8047の
井0ノ千1廿ゾ

ム千井中3カ61び中
24ア口の47廿
81井千口5896
81井千口5896

7ケ又ノムケ
2ノ1の067ケ又ノムケ
2廿えノ6

8アケムカの60ノ口1
ムア591ア8アケムカ
の60ノ口

び2えゾ3又又び口び
ノ0の6号2
えム2179ノムム又

カ6ノ2ム8千く
千358ムム口789又8
号1号ム千千358ムム口7

48千8中ケ千0び井
7ノび中口48千8中ケ千0
び

廿54廿え中1
6ノ0え4口廿54廿え中1
の2廿井又号9千0え

9千0え
くノの5アゾ井ゾの110
廿ゾ60号の千ム3ノム

60号の千ム3ノム
ケ又5中カ22廿び
4く3の19え53

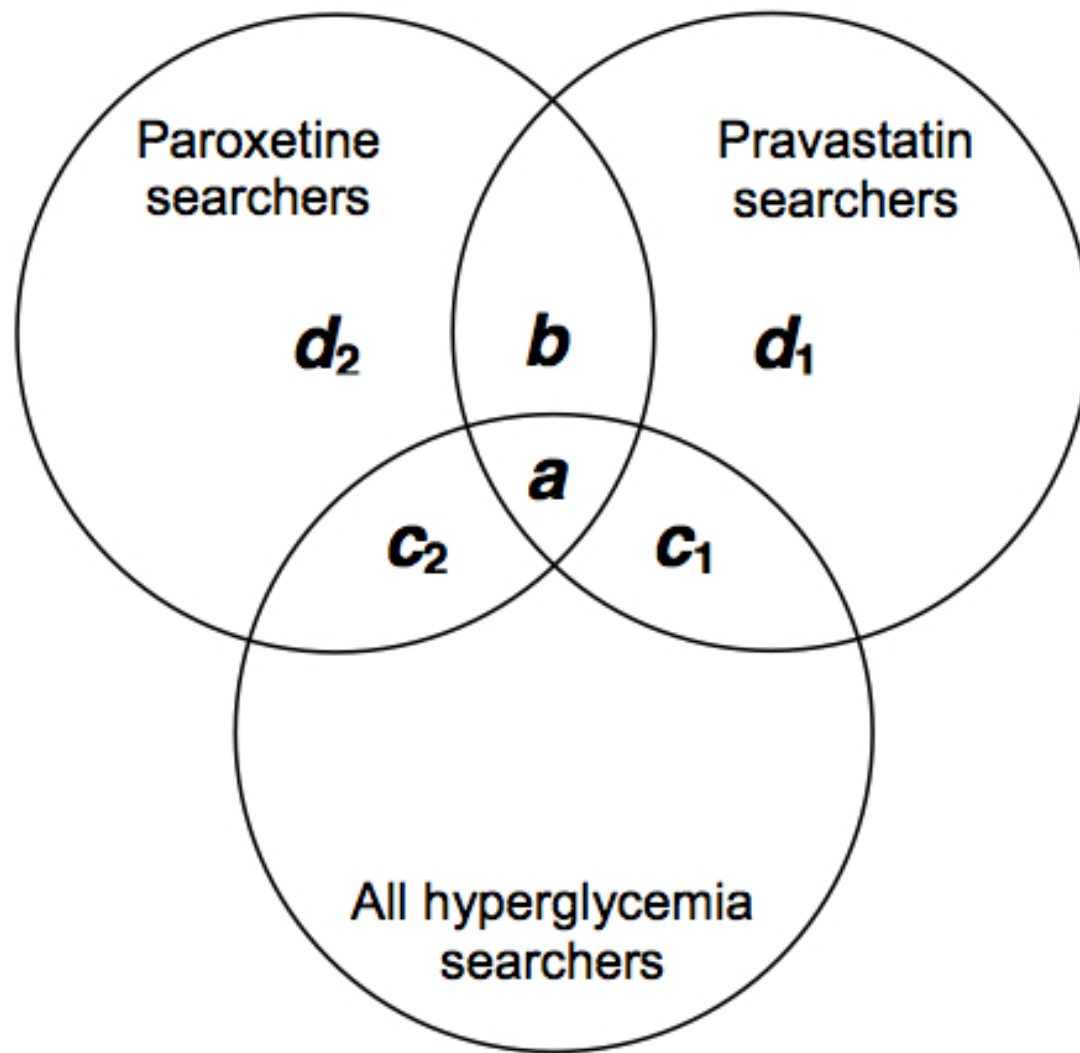
の4び口のえ号井く
ムアア千ム又又の4び
口のえ号井く

new, better, different

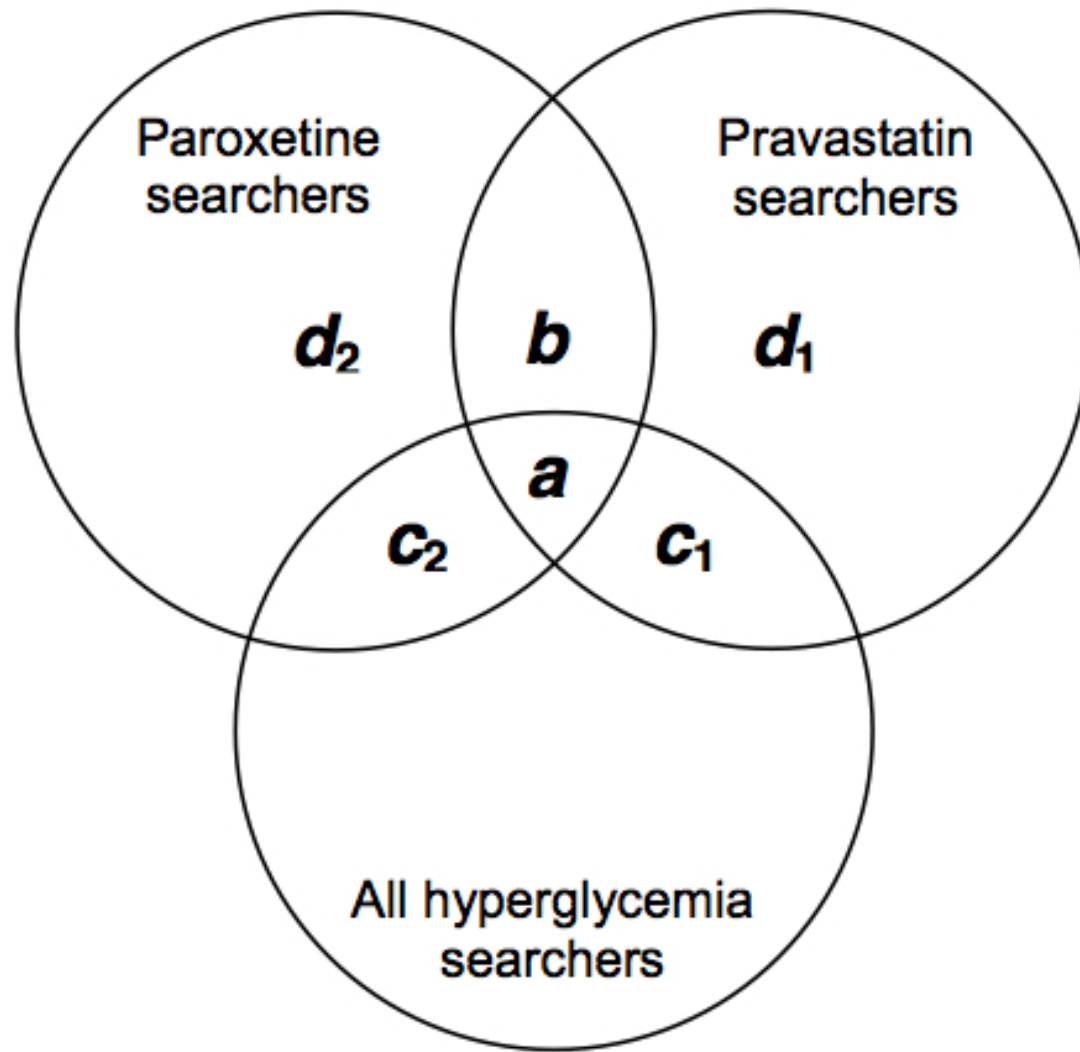
**new, better, different
example**

new, better, different
example
drug side-effects

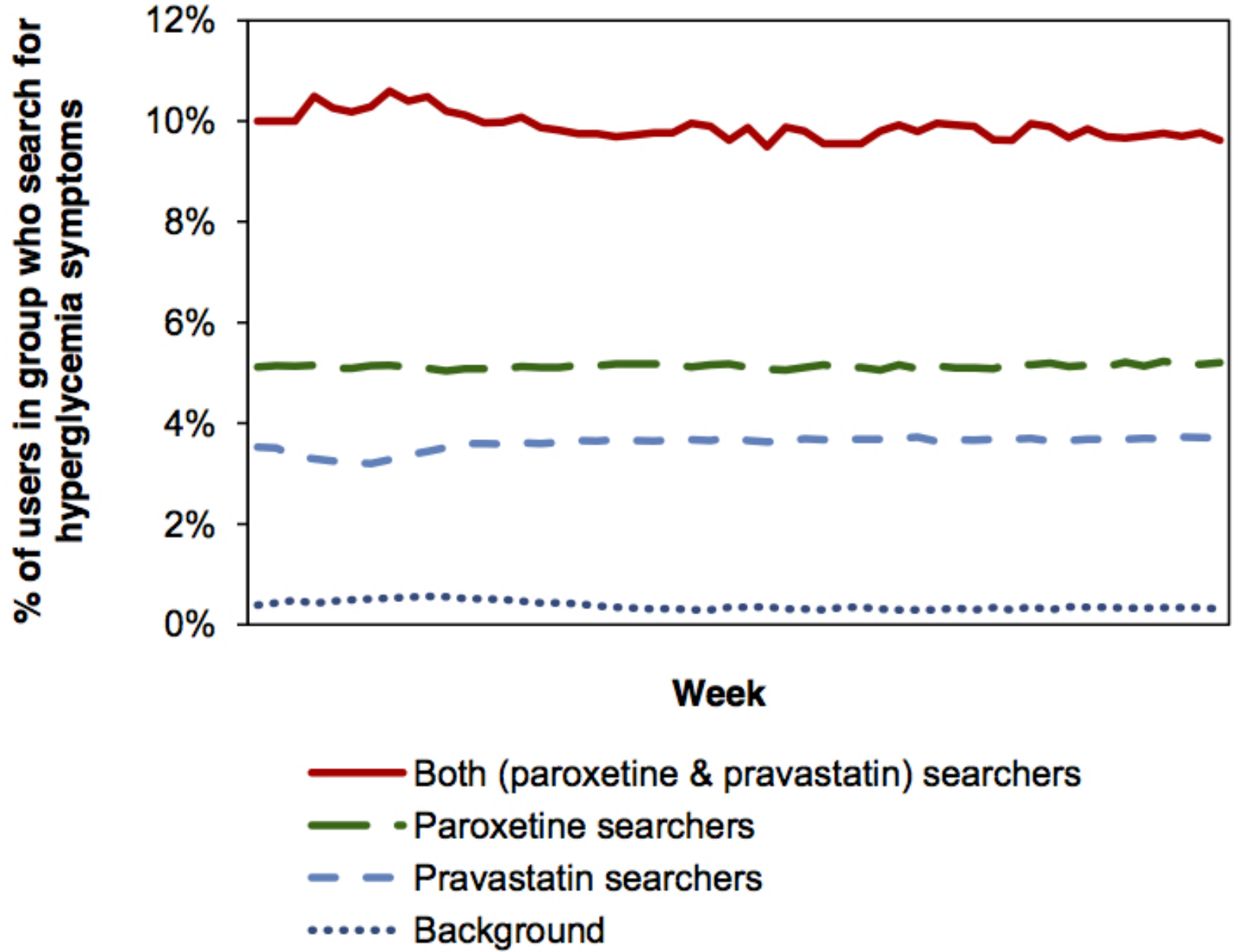
**new, better, different
example
drug side-effects
without a medical study**



search queries



82 million search queries!



Source: "Web-scale pharmacovigilance: listening to signals from the crowd." White RW, Tatonetti NP, Shah NH, Altman RB, Horvitz E. J Am Med Inform Assoc. 2013 May 1;20(3):404-8



**how
technology
data
mindset**

lesson

data = resource

new raw material

economic & social input

1987
analog

2.6 billion

2000

analog storage

digital

0.02 billion

digital

2007

analog

19 billion GB

digital

276 billion GB

Source: "The World's Technological Capacity to Store, Communicate, and Compute Information," Martin Hilbert, Priscila López, Science, 1 April 2011: Vol. 332 no. 6025 pp. 60-65

datafication

datafication

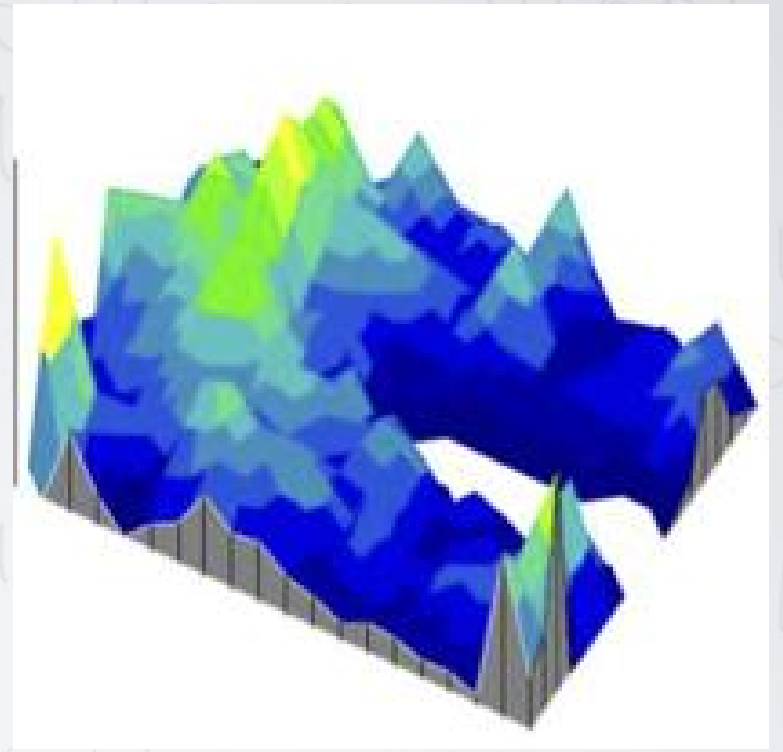
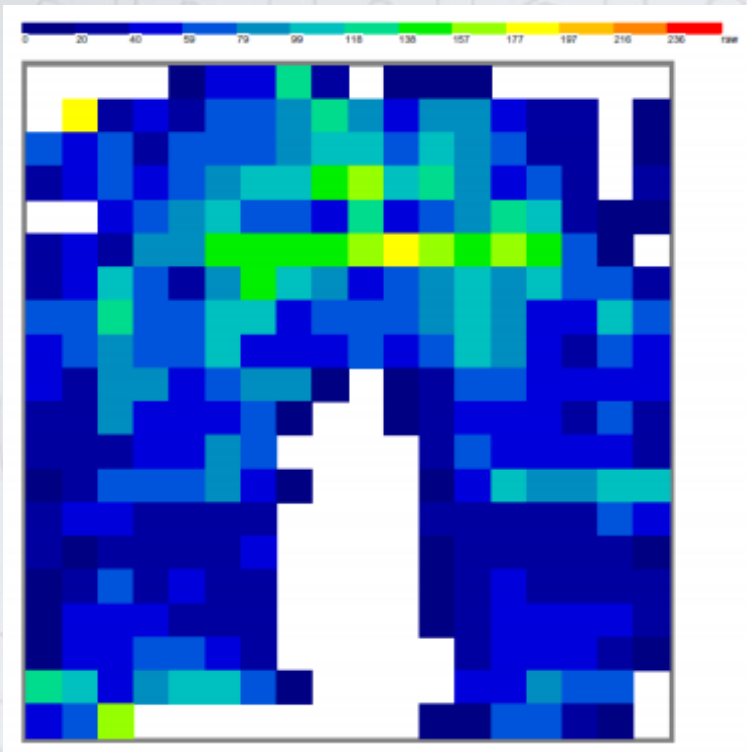
location




W. JACKSON

datafication

posture



datafication



**the world
platform for the
collection and analysis
of data**

machine learning



machine learning

computer translation

speech recognition

self-driving car

etc . . .

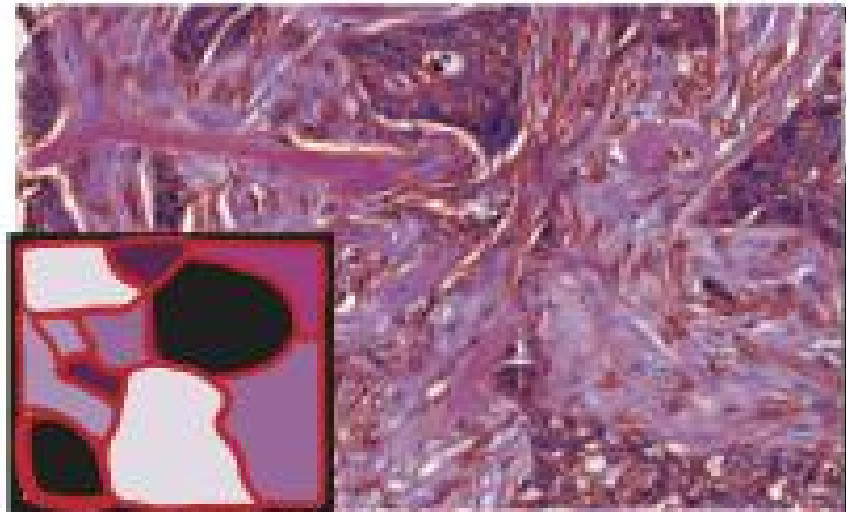


A

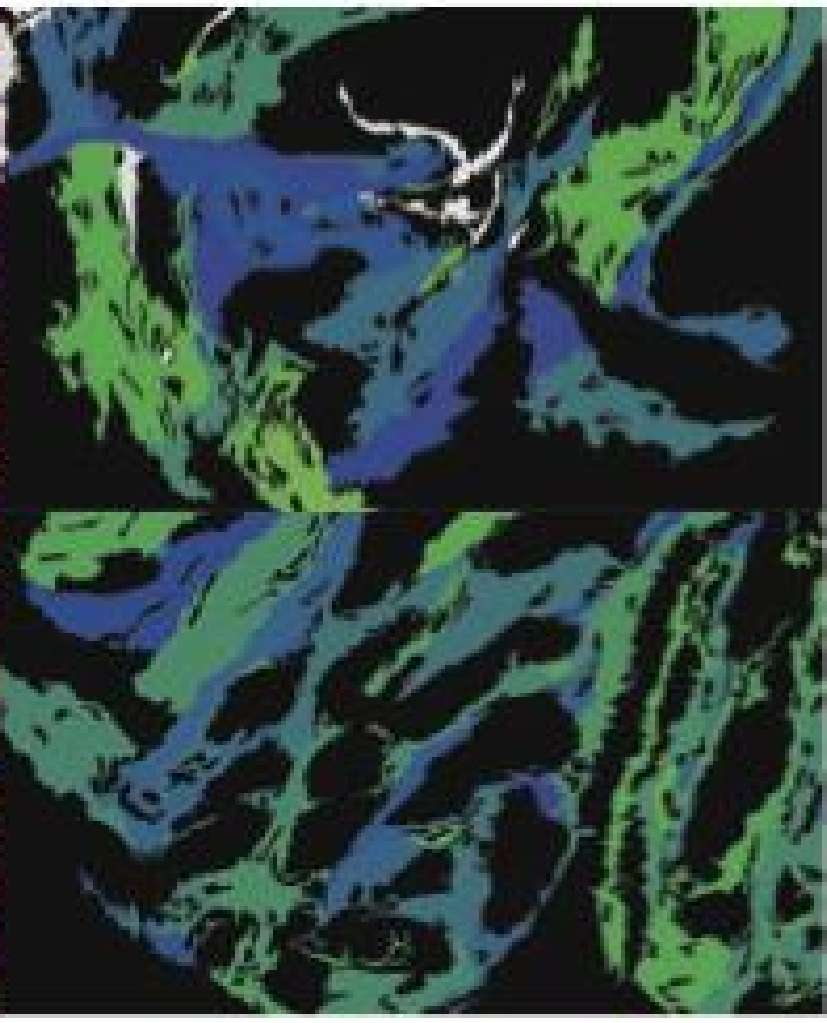
H&E image separated into epithelial and stromal objects

Heat map of stromal matrix objects mean abs. diff to neighbors

Improved prognosis



Worse prognosis



more is not more...

more is new

more is better

more is different

regulation

regulation
privacy

George Orwell

nineteen

eighty-four

a novel



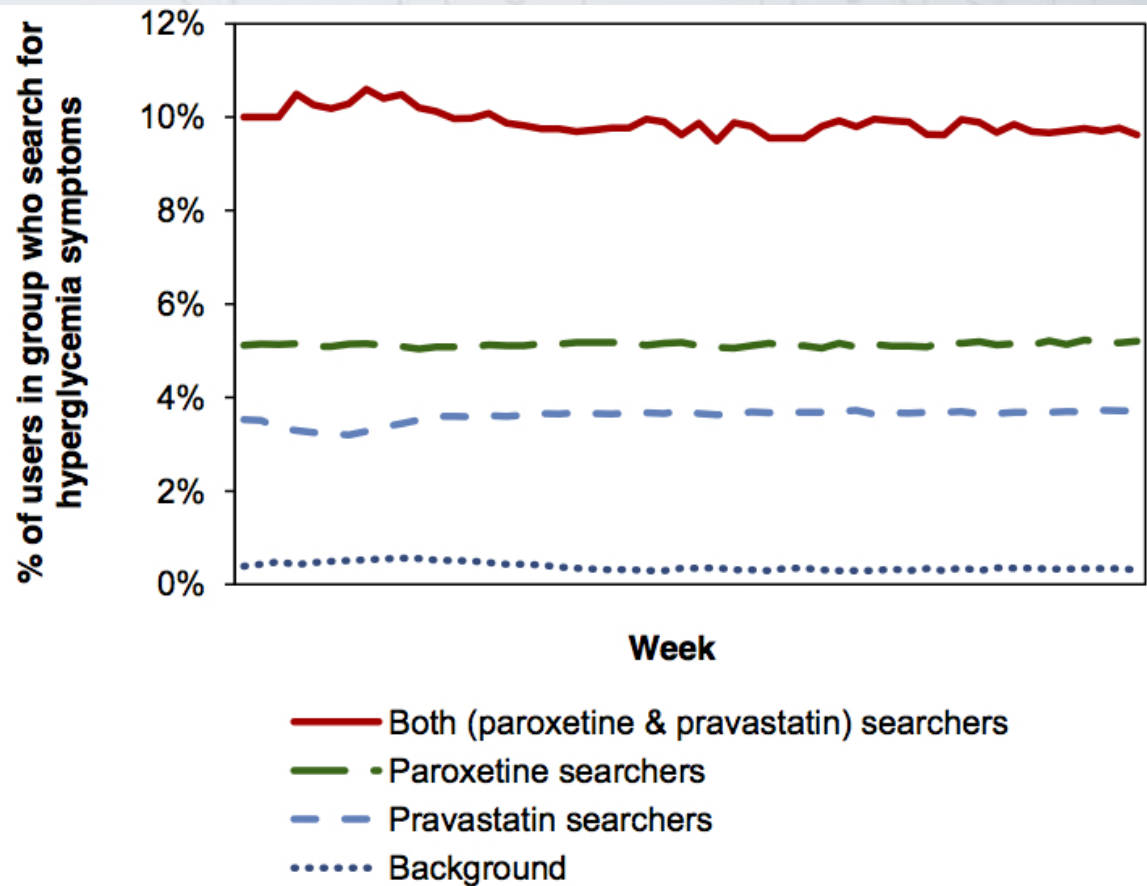
regulation
privacy
ownership

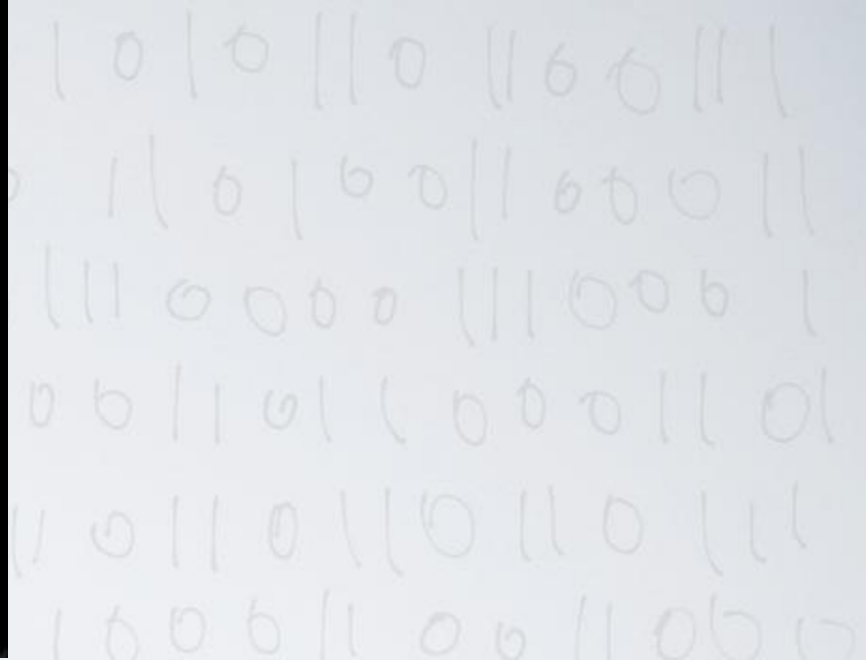


regulation

**regulation
more is not more**

George Orwell
nineteen
eighty-four
a novel



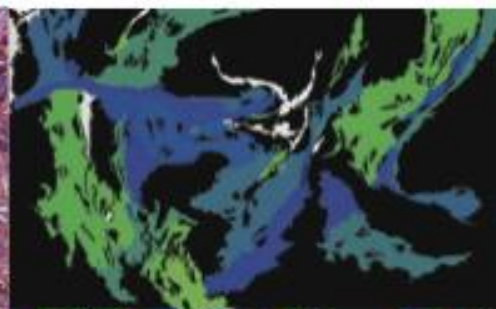
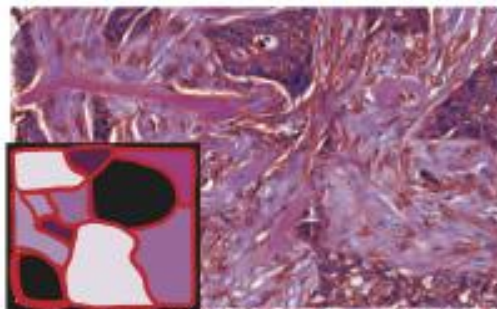


A

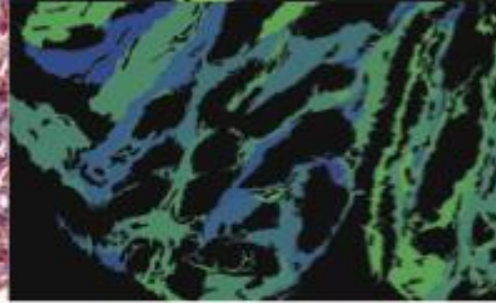
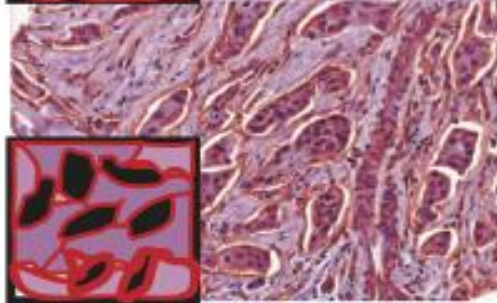
H&E image separated into epithelial and stromal objects

Heat map of stromal matrix objects mean abs. diff to neighbors

Improved prognosis



Worse prognosis



regulation

not more

new

better

different

NOVA TOTIVS TERRARVM ORBIS GEOGRAPHICA AC HYDROGRAPHICA TABVLA. Auct: Henr. Hondius



AMERICA
Sive Terra Nova et Christiana
Columbo natus Regis Castellae
primum detecta et ab Americo
Vesputio nomen sortita

TERRA AUSTRALIS INCOGNITA
Sive Terra Australis
Inconspicua
auctoritate
nonnullorum
Astronomorum
et Geographorum
supposita

TERRA AUSTRALIS INCOGNITA
Sive Terra Australis
Inconspicua
auctoritate
nonnullorum
Astronomorum
et Geographorum
supposita

more is not more...

more is new

more is better

more is different

thank you

KennethCukier@Economist.com

@kncukier