

MANAGEMENT REVIEW

Menaxhim pa teatër.

STRATEGJIA

PILOTET DHE EKSPERIMENTET PARA NDRYSHIMIT TË MADH

Një ide që priti gjashtë muaj, sa pak ide të provuara përmirësojnë diçka, Modeli i Përmirësimit, pse efektet tkurren në shkallë të gjerë, kur t'i besosh një shifre, dhe një kartë për planin e një piloti.



12%

më shumë të ardhura te Bing nga një ndryshim i vogël te titujt e reklamave, një ide që kishte pritur më shumë se gjashtë muaj që dikush ta provonte.

KOHAVI & THOMKE, 2017

BRENDA

TEMA KRYESORE

Gjashtë muaj në sirtar

FAQE 03

ÇFARË THOTË KËRKIMI

Pse rezultatet tkurren në shkallë të gjerë

FAQE 06

MJETI I NUMRIT

Karta e planit të pilotit

FAQE 08

NR. 69 · STRATEGJIA

Në këtë numër

Ndryshimet e mëdha zakonisht vendosen me bindje: një plan, një buxhet, një datë nisjeje. Ky numër flet për ta provuar ndryshimin fillimisht në të vogël, përballë një grupi krahasimi dhe me një rregull vendimi të rënë dakord para nisjes, që të vendosë rezultati, jo mendimi më i zëshëm.

Te Bing, një ide e vogël priti më shumë se gjashtë muaj dhe pastaj i rriti të ardhurat me 12%. Te Microsoft, vetëm një nga tri ide të provuara përmirësoi atë që synonte. Modeli i Përmirësimit e provon një ndryshim në të vogël: një ditë, një njësi. Në 126 prova qeveritare, efektet ishin shumë më të vogla se në studimet e botuara. Katër kontrolle tregojnë kur t'i besosh një rezultati, dhe një kartë ndihmon për planin e pilotit.

Stiven Janaqi REDAKTOR · STIVEN CATALYST

03 TEMA KRYESORE
Gjashtë muaj në sirtar

04 SHIFRAT
Shumica e ideve nuk funksionojnë

05 MODELI
Tri pyetje, cikle të vogla

06 ÇFARË THOTË KËRKIMI
Pse rezultatet tkurren në shkallë të gjerë

07 SI MATET
Shifra që t'u besosh

08 MJETI I NUMRIT
Karta e planit të pilotit

09 BURIMET
Burimet dhe metoda

SI LEXOHEM KY NUMËR

Shifra

Çdo shifër ka burimin dhe vitin poshtë faqes së vet.

Leximi ynë

Ku flet redaksia dhe jo kërkimi, shkruhet qartë.

Praktika

Hapat dhe karta janë propozime për t'u provuar, jo rezultate kërkimi.

TEMA KRYESORE

Gjashtë muaj në sirtar

Në 2012, një punonjës i Microsoft-it që punonte për motorin e kërkimit Bing propozoi të ndryshonte mënyrën si shfaqeshin titujt e reklamave. Një inxhinieri do t'i mernte disa ditë, por ishte një nga qindra ide, dhe menaxherët e programit i dhanë përparësi të ulët.

Ideja priti më shumë se gjashtë muaj. Pastaj një inxhinier, duke parë sa pak kushtonte kodi, nisi një test të thjeshtë A/B: përdoruesit u ndanë rastësisht mes titujve të vjetër dhe të rinjve. Brenda pak orësh të ardhurat u rritën aq shpejt sa ndezën një alarm “shumë mirë për të qenë e vërtetë”. Zakonisht alarme të tilla tregojnë një gabim në kod. Këtë herë jo.

6+

muaj priti ideja, me përparësi të ulët

12%

më shumë të ardhura, pa dëmtuar treguesit kryesorë të përdoruesve

100+ mln \$

në vit vetëm në SHBA, sipas vlerësimit të autorëve

LEXIMI YNË

Asnjë mbledhje nuk do ta kishte renditur saktë këtë ide. Një test i lirë dha brenda pak orësh përgjigjen që muajt e renditjes së përparësive nuk e kishin dhënë.

Burimi: Ron Kohavi & Stefan Thomke, Harvard Business Review 95(5), 2017 (përmes HBR reprint R1705E)

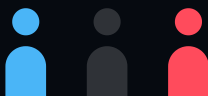
Kohavi dhe Thomke e quajnë idenë që solli më shumë të ardhura në historinë e Bing-ut. 12% u mat në test; 100 milionë dollarët janë shifra e tyre vjetore për SHBA-në.

SHIFRAT

Shumica e ideve nuk funksionojnë

Kompanitë që provojnë pothuajse çdo ndryshim mësojnë të njëjtin mësim, shkruajnë Ron Kohavi dhe Stefan Thomke: shumica e ideve të reja dështojnë në eksperimente, dhe edhe ekspertët shpesh gabojnë se cilat do të shpërblejnë. Te Microsoft në tërësi:

IDETË E PROVUARA ME EKSPERIMENTE TE MICROSOFT, SIPAS REZULTATIT



1 nga 3

dalin efektive

1 nga 3

nuk ndryshojnë asgjë

1 nga 3

kanë rezultat negativ

Në produkte të optimizuara mirë si Google dhe Bing, vetëm rreth 10–20% e eksperimenteve japin rezultat pozitiv; te Slack, rreth 30% e eksperimenteve për të ardhurat (2019). Sipas Thomke-s, mesatarja është rreth 10%.

LEXIMI YNË

Nëse dy nga tri ide të besueshme nuk ndryshojnë asgjë ose i përkeqësojnë gjërat, të zbatosh një ndryshim të madh pa e provuar është bast, jo vendim.

Burimet: Ron Kohavi & Stefan Thomke, Harvard Business Review 95(5), 2017 (përmes HBR reprint R1705E); Ron Kohavi, Diane Tang & Ya Xu, Cambridge University Press, 2020 (përmes Chapter 1, published by the authors); Stefan H. Thomke, Harvard Business Review Press, 2020 (përmes Michael Blanding, "Creating the Experimentation Organization", HBS Working Knowledge, 2019)

Produkte online, ku testet janë të lira; për repartet e punës nuk gjetëm pjesë të krahasueshme. Slack: një postim që citojnë Kohavi, Tang & Xu; 10% e Thomke-s: një intervistë (2019).

MODELI

Tri pyetje, cikle të vogla

Modeli i Përmirësimit (Model for Improvement), nga Associates in Process Improvement dhe i shtjelluar te The Improvement Guide (Langley dhe kolegët, 2009), bën tri pyetje para se të ndryshohet diçka përgjithmonë.

01**Qëllimi**

Çfarë po përpiqemi të arrijmë?

02**Matja**

Si do ta dimë që një ndryshim është përmirësim?

03**Ndryshimi**

Çfarë ndryshimi mund të bëjmë që të sjellë përmirësim?

PROVIMI I NDRYSHIMIT: CIKLI PLAN-DO-STUDY-ACT

01**Planifiko****02****Bëj****03****Studio****04****Vepro**

Planifiko ndryshimin, provoje, studio rezultatet, vepro sipas asaj që mësove. Prova është e vogël me qëllim: shembulli i Institute for Healthcare Improvement është një pacient, një ditë. Pas disa ciklesh në kushte të ndryshme, si dita e nata, ndryshimi bëhet i përhershëm dhe shtrihet.

LEXIMI YNË

Piloti nuk është një nisje e butë. Është një pyetje me afat: a e bën ky ndryshim, këtu, më mirë se ajo që bëjmë sot?

Burimi: Langley, Moen, Nolan, Nolan, Norman & Provost, Jossey-Bass, 2009 (përmes Institute for Healthcare Improvement, "Model for Improvement")

Pyetjet dhe hapat ndjekin faqen e Institute for Healthcare Improvement për modelin; vetë libri nuk u pa. Leximi është i redaksisë.

Më shumë te esaja

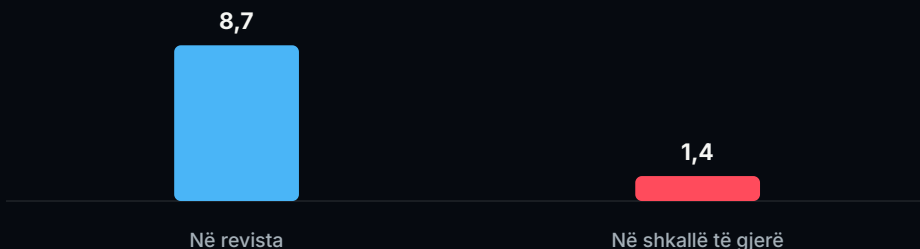
DMAIC pa black belt

ÇFARË THOTË KËRKIMI

Pse rezultatet tkurren në shkallë të gjerë

Stefano DellaVigna dhe Elizabeth Linos mbledhën 126 prova të rastësishme me 23 milionë njerëz, të gjitha të kryera nga dy “njësi nudge” të mëdha të qeverisë amerikane, dhe i krahasuan me provat nudge të botuara në revista akademike.

EFEKTI MESATAR I NJË NUDGE-I TE PJESËMARRJA, PIKË PËRQINDJEJE



Provat e botuara ishin të vogla: mediana e një grupi trajtimi ishte 484 veta, kundrejt 10.006 te njësitë nudge. Provat e vogla dallojnë vetëm efekte të mëdha, dhe revistat botojnë kryesisht atë që dallohet. Botimi selektiv, i rënduar nga fuqia e ulët statistikore, shpjegon rreth 70% të diferencës; dallimet mes vetë nudge-ve shpjegojnë pjesën më të madhe të mbetjes.

LEXIMI YNË

Historia e suksesit që lexon është ajo që u botua. Planifiko sikur shtrirja e ndryshimit të japë më pak se piloti që të bindi.

Burimi: Stefano DellaVigna & Elizabeth Linos, *Econometrica* 90(1), 2022

Pjesëmarrja: pjesa e atyre që bëjnë atë që u kërkohet, si regjistrimi ose vaksinimi. Efektet dhe 70% nga *Econometrica* (2022); madhësitë e grupeve nga versioni i punës i 2020.

SI MATET

Shifra që t'u besosh

“Të marrësh shifra është e lehtë; të marrësh shifra që mund t'u besosh është e vështirë”, shkruajnë Kohavi dhe Thomke. Katër kontrolle para se të vendosësh rezultati i pilotit:

1

Krahaso me një grup kontrolli

Te Yahoo, një studim vëzhgues e vlerësoi efektin e reklamave të kërkimit për markën 871–1.198%; një eksperiment i kontrolluar, 5,4%.

2

Cakto një masë kryesore dhe kufijtë mbrojtës

Para nisjes, jo pas; kufijtë mbrojtës janë ato që s'duhet të përkeqësohen.

3

Vazhdo të paktën një javë të plotë

LinkedIn mat të paktën një javë: një ditë e vetme i mbipërfaqëson përdoruesit më aktivë.

4

Kontrollo ndarjen, dysho për befasinë

Grupet me madhësi të paplanifikuar shpesh e zhvlerësojnë rezultatin. Ligji i Twyman-it: çdo shifër që duket interesante ose ndryshe zakonisht është e gabuar.

SHEMBULL HIPOTETIK, NJË RRUGË E RE MBLEDHJEJE NË NJË MAGAZINË

Krahasimi zona A me rrugën e re, zona B si më parë, po ato dy javë

Masa kryesore rreshta të mbledhur në orë: A +6%, B +4%

Kufiri mbrojtës gabimet e mbledhjes: të pandryshuara në të dyja

Përballë zonës B, rruga vlen rreth 2 pikë, jo 6. Numrat janë të shpikur.

Burimet: Ron Kohavi & Stefan Thomke, Harvard Business Review 95(5), 2017 (përmes HBR reprint R1705E); Ron Kohavi, Diane Tang & Ya Xu, Cambridge University Press, 2020 (përmes Chapter 1, published by the authors); Ya Xu, Weitao Duan & Shaochen Huang, KDD '18 (ACM), 2018

Kontrollet sipas Kohavi & Thomke (2017), Kohavi, Tang & Xu (2020) dhe, për javën, Xu et al. (2018). Shembulli është i redaksisë.

MJETI I NUMRIT

Karta e planit të pilotit

Plotësoje para nisjes, me ata që do ta drejtojnë pilotin. Pastaj shtrihe me hapa: fillimisht një grup i vogël, një fazë matjeje, pastaj të gjithë.

01 Hipoteza nëse ndryshojmë X, Y përmirësohet rreth Z, sepse ...

02 Çfarë ndryshon një ndryshim, aq i qartë sa ta përsëriste kushdo

03 Ku dhe sa gjatë një zonë, turn ose ekip; të paktën një javë e plotë normale

04 Çfarë matet një masë kryesore, si numërohet, dhe kufijtë mbrojtës

05 Grupi i krahasimit ku vazhdon mënyra e sotme, në të njëjtat ditë

06 Rregulli i vendimit i caktuar para nisjes: shtrihe, ndryshoje ose ndale, dhe kush vendos

Burimet: Langley, Moen, Nolan, Nolan, Norman & Provost, Jossey-Bass, 2009 (përmes Institute for Healthcare Improvement, "Model for Improvement"); Ya Xu, Weitao Duan & Shaochen Huang, KDD '18 (ACM), 2018; Stefan H. Thomke, Harvard Business Review Press, 2020 (përmes Michael Blanding, "Creating the Experimentation Organization", HBS Working Knowledge, 2019)

Praktikë e propozuar nga redaksia, sipas Modelit të Përmirësimit, Xu et al. (2018) dhe rregullit të Thomke-s: një hipotezë që provohet, një grup kontrolli, dhe t'u përmbahesh rezultateve.

Hape mjetin: Sigma dhe karta e kontrollit stivencatalyst.com/sq/tools/sigma-control-chart/

BURIMET DHE METODA

Çdo shifër ka një burim.

Shifrat e këtij numri vijnë nga burimet më poshtë. Viti tregon sa e re është secila.

The Surprising Power of Online Experiments

Ron Kohavi & Stefan Thomke, Harvard Business Review 95(5), 2017 · përmes HBR reprint R1705E

<https://hbr.org/2017/09/the-surprising-power-of-online-experiments>

Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing

Ron Kohavi, Diane Tang & Ya Xu, Cambridge University Press, 2020 · përmes Chapter 1, published by the authors

<https://experimentguide.com/>

Experimentation Works: The Surprising Power of Business Experiments

Stefan H. Thomke, Harvard Business Review Press, 2020 · përmes Michael Blanding, "Creating the Experimentation Organization", HBS Working Knowledge, 2019

<https://www.library.hbs.edu/working-knowledge/creating-the-experimentation-organization>

The Improvement Guide: A Practical Approach to Enhancing Organizational Performance (2nd ed.)

Langley, Moen, Nolan, Nolan, Norman & Provost, Jossey-Bass, 2009 · përmes Institute for Healthcare Improvement, "Model for Improvement"

<https://www.ihl.org/resources/how-improve-model-improvement>

RCTs to Scale: Comprehensive Evidence from Two Nudge Units

Stefano DellaVigna & Elizabeth Linos, Econometrica 90(1), 2022

<https://doi.org/10.3982/ECTA18709>

SQR: Balancing Speed, Quality and Risk in Online Experiments

Ya Xu, Weitao Duan & Shaochen Huang, KDD '18 (ACM), 2018

<https://doi.org/10.1145/3219819.3219875>

METODA EDITORIALE

Çdo shifër u kontrollua për vitin, botuesin dhe çfarë mat saktësisht. Kur faqja e botuesit nuk hapej, shifra u kontrollua te përmbledhje të pavarura dhe shënohet me “përmes”. Interpretimi i redaksisë shënohet “Leximi ynë”. Shifrat që nuk u konfirmuan nuk janë në numër.



MANAGEMENT REVIEW

Menaxhim pa teatër.

Çdo numër, një pyetje e menaxhimit, e kontrolluar me kërkimet më të mira.

TË GJITHË NUMRAT

stivencatalyst.com/sq/magazine/management-review.html