



BIG Data = BIG Deal

תופעת הביג דאטה, ההזדמנויות העסקיות, האתגרים ואסטרטגיית המימוש • מצורף בונוס, מבוסס בשורה עתיקה: "לחשוב בגדול, להתחיל בקטן" • מידע גדול זה עסק גדול • מאמר אחד משניים

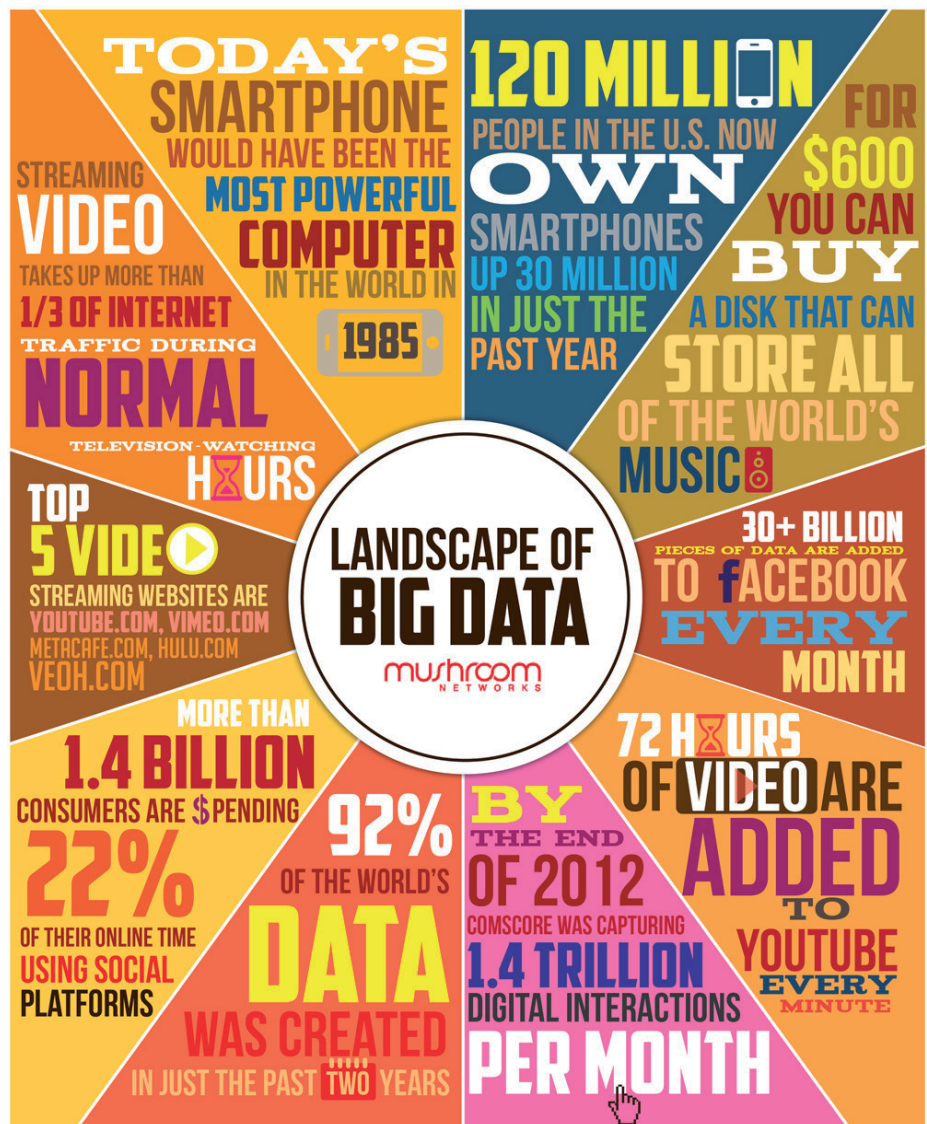
בבתי חולים, מידע על התחבורה בכבישים, מידע מחיישנים רפואיים ועוד. בנוסף למידע הרב שנוצר על ידי אנשים, יותר ממחצית המידע הזמין כביג דאטה הוא מידע אוטומטי שמתעד שימוש שלנו באינטרנט או בחיישנים שונים, אשר מכונה Dark Data ו־Shadow Data.

לפי גרטנר, תוכן חברתי שנוצר על ידי גולשים ברשת הוא התורם העיקרי למידע חדש בארגונים ויגיע לסך 20% חדירה בארגונים. בעוד שכיום, עדיין רוב סוגי המידע הם טקסטואליים, מתחזקת המגמה של ייצור קבצי וידאו לתיעוד תכנים חברתיים. בנוסף: מתרחב מנעד המכשירים שמייצרים תכנים כלשהם (טלוויזיות, צמידים רפואיים, חיישני תנועה ועוד). למעשה, לרוב פריטי המידע המיוצרים כיום מצטרפות תכונות גיאוגרפיות (סמיקום) שמעצמות את ההזדמנויות הגלומות בנייתן. לפי חטיבת הסטטיסטיקה של האו"ם, בשנת 2020 יהיו בעולם 50 מיליארד "מכשירים" (Internet of Things) שמחוברים ומייצרים אותות מידע תדיר. זרם בלתי פוסק של מידע.

אז מה?

לפי אותו דו"ח של האו"ם, כרגע השימוש בביג דאטה מתרכז בשני תחומים עיקריים: בתעשיות ביטחוניות (למשל כిפת ברזל) ושימושים מסחריים בארגונים. למרות ש-70% מהמידע נוצר על ידי אנשים פרטיים, 85% ממנו כבר מטופל או נמצא באחריות של ארגונים. הארגונים מתחילים להבין שהיצע המידע הזמין בעולם מסביבנו הוא אינסופי בכמות ובמגוון, ומשכר בפוטנציאל שלו. קיימת תחושה של רחש אינסופי, שאנחנו - כנראה - חייבים להקשיב לו. למה? כי זה משפר את יכולת החיזוי, לצורך קבלת החלטות. קחו לדוגמה שימושי חיזוי הונאה או קביעת סיכוני אשראי. המודלים הקלאסיים משתמשים ב-10-15 מדדים לחיזוי, אך ברור, שכלל שנכניס באופן מושכל למשוואה מדדים רבים יותר, כך כנראה נוכל לדייק יותר בחיזוי. חברת ZestFinance, חברה טכנולוגית בתחום חישוב סיכוני האשראי, שנבחרה לאחת החברות החדשניות בתחום, פיתחה מודל סיכון שמשלב ניתוח מידע מבוסס ביג דאטה עם Machine Learning. החברה מספקת

מה גדול הביג דאטה? מאמרים בנושאי ביג דאטה מתחילים בדרך כלל בסטטיסטיקות מסחררות על כמות המידע הנוצר בעולם ועל כמה מהר אנחנו מכפילים את הכמויות. להלן כמה מספרים:



מקור: mushroom networks

וזה עוד לא הכל. ביג דאטה אינו רק כמות (Volume). הוא מקיף סוגים רבים של מידע (Variety): תכנים חברתיים, דפוסי שימוש בדפי אינטרנט, תוצאות חיפוש במנועי חיפוש, אותות שימוש בטלפון סלולרי, תנועות שימוש בכרטיסי אשראי, טרנזקציות ביטוח, סטטיסטיקות ביקורים