Agenda

What is big data?

Why would an organization care?

How to compete based
on big data?

Big data definition

"One man's noise is another man's data"
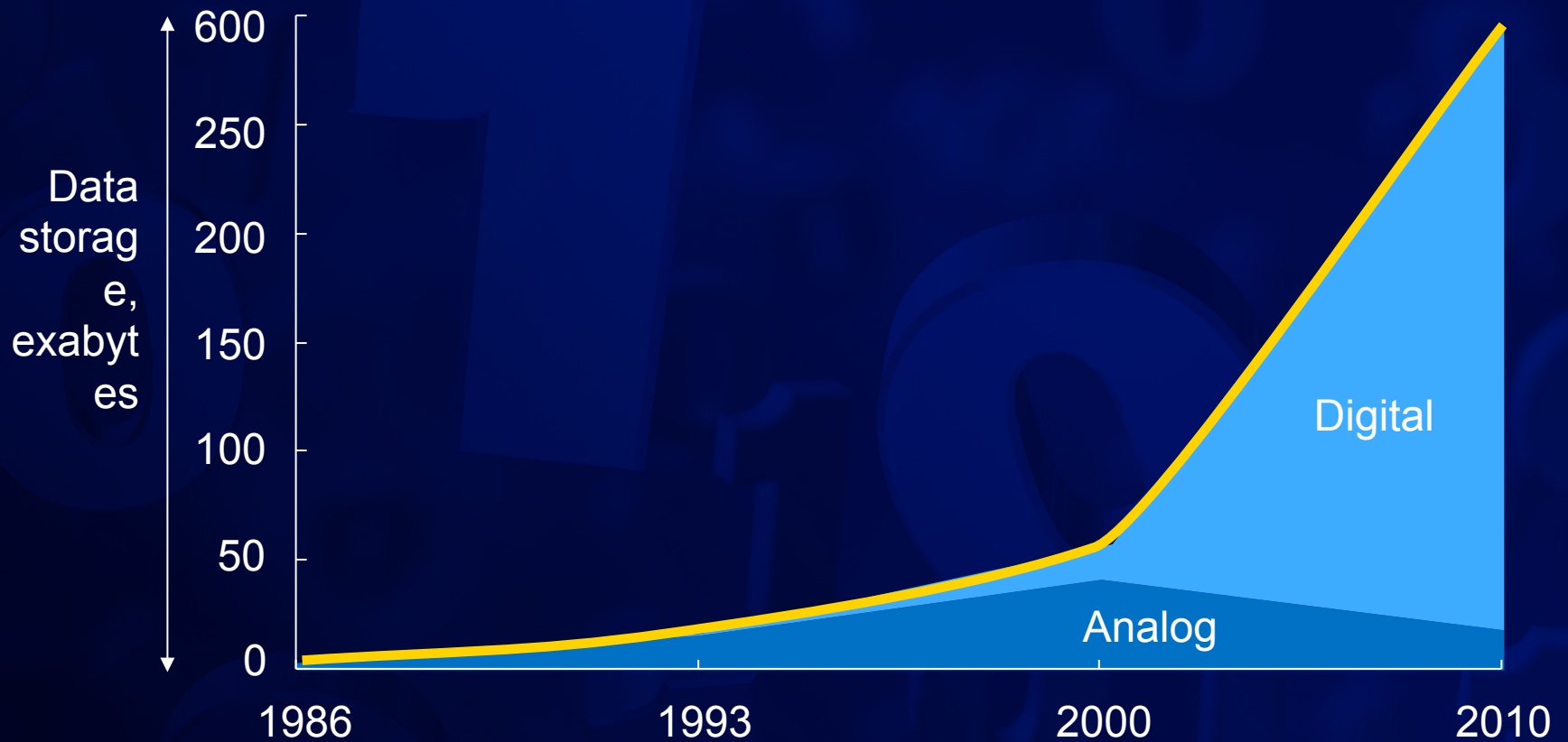**Bill Stensrud, promise and peril of big data, Aspen Institute**

V for volume: huge amount of data accessible

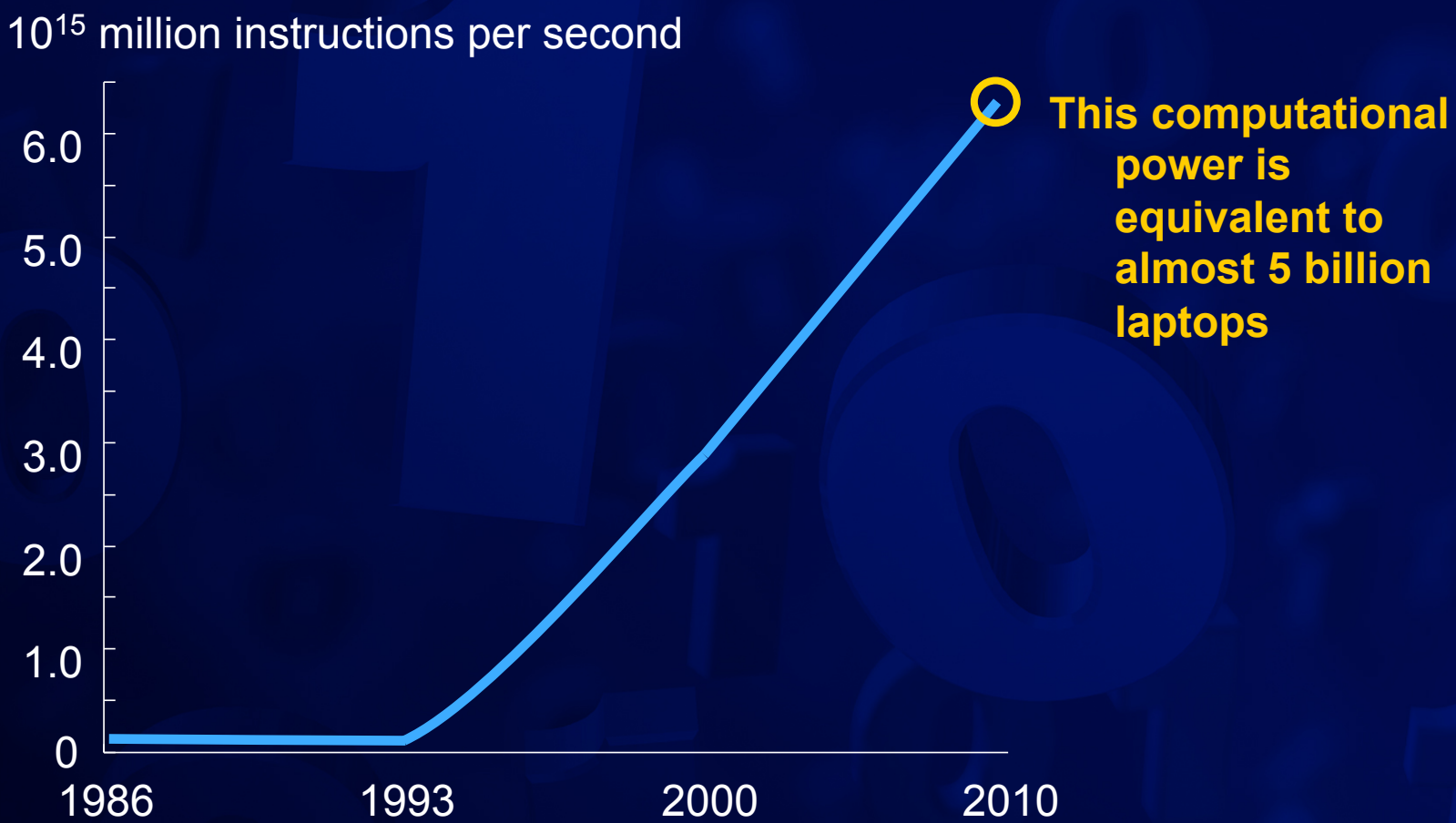**V** for velocity: huge computational power to leverage the amount of data

V for variety: large variety of un-structured sources

Data storage power evolution

Global installed, optimally compressed, storage

Data storage, exabytes

600
250
200
150
100
50
0

1986   1993   2000   2010

Digital

Analog

Computation capacity power evolution

10^{15} million instructions per second

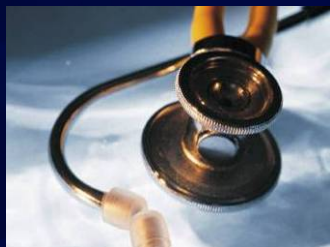This computational power is equivalent to almost 5 billion laptops

Big data – Agenda

What is big data?

Why would a company care?

How to build competitive advantage based on big data?

Big data can generate significant value across sectors

US health care
- $300 billion value per year
- ~0.7 percent annual productivity growth

US retail
- 60+% increase in net margin possible
- 0.5–1.0 percent annual productivity growth

Europe public sector administration
- €250 billion value per year
- ~0.5 percent annual productivity growth

Manufacturing
- Up to 50 percent decrease in product development, assembly costs
- Up to 7% reduction in working capital

Global personal location data
- $100 billion+ revenue for service providers
- Up to $700 billion value to end users

But more? ….

- **Big data is the new name for competitive intelligence**

- **Big data analytics is the new way to conduct and experiment businesses**

Big brother is watching you

12 trillion
- **Google** searches conducted worldwide by 2010

800 million
- **Facebook** active users worldwide, of which 10% of conversations involving a brand

**19%** proportion of **Twitter** comments referring to a brand

…And they do it within your own company

50% of Twitter account holders **tweet during their work**

75% of open source software contributors **are engaged by traditional proprietary software companies**

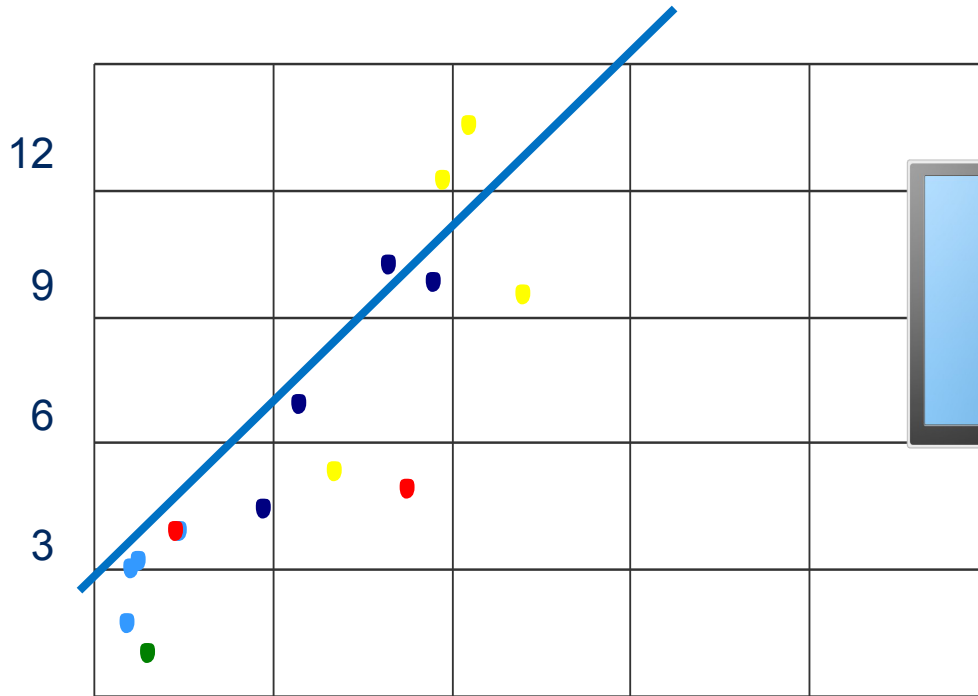# Nowcasting automotive brand sales with Google data

Monthly sales vs. Google search trends at second week of each month, US
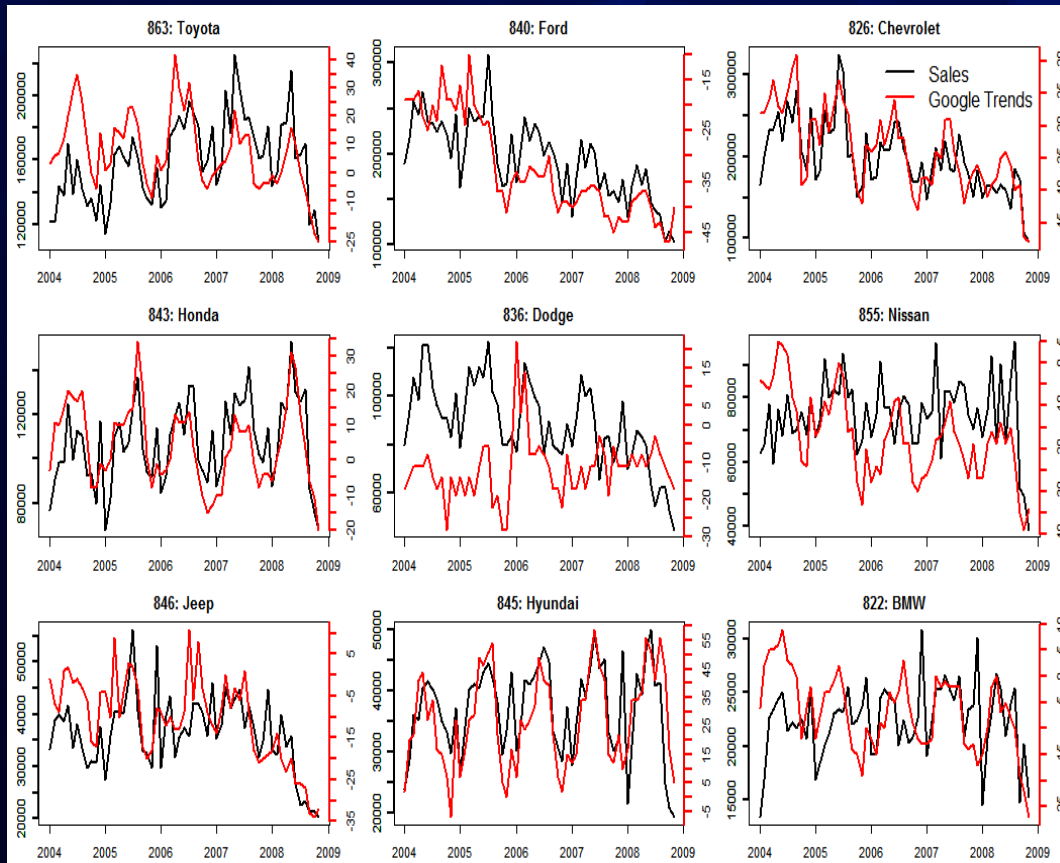
Nowcasting DTV subscription with Google data

Quarterly uptake vs. Google search trends 6 weeks in advance, Europe

Regression fit

$$Y = 2,1 + 0,21.X$$

$$R^2 = 0.78$$

$$P = 0.00$$

Search queries

But more? ….

- **Big data is the new name for competitive intelligence**

- **Big data analytics is the new way to conduct and experiment businesses**

Examples of organizations that have successfully used big data analytics

Engines for growth	Description	Examples
Analytics	▪ Uses business intelligence, advanced analytics to identify customer behavioral insights that translate into marketing programs, product offers, and customer communications	Harrah's, Walmart, amazon.com, TESCO, Capital One, PROGRESSIVE, Kroger, dunnhumby, citigroup, JPMorgan
Partner engine	▪ Helps companies reach customers cost-effectively by leveraging 3rd parties (e.g., Google's Ad-Sense network, Mutual funds networks)	Google, charles SCHWAB, Ameriprise
Research engine	▪ Helps companies to scale research and development pipelines and test ideas with consumers, e.g., P&G runs "Temor," a consumer research panel with 500,000 mothers and 250,000 youth	P&G, Johnson & Johnson
Operations engine	▪ Operations excellence driving lowest cost operations	SOUTHWEST, UPS, Walmart

The game is on

Percentage of enterprises agreeing on future capabilities", 2011

Legend: ☐ Developing countries   ☐ Developed countries

| Capability	Developing	Developed
Decision based on data rather than on opinion and experience	37	32
Hierarchy will become flatter	31	28
The boundaries between employees, customers will blur	30[1]	38
Teams will self-organize	30	34
Financial transparency will increase dramatically	25[1]	19
Strategic priorities will be set from the bottom-up	20	15
Data for decision making will be collected through experiments	18	20
Performance will be evaluated by peers rather than by managers	15	12
Large companies and/or business units will disaggregate	14[1]	8
Data from Internal and voting markets will allocate resources	13[1]	18
Employees will play a much greater role in selecting leaders	13[1]	8
Employees will have much more discretion in choosing which tasks to work on	11	12
Compensation decisions made by peers rather than by managers	4	2
Data will be more strategic,(total of above)	48	49
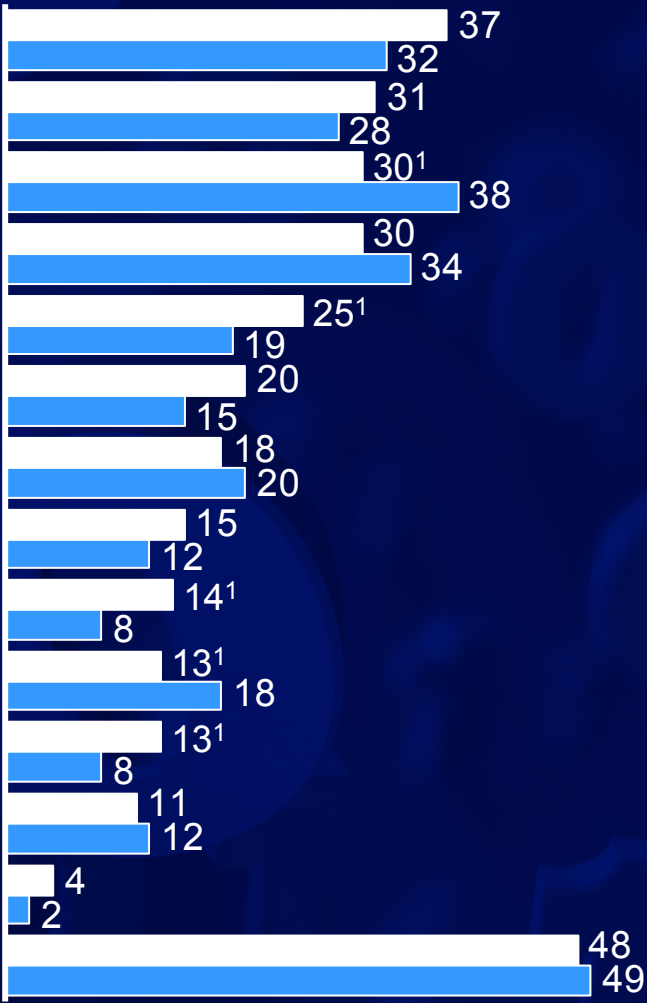
Big data – Agenda

What is big data?

Why would an organization care?

How to build competitive advantage based on big data?

Five ways for big data to create transformational value
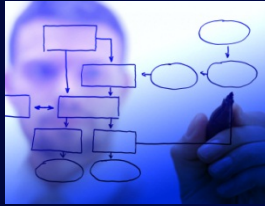
1 **Create transparency**

2 **Expose variability and enable experimentation**

3 **Segment populations to customize actions**

4 **Replace/support human decision-making with automated algorithms**

5 **Innovate new business models, products, and services**

Zillow®
Your Edge in Real Estate

- Mines publicly available data on local real estate markets – "zillions of data points"

- Applies a propriety statistically driven formula to compute an estimated market value

- Refines algorithm by back-testing historic "Zestimates" with the actual transaction prices of homes that sold

Hayes Valley
15 results.

☑ 🏠 For Sale	(15)
☐ 🏠 For Rent	(1)
☐ 🏠 Make Me Move	(11)
☐ 🏠 Recently Sold	(32)

Price
Min — Max

Beds   Baths
0+   0+

💧 Save   ▶ More filters

Sort: Featured   🗔 Map ▾   ✉ Share ▾   📧 E-mail alerts

7 Broderick St, San Francisco, CA   (Buena Vista Park)
🏠 House For Sale: $1,895,000   Beds: 4   Days on Zillow: 24
Zestimate®: $1,826,200   Baths: 3.5   Built: 1900
Mortgage: $7,684/mo   Sqft: 2,743   Single Family
▶ See current rates   Lot: 2,866   Price/sqft: $690
22 Photos ▾

2. Enable experimentation: Continuous A/B testing at Amazon

Selected examples of how "Add-to-Cart" functionality has evolved over time,
   as guided by A/B testing:

3. Segment populations: identify highest potential micro-segments

Change in duration of unemployment (months) by labor market instrument

Legend:
- Market customers
- Activation customers
- Support customers
- Counseling customers

Integration subsidy
- 2.5
- 1.9
- 0.5
- 3.0

Language courses
- 4.8
- 4.1
- 0.5
- 3.2

Job creation measures
- 15.1
- 13.8
- 11.7
- 10.6

Future vocational training
- 11.4
- 11.3
- 8.6
- 6.8

Bridging allowance
- -3.1
- 0.5
- -9.6
- -6.2

Short-term training
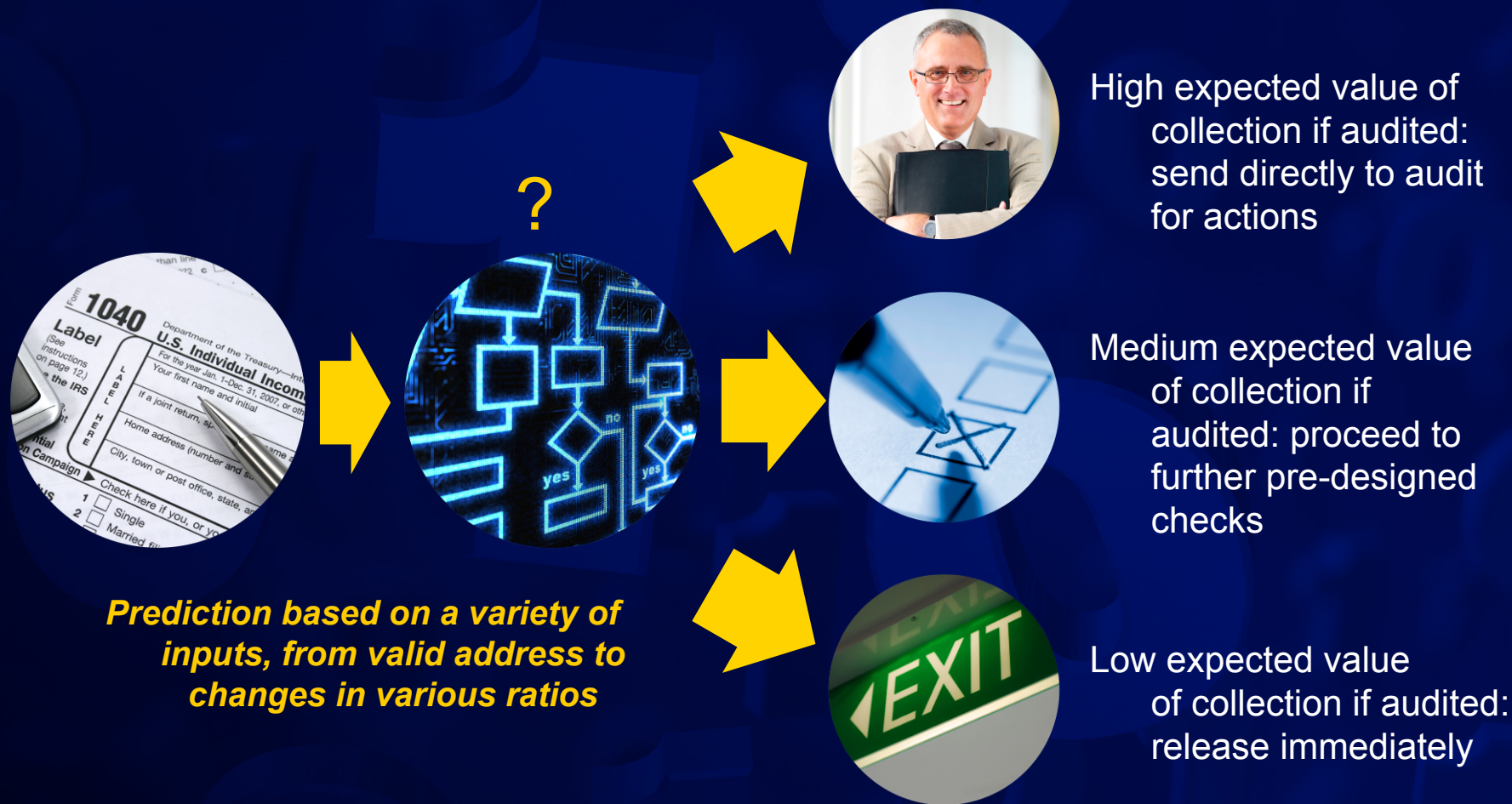- 5.9
- 4.9
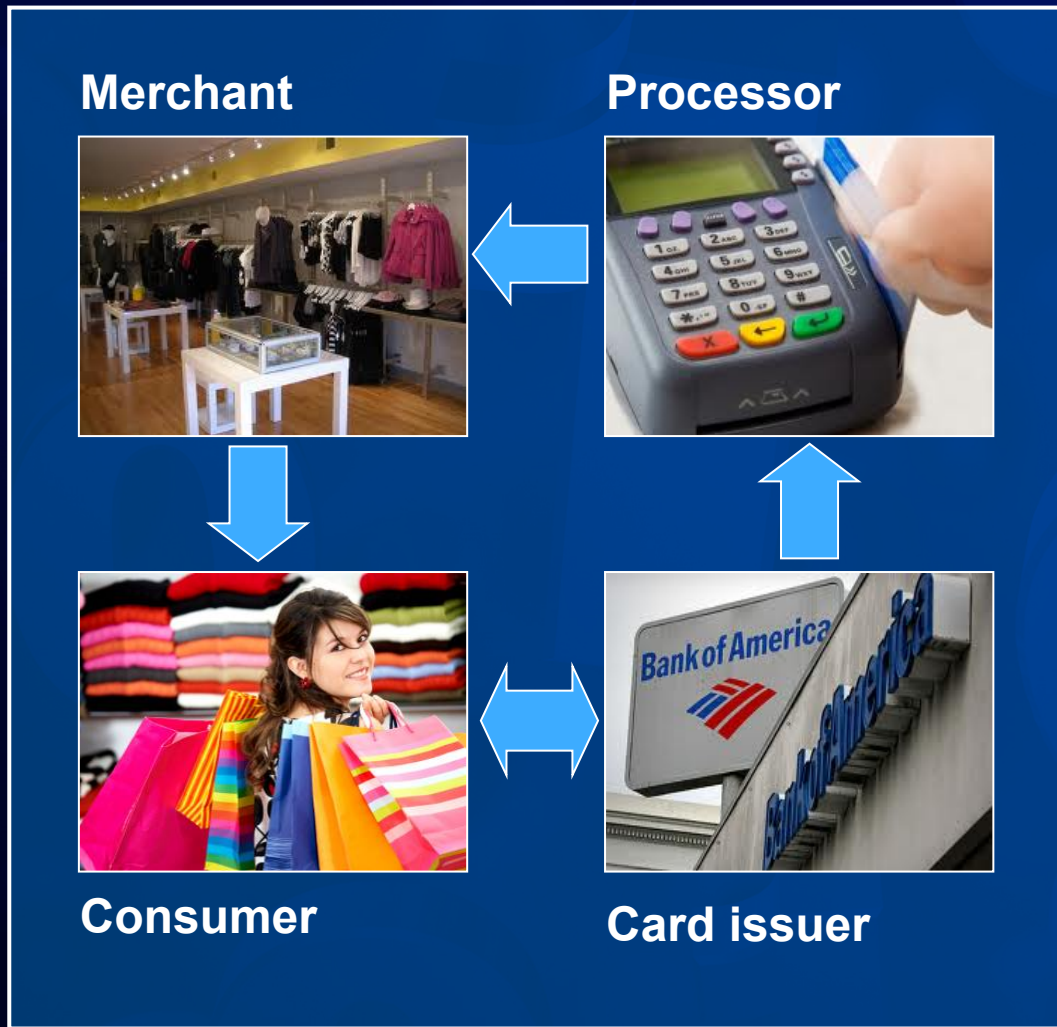- 3.4
- 2.8

Bundesagentur für Arbeit
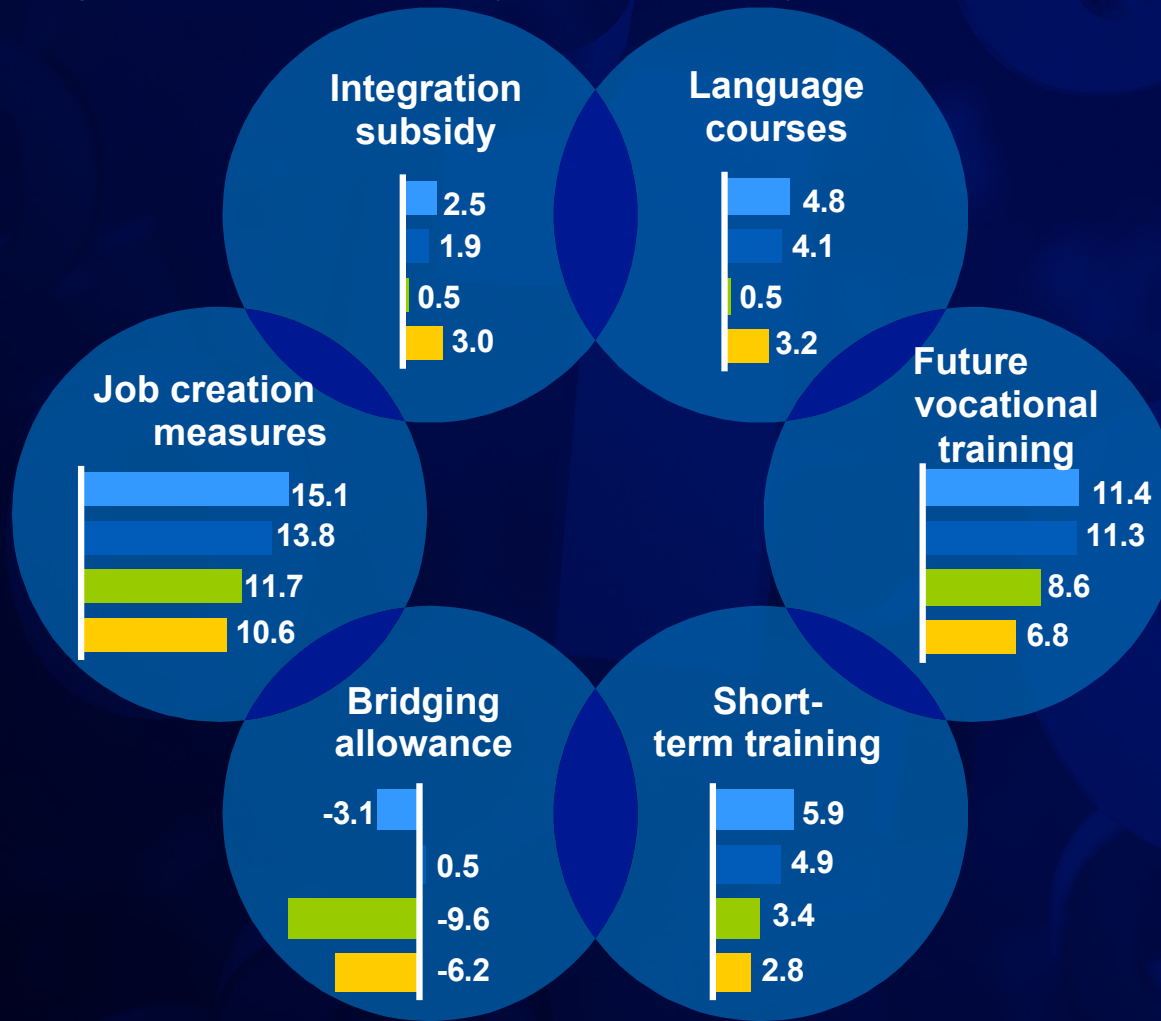
- Micro segmented approach to unemployed using a amongst other tools a skill-will matrix
- Targeting of back-to-work efforts and removing ineffective programs
- Total cost saving of €10 billion removed during the 3 years of the program

4. Automatic algorithms: Supporting human decision making: determining tax audit strategy

Prediction based on a variety of inputs, from valid address to changes in various ratios

High expected value of collection if audited: send directly to audit for actions

Medium expected value of collection if audited: proceed to further pre-designed checks

Low expected value of collection if audited: release immediately

Merchant

Processor

Consumer

Card issuer

MasterCard is finding ways to monetize the transactional data from cardholders. Combining data with analytic skills, **MasterCard Advisors**, a subsidiary of MasterCard, provides various information and analytic services for merchants

Typical issues to tackle re big data

Data policies
- Privacy concerns
- Data security issues
- Intellectual ownership and liability issues

Technology & techniques
- Deployment of technologies
- Legacy system or inconsistent data formats
- Ongoing innovation

Organizational change & talent
- Shortage of talent
- Leadership that understands big data
- Aligned workflows and incentives

Access to data
- Access to "foreign" data
- Integrating with own proprietary data

Industry structure
- Sectors lacking competitive forces
- Misalignment of incentives to invest in big data