The **DIGITAL UNIVERSE** of OPPORTUNITIES

E M C D I G I T A L

UNIVERSE

RICH DATA
& the Increasing
Value of the
INTERNET OF THINGS

GET STARTED

The **Digital Universe** Is Huge —And **Growing Exponentially**

4.4
ZETTABYTES

In 2013, there were almost as many bits in the Digital Universe as stars in the physical universe

44
ZETTABYTES

Source: IDC, 2014
* iPad Air – 0.29" thick, 128 GB

2013

If the Digital Universe were represented by the memory in a stack of tablets, in **2013** it would have stretched two-thirds the way to the Moon*

2020

By **2020**, there would be 6.6 stacks from the Earth to the Moon*

IDC
Analyze the Future

Emerging Markets Will Surpass Mature Markets by 2017

% of TOTAL
D I G I T A L
UNIVERSE

90%
80%
70%
60%
50%
40%
30%
20%
10%
0%

■ Emerging Markets
■ Mature Markets*

*US, Western Europe, Japan,
Canada, Australia, NZ
Source: IDC, 2014

2012   2013   2014   2015   2016   2017   2018   2019   2020
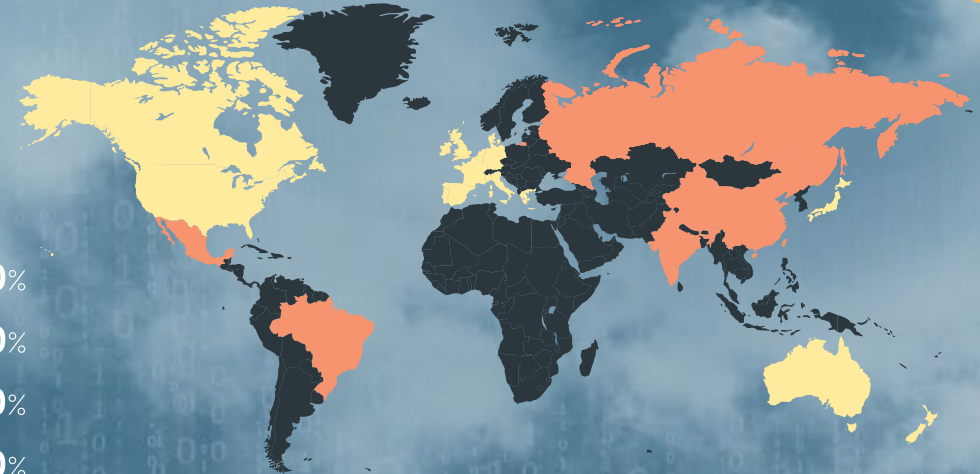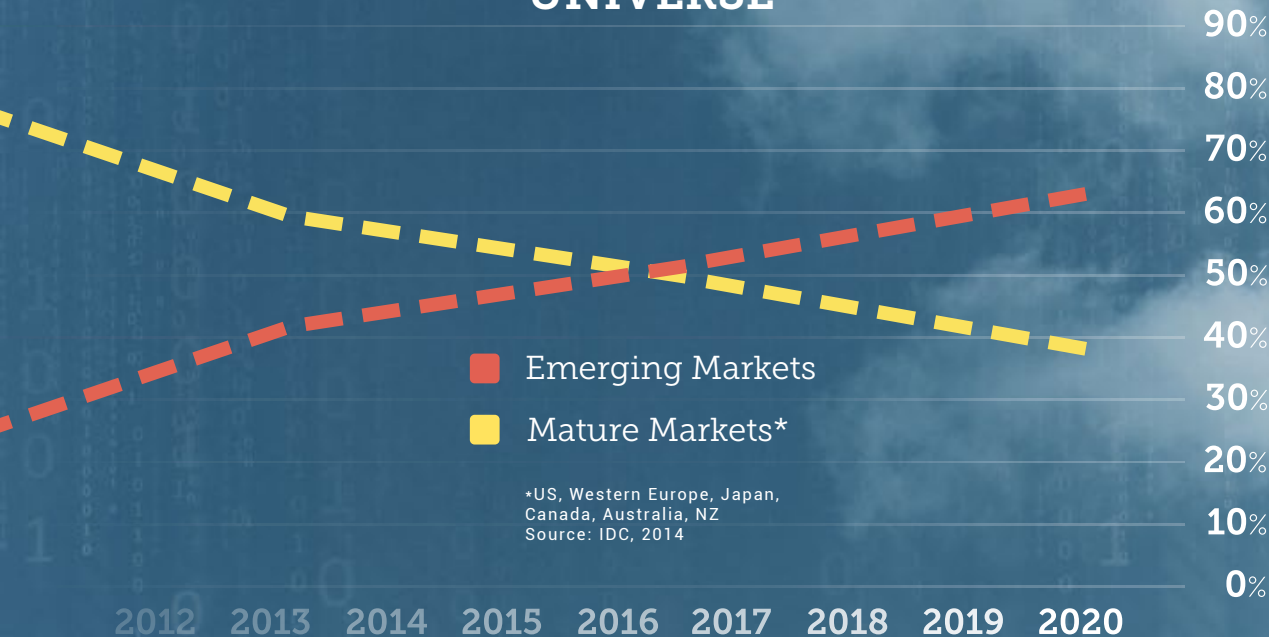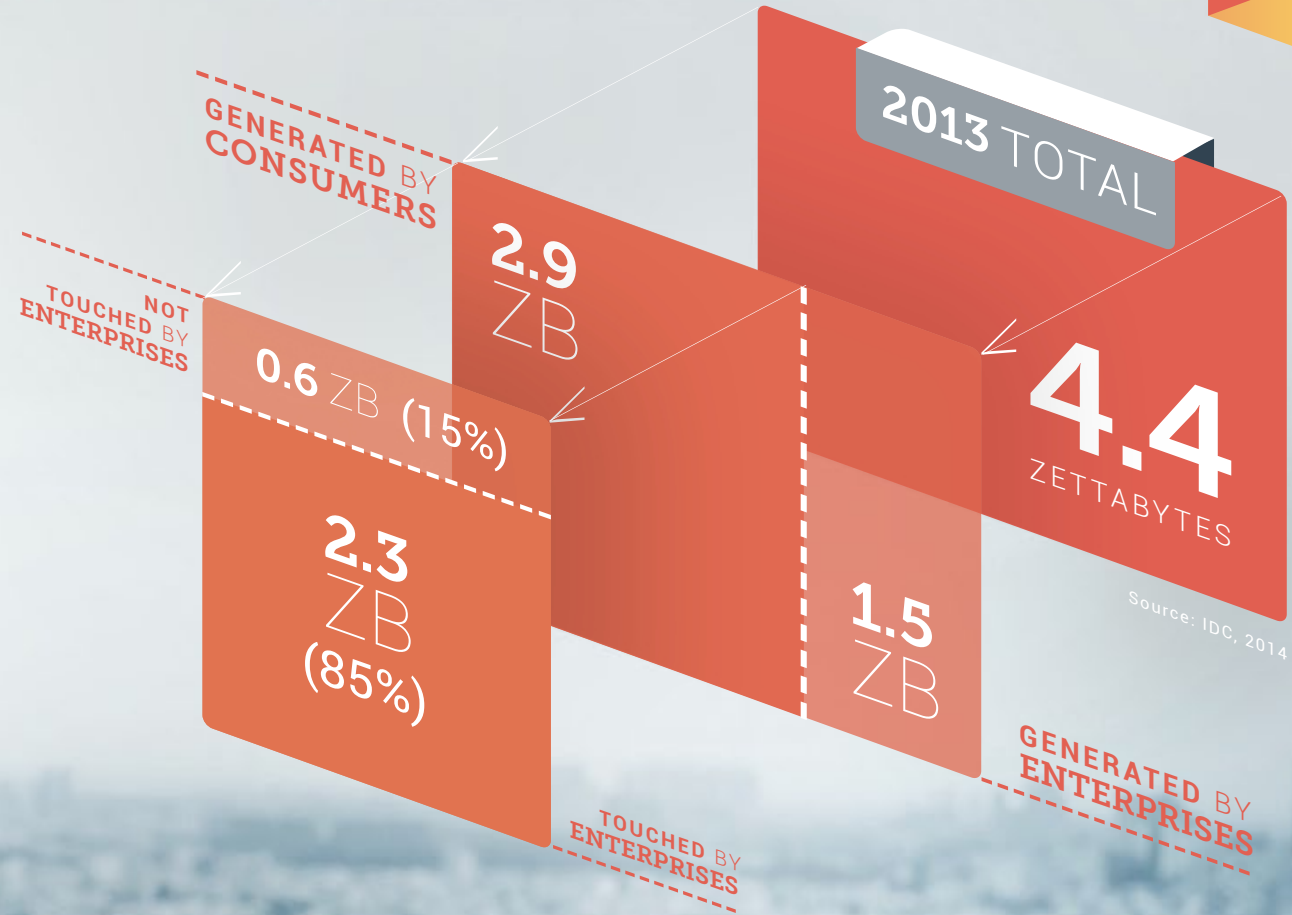
In **2013**, mature markets represented **60%** of the Digital Universe

By **2020**, that will flip-flop, with emerging markets (including China, Brazil, India, Russia, and Mexico) representing **60%**

IDC
Analyze the Future

3

2/3 of DU Is Created **by Consumers,**
but **Enterprises** Are Responsible for **85%** of This

GENERATED BY
CONSUMERS

2013 TOTAL

2.9 ZB

NOT TOUCHED BY ENTERPRISES

4.4 ZETTABYTES

(5%)

2.3 ZB (85%)

1.5 ZB

Source: IDC, 2014

TOUCHED BY ENTERPRISES

GENERATED BY ENTERPRISES

The **Internet** of **Things** Is **Exploding**

The "Internet of Things" is fueled as analog functions managing the physical world migrate to digital functions

It consists of adding computerization, software, and intelligence to things as varied as cars, toys, airplanes, dishwashers, turbines, and dog collars

While not all "things" are connected to the Internet, **20 billion** of them were in 2013, and **32 billion** will be by 2020
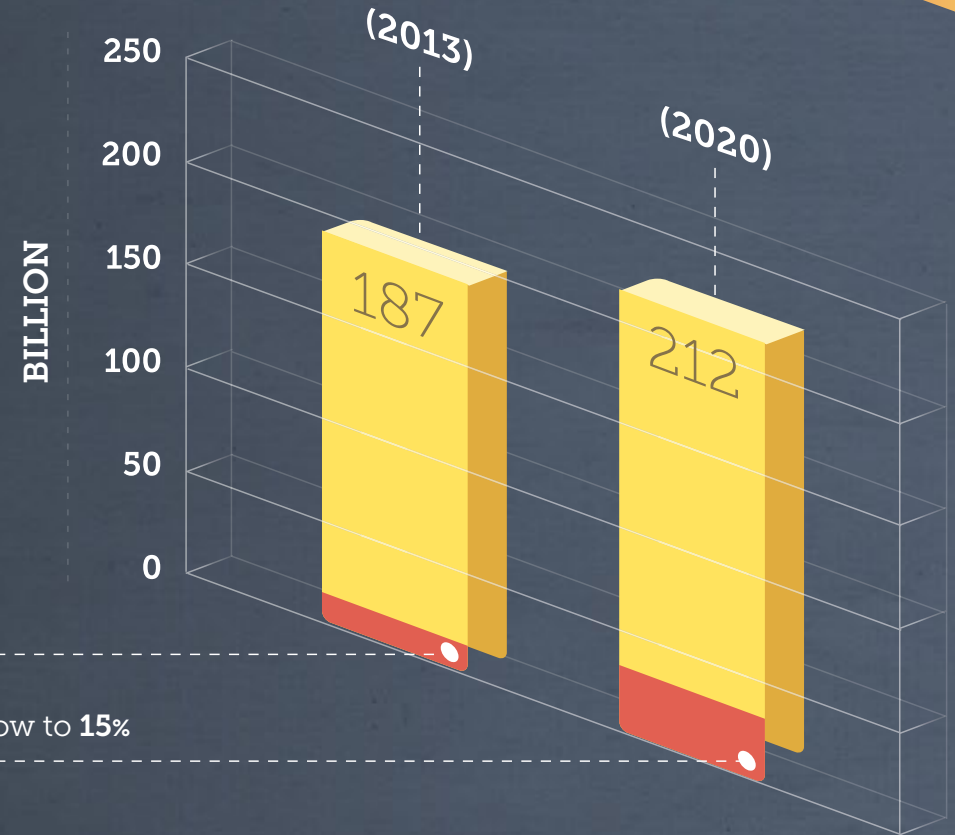
Total Number of Connectable Things

In **2013**, connected "things" were **7%** of the total

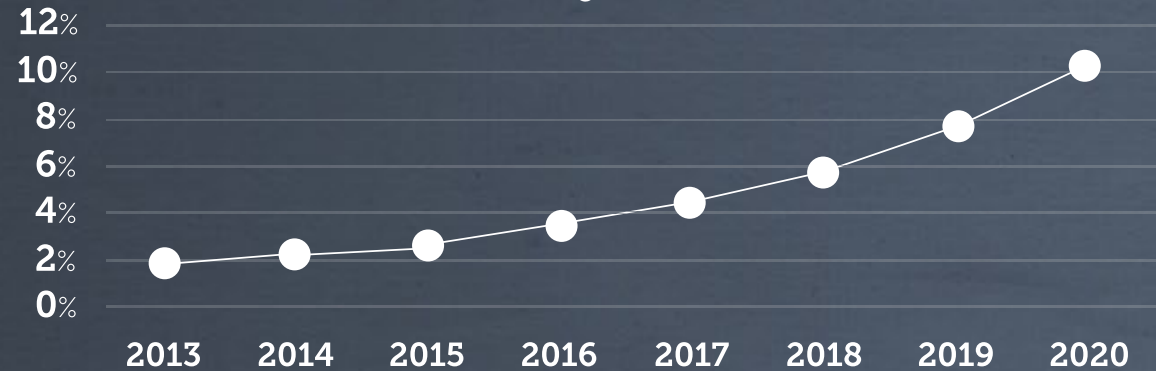By **2020**, that number will grow to **15%**

Source: IDC, 2014

BILLION

250
200
150
100
50
0

(2013)

(2020)

187

212

The Internet of Things Will Contribute an **Increasingly Large** Amount to the Digital Universe

IoT Embedded Systems as % of the DU

Source: IDC, 2014

The network connecting devices in the Internet of Things is characterized by automatic provisioning, management, and technology

It includes:

- Intelligent systems and devices
- Connectivity enablement
- Platforms for device, network, and application enablement
- Analytics and social business
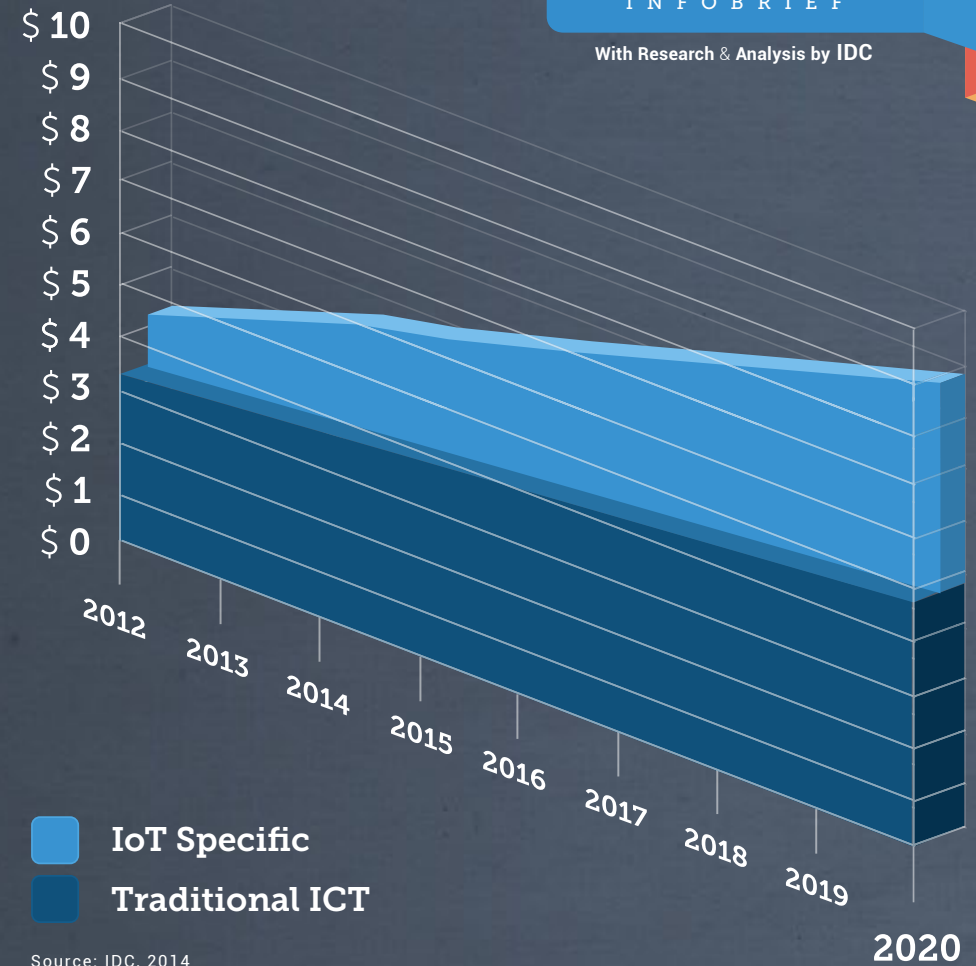- Vertical industry solutions

The Internet of Things Will **Subsume** the Information and Communication Technology Industry

Over time, the Internet of Things (IoT) will grow to subsume the traditional Information and Communication Technology (ICT) industry

IoT is growing over three times as fast as traditional ICT, and by 2020 will nearly equal all other ICT spending

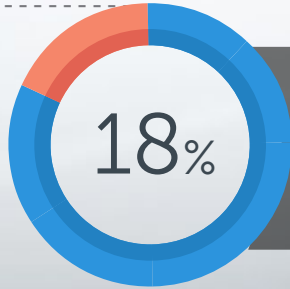Buyers and users of IoT technology and services will realize huge business benefits
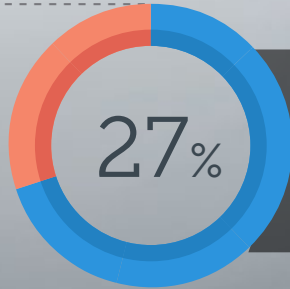

TRILLION

$ 10
$ 9
$ 8
$ 7
$ 6
$ 5
$ 4
$ 3
$ 2
$ 1
$ 0

2012 2013 2014 2015 2016 2017 2018 2019 **2020**

IoT Specific
Traditional ICT

Source: IDC, 2014

Mobility Is a Key Driver of the DU

2014

18%

Mobile "Connected Things" Generate of **18**% of the Digital Universe

2020

27%

In 2020, the figure grows to **27**%

- Generated by Mobile **"Connected Things"**
- Rest of Digital Universe

Mobile "things" include devices such as RFID tags, GPS devices, smart cards, cars, toys, and even dog collars

Source: IDC, 2014

5 Ways IoT Will **Create New Opportunities**

New business models

The IoT will help companies create new value streams for customers, speed time to market, and respond more rapidly to customer needs.

Real-time information on mission-critical systems

Enterprises can capture more data about processes and products more quickly and radically improve market agility.
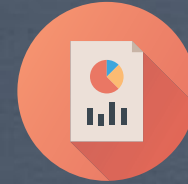
Diversification of revenue streams

The IoT can help companies monetize additional services on top of traditional lines of business.

Global visibility

The IoT will make it easier for enterprises to see inside the business, including tracking from one end of the supply chain to the other, which will lower the cost of doing business in far-flung locales.

Efficient, intelligent operations

Access to information from autonomous endpoints will allow organizations to make on-the-fly decisions on pricing, logistics, and sales and support deployment.

Five Criteria
to Extract
Maximum Value
from Data

The Digital Universe is too big and too varied for companies to make sense of all the data it contains. Fortunately, that isn't necessary. Instead, they need to target the highest value (i.e., "target-rich") data. IDC defines target-rich data using the following criteria:

Easy to access.

Can you obtain the data, or is it hopelessly locked away on end-user PCs, shuttling about on closed-end data processing systems, or trapped in proprietary embedded systems?

Real-time.

Is the data available in real-time, or does much of it come too late to drive real-time decisions and actions?
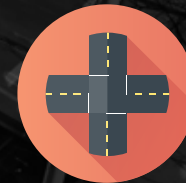
Footprint.

Could top-notch analysis of this data affect a lot of people, major parts of the organization, or lots of customers?

Transformative.

Could this kind of data, properly analyzed and acted upon, actually change a company or society in a meaningful way?

Intersection synergy.

Could this kind of data have more than one of the above attributes?

High-Value Data Is a
Manageable Subset of the Total

The size, diversity, and rapid growth of the Digital Universe can be daunting. Companies face the challenge of implementing predictive analytics, self-service business intelligence and analytics, and easy-to-use tools for data discovery and real-time decision making

The good news: companies don't have to wade through the vastness of the **entire** Digital Universe; they can find the best opportunities by focusing on the highest-value, target-rich data

1.5%

2014

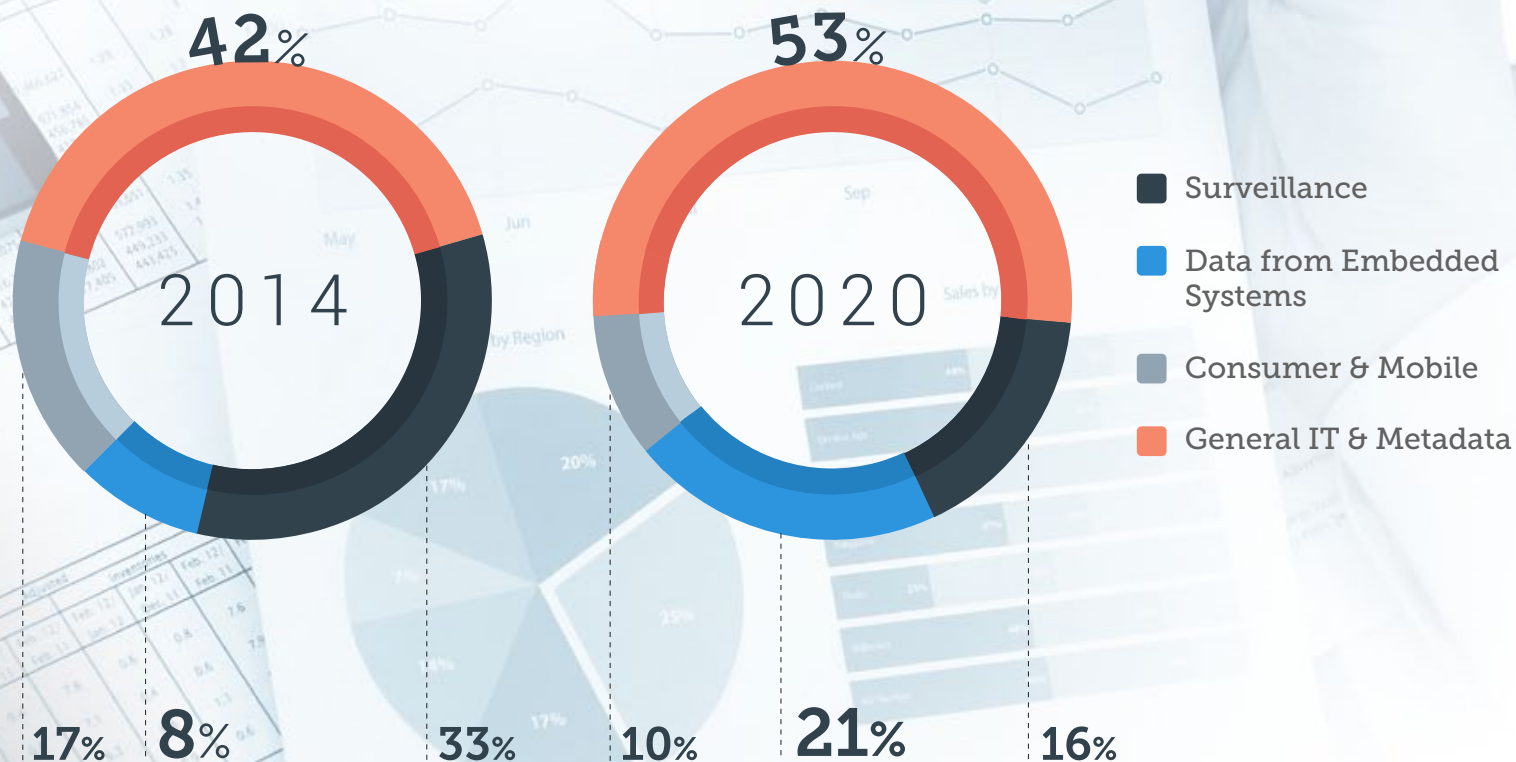At **1.5%** of the total, target-rich data is a much more manageable area of discovery

Source: IDC, 2014

Data from Embedded Systems Will Represent a Larger Percentage of "Target-Rich" Data

General IT and metadata make up the largest portion of "target rich" data and will continue to grow as Big Data projects expand and the base of metadata grows

The biggest growth is data from embedded systems, fueled by growth of the Internet of Things

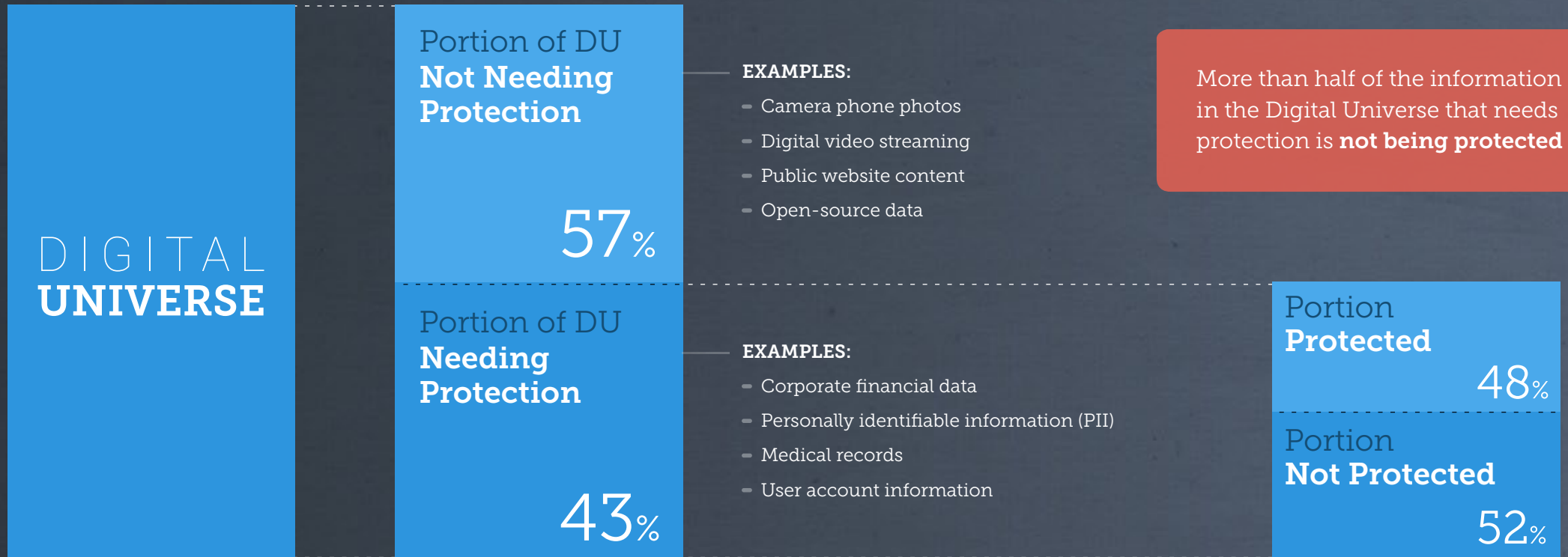The biggest decline is surveillance as the analog-to-digital transition in surveillance winds down

2014

42%

17%

8%

33%

2020

53%

10%

21%

16%

- Surveillance
- Data from Embedded Systems
- Consumer & Mobile
- General IT & Metadata

Source: IDC, 2014

IDC — Analyze the Future

12

Information Security: Much of the Data that Needs to Be Protected Is Not Yet Protected

DIGITAL
UNIVERSE

Portion of DU
Not Needing Protection

57%

EXAMPLES:
- Camera phone photos
- Digital video streaming
- Public website content
- Open-source data

Portion of DU
Needing Protection

43%

EXAMPLES:
- Corporate financial data
- Personally identifiable information (PII)
- Medical records
- User account information

More than half of the information in the Digital Universe that needs protection is **not being protected**

Portion
Protected

48%

Portion
Not Protected

52%

Source: IDC, 2014

Organization of Data: Few Organizations are at the Top of Analytics Maturity

Fewer than 1% of enterprises have achieved the highest level of Big Data and analytic usage

Big Data tends to be unstructured (e.g., in documents and text files), diversely formatted, of uncertain accuracy and unpredictable value, and often demands real-time attention

To maximize Big Data, organizations must implement new technologies and processes to change today's inflexible data structures to more egalitarian and flexible data "lakes"

					PERCENT
0.7	16.5	62.5	20	0.3	
Ad Hoc	Opportunistic	Repeatable	Managed	Optimized	

70
60
50
40
30
20
10
0

Source: IDC, 2014

IDC
Analyze the Future
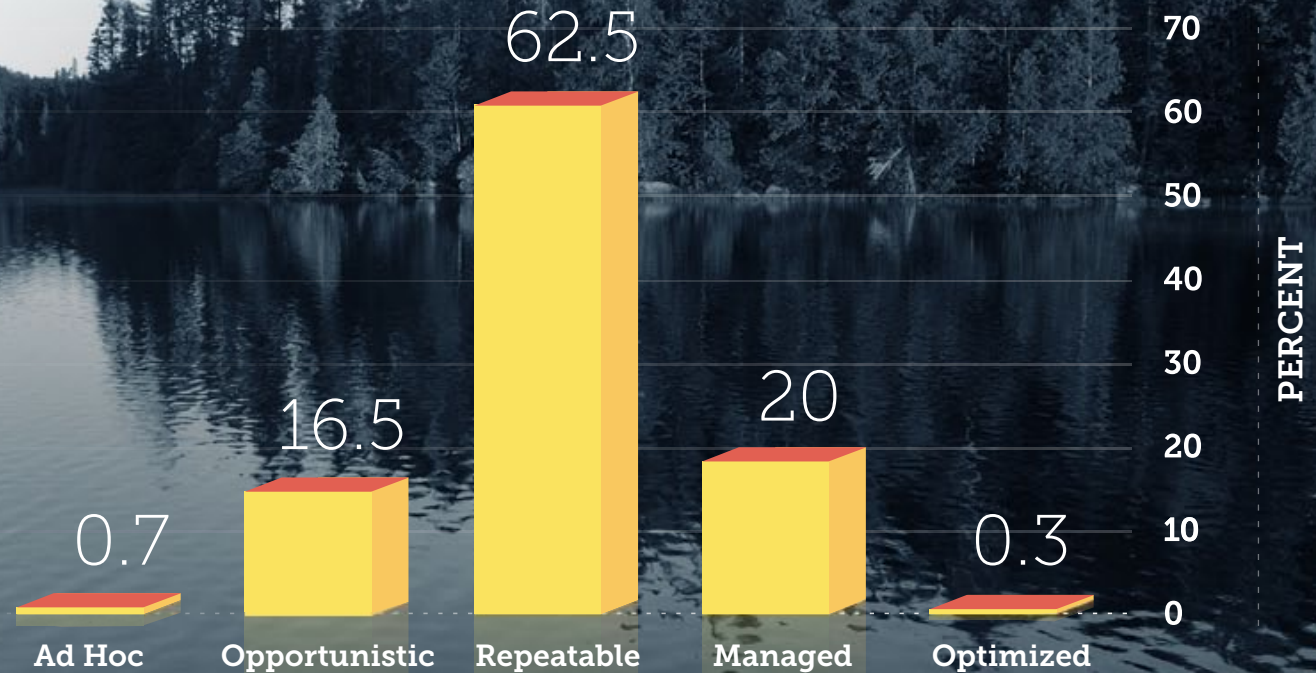
14

Talent Pool: IT Pros
Will Shoulder a Greater
Storage Burden

While much of the
IoT will be self-service and
self-supported, someone still
needs to architect the data
stores, answer helpdesk calls,
and maintain the data farms

More importantly, IT skills
and expertise need to be
upgraded to handle new data
sources and formats, and the
new technologies of today

230
GB
PER
IT PRO

28
MILLION
IT PROS
WORLDWIDE

2014

1,231
GB
PER
IT PRO

36
MILLION
IT PROS
WORLDWIDE
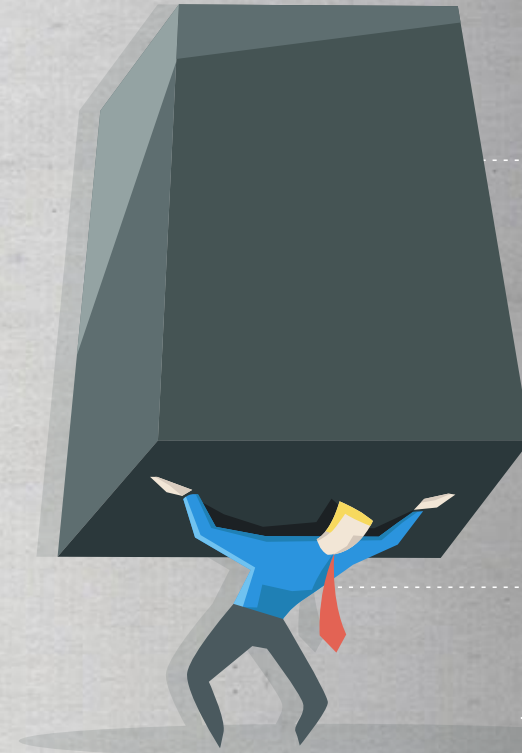
Source: IDC, 2014

2020

Three Steps All Enterprises Must Take

Many of the biggest challenges posed by the digital universe are organizational. Three steps organizations should take to survive and thrive in the new era are:

Define and implement an enterprise-wide data **governance policy.**

Put in place a central governance policy to determine who owns the data, who has the right to access it, where is the data, and what are the compliance, privacy, security, and other risk factors associated with the data.

Assess and select the right **software tools.**

To manage the data deluge, you must choose and deploy the right next-generation software tools for data cleaning, crunching, and consumption, and seamlessly integrate them with legacy systems.

Design and execute a plan for acquiring the required **skills and talent.**

Define the skills and expertise you need today and will need tomorrow and establish the right processes, programs, and incentives to upgrade your workforce.

Methodology

This is the seventh time IDC has conducted the Digital Universe study for EMC. It was—and still is—the only study to estimate and forecast the amount of digital data created annually. It has used the same methodology since its inception, allowing the size of the Digital Universe to be traced all the way back to 2005, when "only" 132 exabytes of data were created and replicated.

Our basic approach to sizing the Digital Universe is to:

- Develop a forecast for the installed base of any of 40 or so classes of device or application that could capture or create digital information.

- Estimate how many units of information—files, images, songs, minutes of video, calls per capita, packets of information—were created in a year.

- Convert the units of information to megabytes using assumptions about resolutions, compression, and usage.

- Estimate the number of times a unit of information might be replicated, either to share or store. Much of this information is part of IDC's ongoing research.

AVAILABLE STORAGE

IDC routinely tracks the terabytes of disk storage shipped each year by region, media, and application.

To determine available storage on hard drives, IDC storage analysts estimated storage utilization on capacity shipped in previous years and added that to the current-year shipments.

For optical and nonvolatile flash memory, we developed installed capacity ratios per device and algorithms to calculate capacity utilization and overwriting. In optical, we found there was much more prerecorded storage than storage that was overwritten by users.