Information Economics

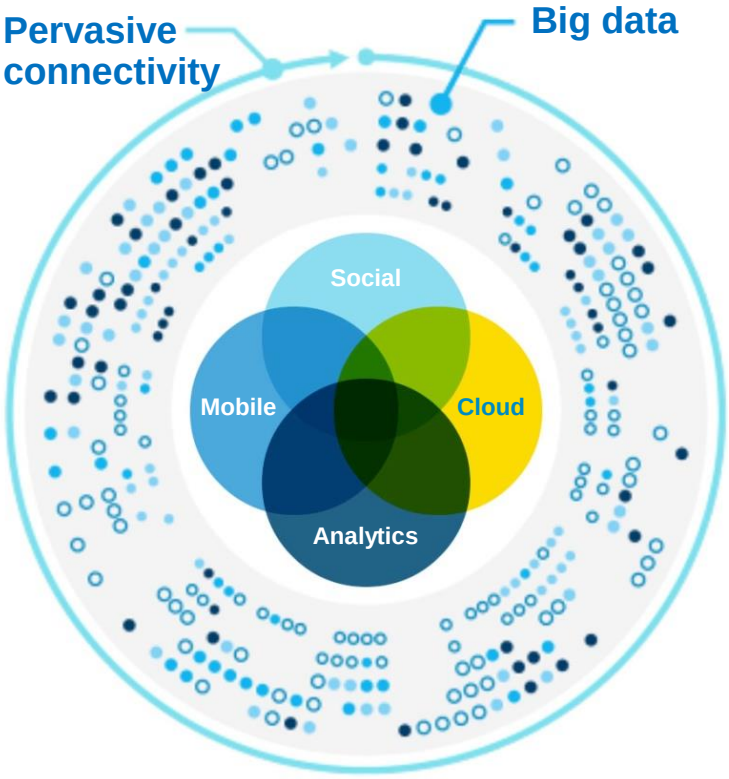
Understanding and getting value from your unstructured data

Making sense of the other 80%

Isuru Fernando, Information Economics Advisor

The emergence of big data and analytics, social, mobile and cloud are fundamentally changing how we live, work and interact

Digital transformation forces



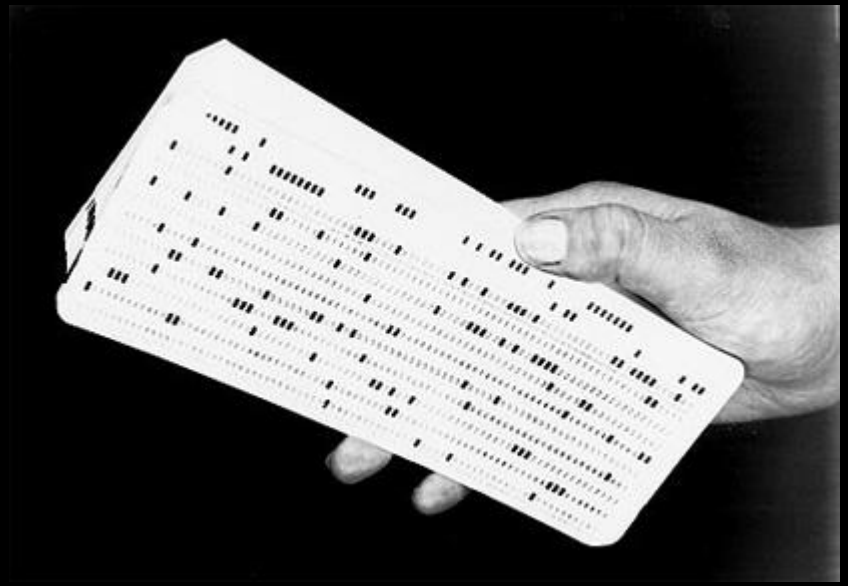
67% of global consumers
want to use mobile devices to complete retail transactions

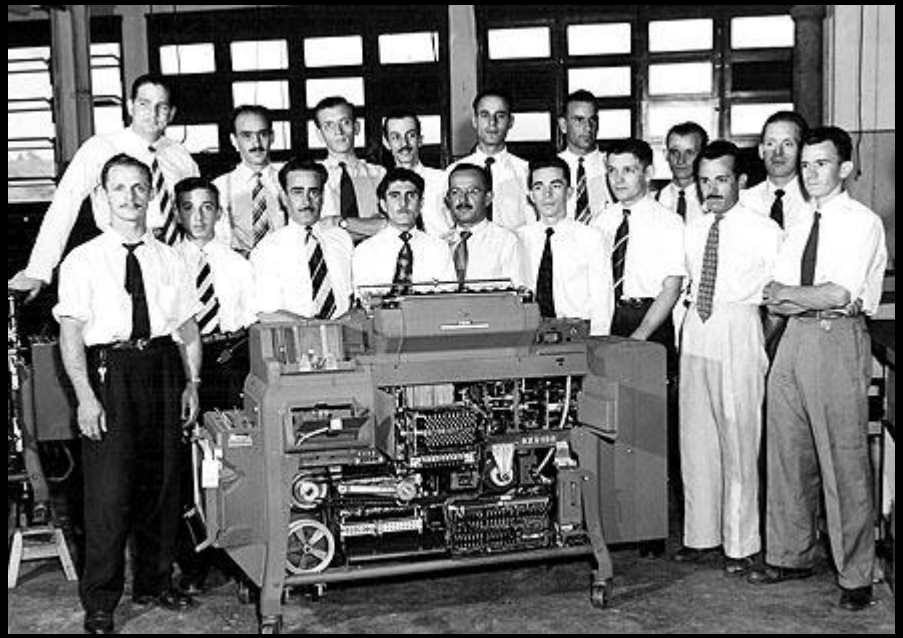
2.5 billion gigabytes of data
generated every day

1/3 of consumer data
will be stored in the cloud by 2016

1.3 billion global mobile workers
by 2015, representing nearly 35% of the worldwide workforce*

Story …





Random Access Method of Accounting and Control

1956: IBM ships the first hard drive in the RAMAC 305 system.

The drive holds 5MB of data at $10,000 a megabyte. The system is as big as two refrigerators and uses 50 24-inch platters.

IBM 305
RAMAC

… and are fueling an explosion of data

In just two days

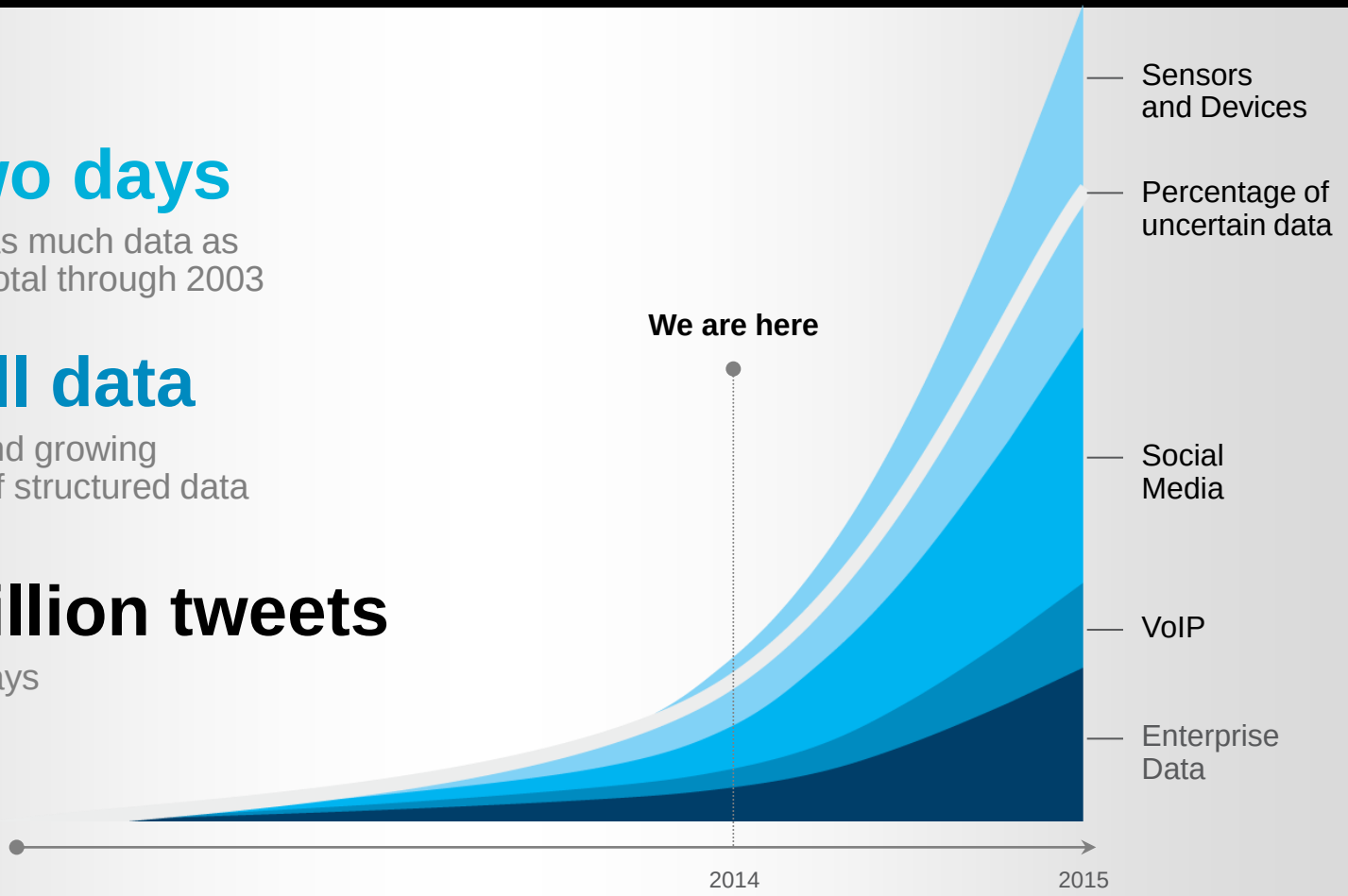
we now generate as much data as
was generated in total through 2003

4/5th of all data

is **unstructured** and growing
15 times the rate of structured data

Over 1 billion tweets

are sent every 3 days



“I think it [data] will be the next generation's natural resource..”

– Ginni Rometty, IBM CEO

“We have for the first time an economy based on a key resource [Information] that is not only renewable, but self-generating. Running out of it is not a problem, but drowning in it is.”

– John Naisbitt

> The best way to reduce the amount of data—delete it.

—Sheila Childs, research VP, Gartner

Progressing from the Possible to the Proven

Government achieves significant cost savings and ability to react to potential threats quickly

Government cuts acoustic analysis from hours to 70 Milliseconds

Utility provider improves prediction of power outages

Utility avoids power failures by analyzing 10 PB of data in minutes

Hospital detects and intervenes in potentially life-threatening conditions

Hospital analyzes streaming vitals to intervene 24 hours earlier

Retailer optimizes inventory levels and product mix

Retailer reduces time to run queries by 80%

Stock exchange reduces time to insights to achieve optimal buying / selling strategies

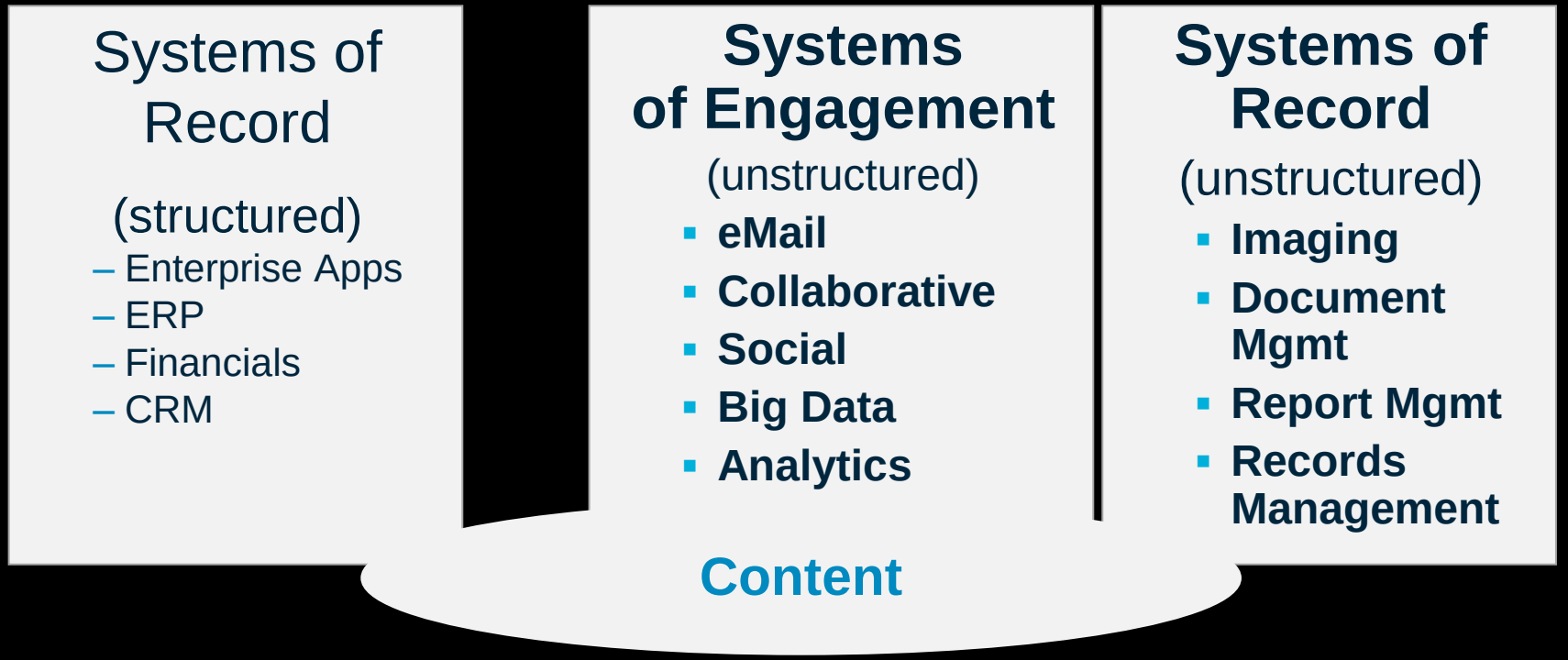
Stock Exchange cuts queries from 26 hours to 2 minutes on 2 PB

Telco provider improves ability to quickly address network issues / opportunities

Telco analyses streaming network data to reduce hardware costs by 90%

Progressing from the Possible to the Proven

In the Enterprise



Content lives everywhere and is growing faster than any other data

Information Economics

Understanding and optimising information value and costs

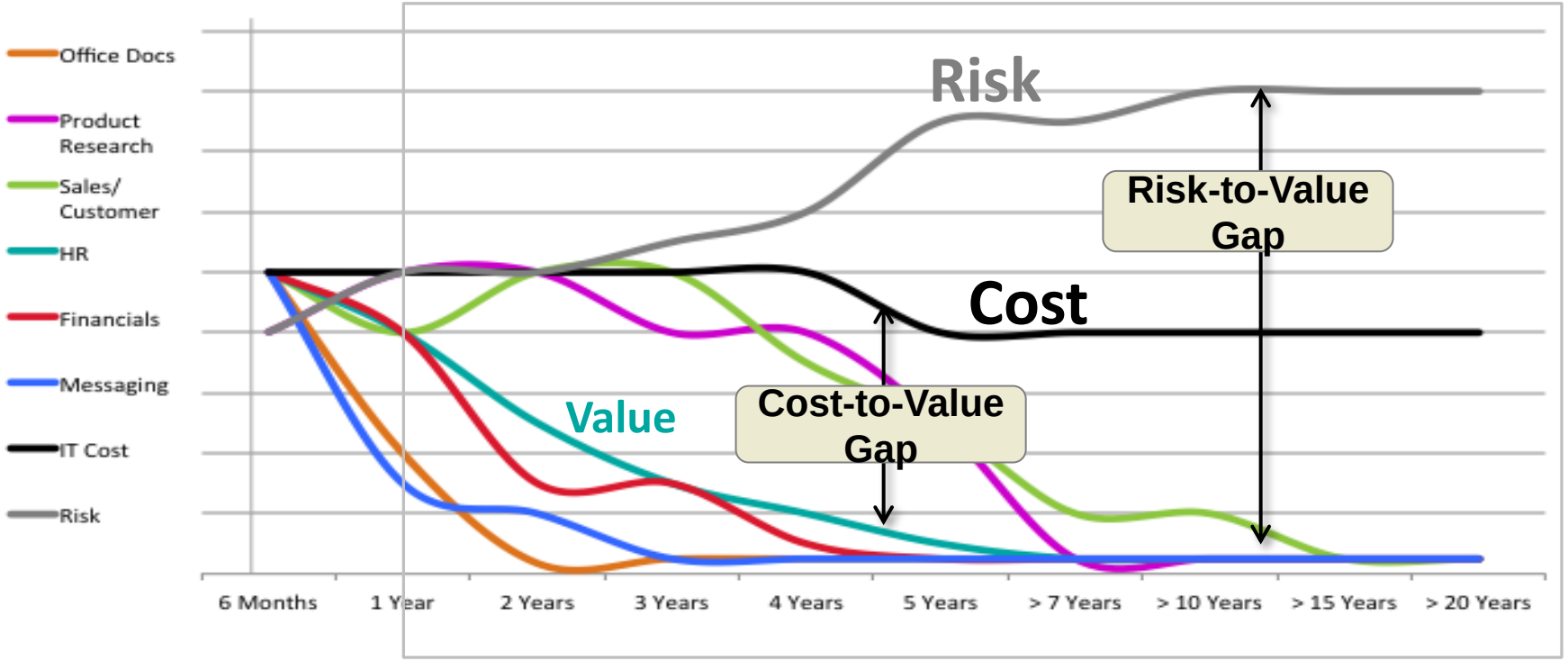
- **Information volume doubles every 18-24 months in most organizations**
- 90% of the world's information was created in the last 2 years[1]

- **$4M to store 1PB and app cost materially adds to run rate**
- Data storage consumes growing share of IT budget thus impacting transformation budgets

- **It costs $18,000 to do e-discovery on 1 gigabyte**[3]
- e-discovery consumes as much as half of the litigation budget

How do we use new technologies and best practices to get more value?

Information value declines over time, but cost and risk do not



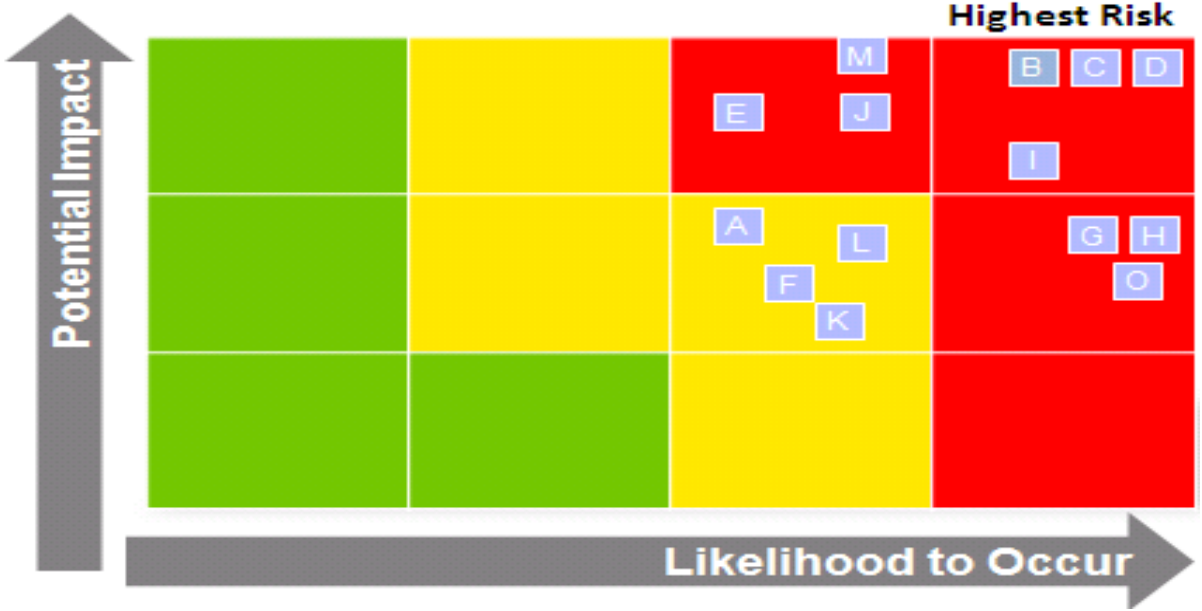
1 Source: Big data: The next frontier for innovation, competition, and productivity McKinsey & Company, 2011 Study
2 CGOC 2012 Summit Survey

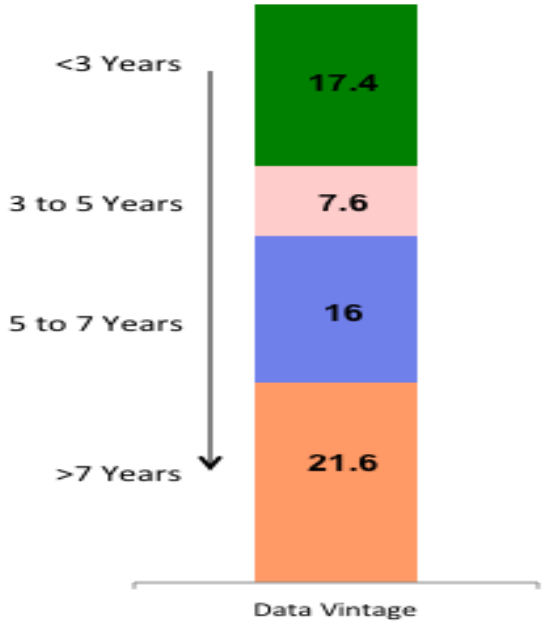
Data that has **aged past its usefulness**

Data that is **consuming COST without providing value** to the organization

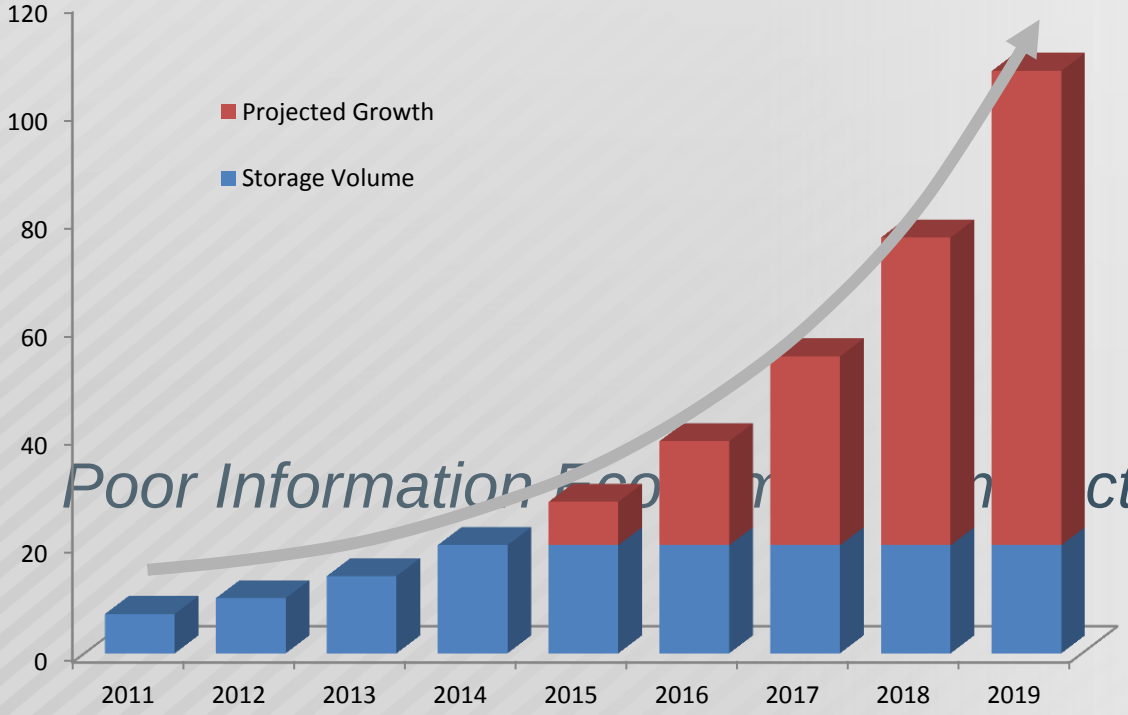
Data of **low value that still carries RISK** in legal actions

Likelihood and Impact of Risk are greatest for data kept longest





- *Growth is out of control*



Poor Information Economics Impacting CIOs, Risk & Legal

40%-60% Annual Growth

Storage Costs are Consuming IT Budget

Supply-Side Initiatives are only a temporary stop-gap

- **Virtualization**
- **Over-Allocation**

Addressing Information Economics Challenges

Can't get value

Over pays for duties

Inability to implement

Over supplies legal, under supplies business

Business under-served, legal flooded with data, and IT over-pays for infrastructure service

“The best way to reduce the amount of data—delete it.”

—Sheila Childs, research VP, Gartner

Excess information = higher cost and greater risk

Dispose of unnecessary data = **reduce** cost and risk

Legal hold: Collect and retain evidence

1%

Regulatory record keeping:
Hold and dispose of data

5%

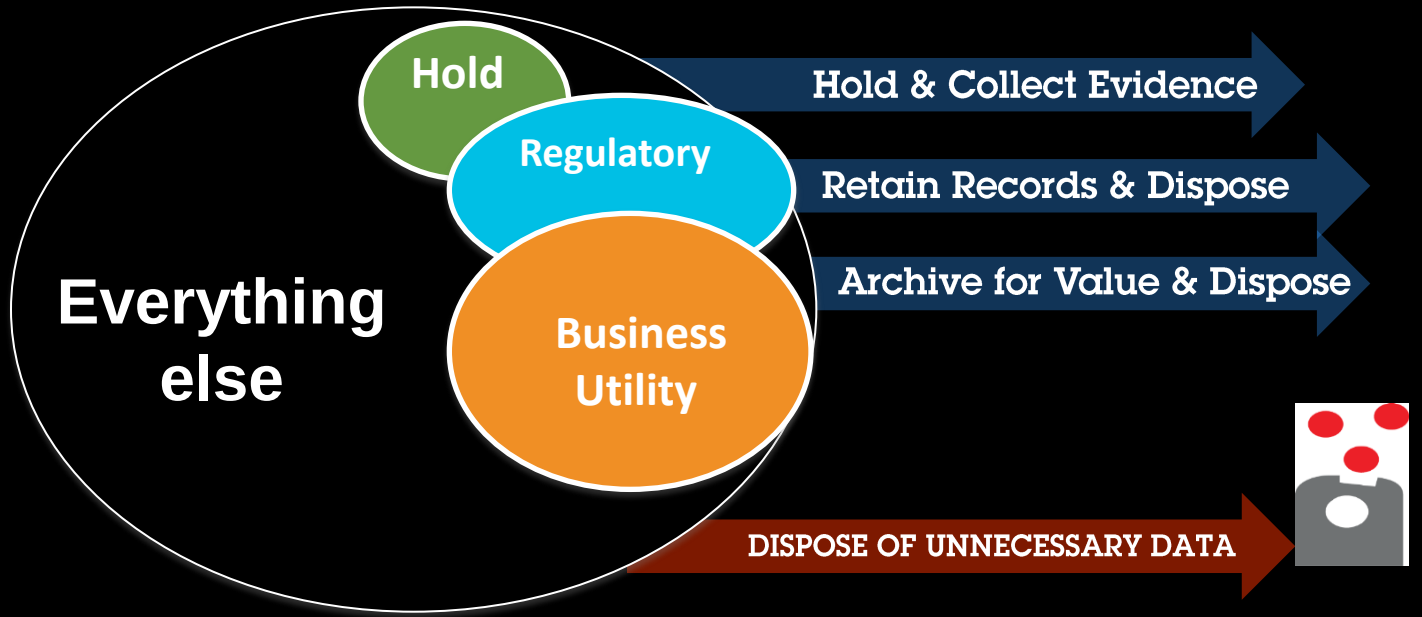
Business value:
Archive records for value

25%

69%

Information without duty and value = unnecessary data

Simple Proposition: Identify & Dispose of Unnecessary Data



Cost & Risk Reduction Enables Disposal

Cost Reduction Normalizes growth curve

%s based on CGOC Summit 2012 Survey

Where do we start?

GREENPEACE

A court Vetted Methodology and Solution

Supported largest litigation case in world by identifying, collecting and analyzing 132 TB of data to produce 200GB of relevant data.

PROBLEM: For *The Deep Water Horizon* matter, look across 132TB's, 3 continents and 8 locations. Collect 1TB to a preservation location in Houston. Full text indexing and apply additional terms to reduce to the smallest defensible data set which was sent out for production review by outside counsel. Final data set was approximately 200 GB's."

SOLUTION: Enable a 100:1 reduction in collection process in less than 2 weeks.

ROI: Saved million of dollars, responded to every DOJ request; substantially lowered outsourced review costs and built a defensible audit trial.

Novartis AG Increase Disposition Capability Ten-fold

10 X
Disposition Capability

6 Weeks
Transformed Retention Maturity Level from 1 to 4

Business Problem:

- Required an efficient, defensible approach to retain information of business value or for regulatory requirement
- Preserve information needed for litigation
- Discard unnecessary information

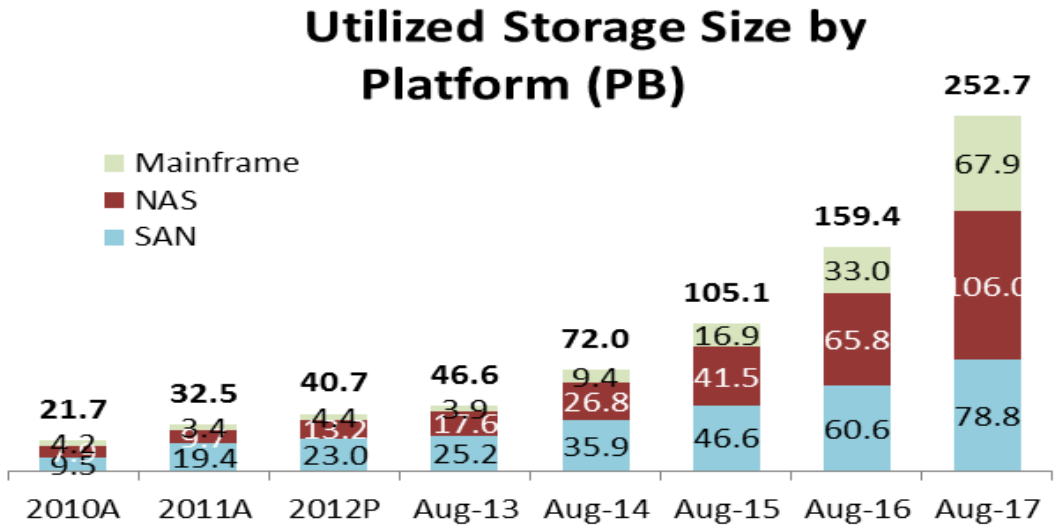
Solution Benefits:

- Ten-fold increase in ability to dispose of unnecessary information
- Reduce litigation and compliance risk with defensible, routine disposal of unnecessary information
- Lower litigation and regulatory compliance risk
- Lower cost with defensible, routine disposal of unnecessary data not needed for legal or business reasons

Financial Services Case Study - Cost

FIRM utilized storage requirements were growing more than 6X from 40PB to 250PB in 5 years

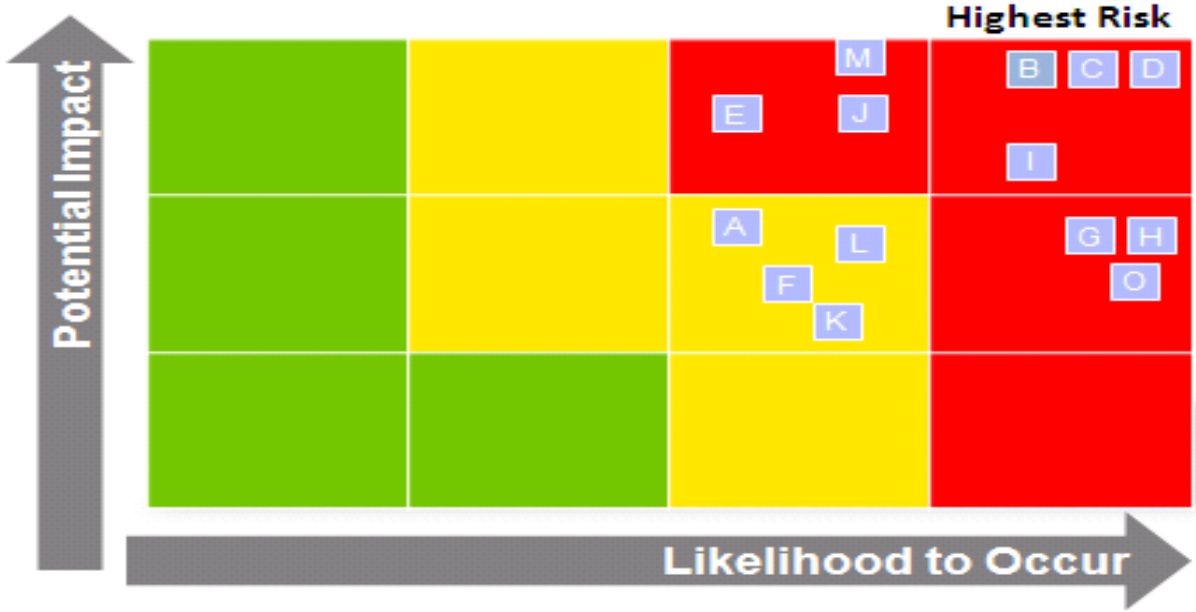
- SAN dominated storage
  - Roughly doubled every year since 2008
  - Grown more than 10X from 2007-2011
  - Makes up 60% of storage vs 10% in 2007



Financial Services Case Study – Risk

“Demand Management” processes had not matured to reflect increasing volumes

- Difficulty disposing of unnecessary data
- Complexity in applying legal holds
- Challenge to ensure record keeping compliance
- Inefficiencies in data management
- Inability to align IT with information value



16 key processes as manual, often ad hoc - Demand and supply processes - 12 at level 1 (ad hoc, manual), 4 at level 2 (common but manual) of 4, represented in A-O.

Find the data that matters

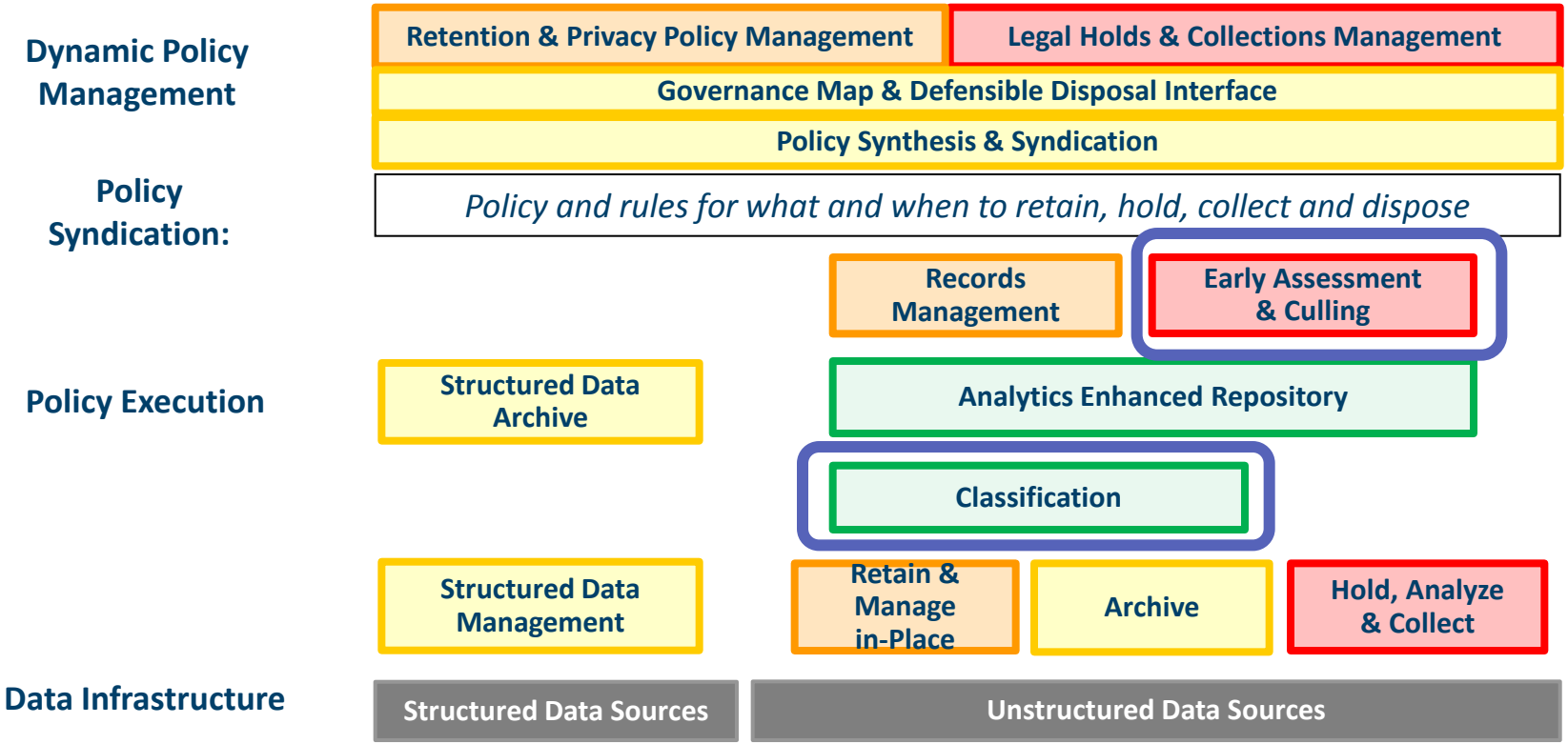
Get rid of old, obsolete data

Identify sensitive and toxic content

Stratify information to accelerate

Business process readiness

Solution Vision to Improve Information Economics

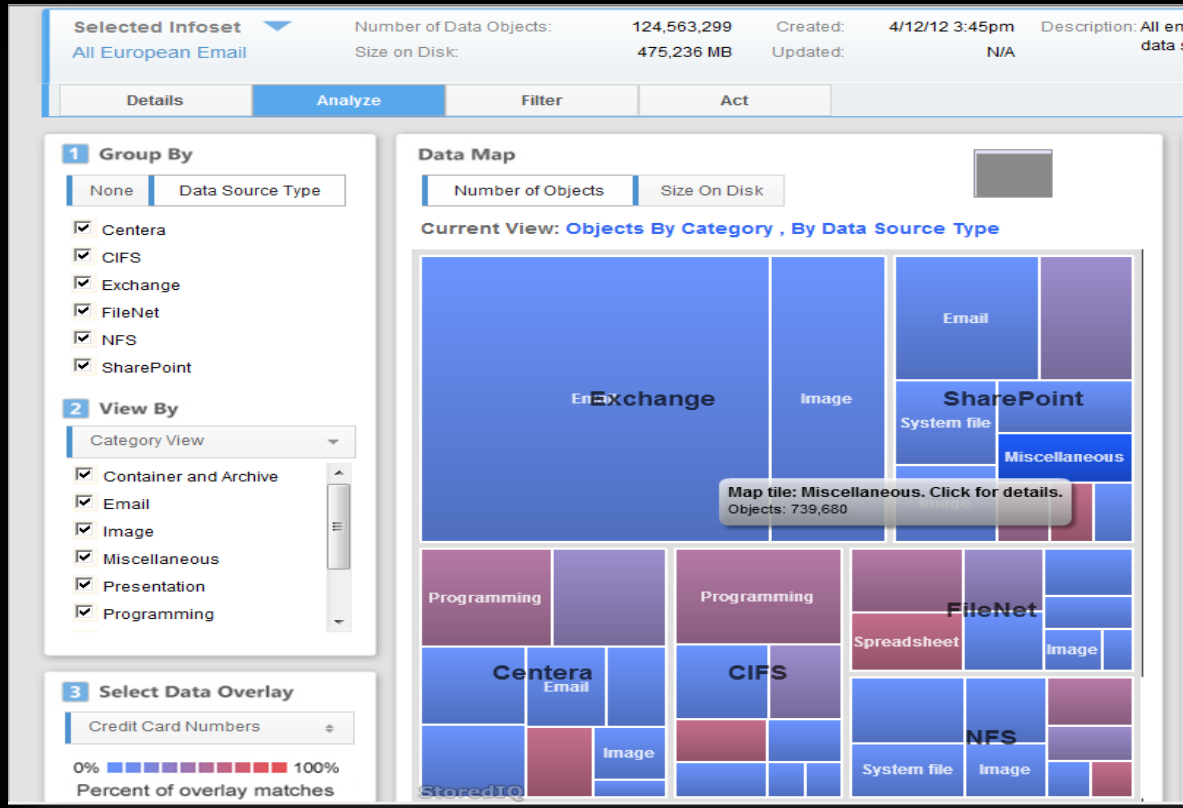


Data about Data

Instrumentation for Identification & Assessment

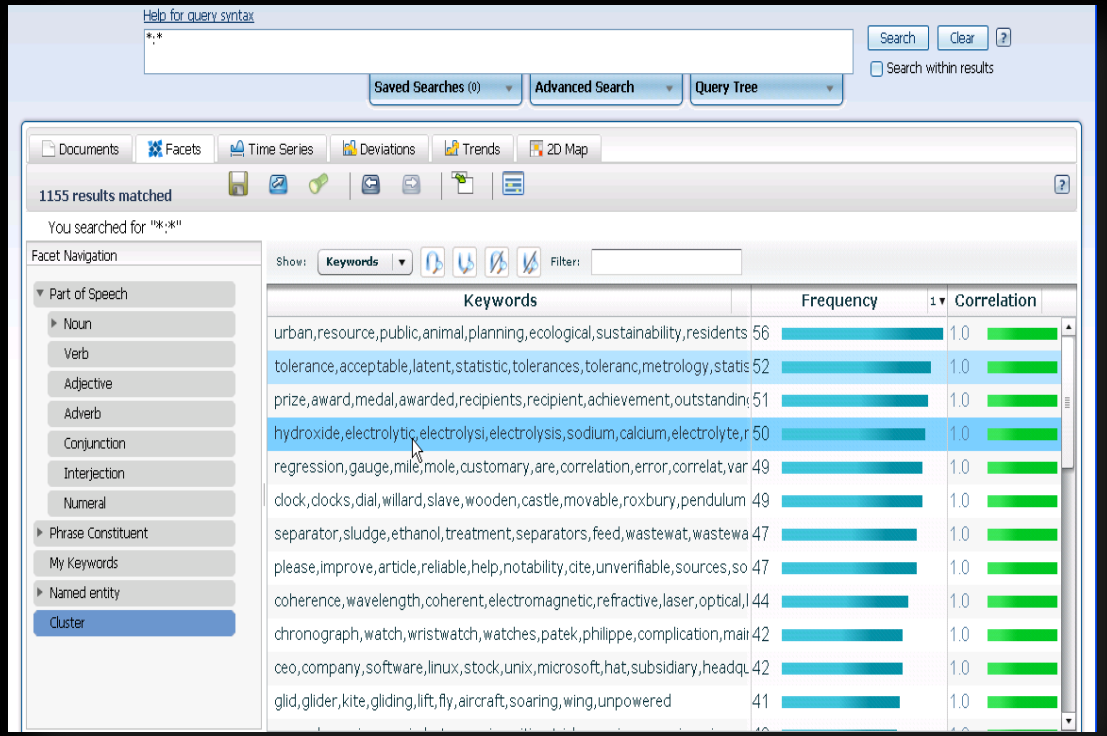
Crawl and identify information types and key data in unstructured documents while “Managing in Place”

- Redundant, Obsolete & Trivial
- 15-25% of data
- Privacy
- Compliance



Automated Classification of Unstructured Information

Artificial intelligence based classification to perform fine grained analytics, scoring & similarity analysis



IBM Experience

- **Defined the Defensible Disposal market** which is now championed by several vendors and analysts in the industry.
- **Provides strategy, expertise, and technology** to operationalize Defensible Disposal projects preferred by major clients in Banking, Insurance, Financial Markets, Oil & Gas, Life Sciences.
- **Experts** to assist in Defensible Disposal initiatives:
  - Experts with insight into information stakeholders' functions including IT, Legal, Records Information Management, Privacy, Security, and the Business.
  - Open Industry Leadership, Contribution, and Insight
    - Founders of the CGOC
    - EDRM Advisory Board Membership
    - IGRM Co-Leadership and working group members
    - Watson, Big Data and Analytics Innovation

CGOC
THE COUNCIL

Information Governance
Aligning Information Stakeholders to Improve Information Economics

IBM

Home | Events | Resources | Blog | News | Supporters | Faculty | About CGOC | Join | Search

WEBCAST

Putting Information to Work: Key Insights from CGOC Summit 2014

June 5th | 1pm E

- Join over 2000 legal, records and IT practitioners from companies such as Bank of America, ExxonMobil, GE, Kaiser Permanente, Liberty Mutual, Novartis, Procter & Gamble and Toyota.
- The CGOC is a practitioner community focused on the intersection and challenges of ediscovery, retention, information governance and management.
- Its mission is to provide executives the opportunity to benchmark and exchange case studies; its practice groups focus on discreet areas in preservation, retention, and information governance to deliver work products that help our members best approach the challenges in maintaining best-in-class programs.
- Quarterly newsletters and monthly in-person and web-based meetings keep you informed and provide you the opportunity to engage with your peers.

Featured Resources

Disposing of Digital Debris, Information Governance Strategy and Practice in Action (EDRM/CGOC White Paper, 2014)

Information Lifecycle Governance Requirements Kit (CGOC Publication, 2014)

Information Lifecycle Governance Leader Reference Guide - Second Edition (CGOC Publication, 2014

Maximizing the Value of Information Technology (Report prepared by CFO Research with AlixPartners)

More »

Thank you

Isuru Fernando, Information Economics Advisor

fernando@nz.ibm.com

@wizzy