

SOCIAL MEDIA HEALTHCARE ANALYTICS

**DARWESH MANGAT – PRESALES HEALTHCARE
DARWESH.MANGAT@SAS.COM**



ANALYTICS

EVERYONE IS AN ANALYST

SMALL SILOS



- Small silos of analytics expertise
- Limited data needed
- Primarily desktop based

EVERYWHERE



- All People
- All Data
- Enterprise Wide

Lack of consistency and sharing
Need for governance
Duplication of effort

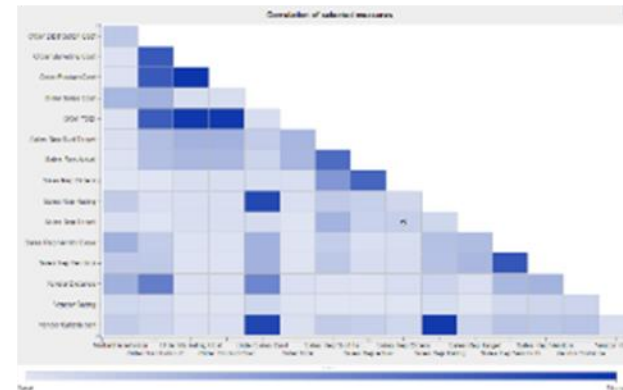
BI/VISUALIZATION CITIZEN DATA SCIENTIST

BI

EMPLOYEE	AGE	DEPARTMENT	EMPLOYEE	AGE	DEPARTMENT	EMPLOYEE	AGE	DEPARTMENT	EMPLOYEE	AGE	DEPARTMENT
EMP001	35	SALES	EMP002	42	MARKETING	EMP003	28	FINANCE	EMP004	55	OPERATIONS
EMP005	48	HR	EMP006	31	LEGAL	EMP007	62	COMPLIANCE	EMP008	22	RESEARCH
EMP009	51	IT	EMP010	38	CONSTRUCTION	EMP011	45	MANUFACTURING	EMP012	29	TRANSPORTATION
EMP013	33	ENERGY	EMP014	41	TELECOMMUNICATIONS	EMP015	58	UTILITIES	EMP016	25	WATER SUPPLY
EMP017	44	AGRICULTURE	EMP018	36	MINING	EMP019	65	CONSTRUCTION	EMP020	27	TRANSPORTATION
EMP021	39	MANUFACTURING	EMP022	47	CONSTRUCTION	EMP023	34	TRANSPORTATION	EMP024	52	WATER SUPPLY
EMP025	43	AGRICULTURE	EMP026	37	MINING	EMP027	61	CONSTRUCTION	EMP028	26	TRANSPORTATION
EMP029	46	MANUFACTURING	EMP030	40	CONSTRUCTION	EMP031	32	TRANSPORTATION	EMP032	54	WATER SUPPLY
EMP033	49	AGRICULTURE	EMP034	35	MINING	EMP035	63	CONSTRUCTION	EMP036	28	TRANSPORTATION
EMP037	42	MANUFACTURING	EMP038	44	CONSTRUCTION	EMP039	31	TRANSPORTATION	EMP040	56	WATER SUPPLY
EMP041	38	AGRICULTURE	EMP042	46	MINING	EMP043	64	CONSTRUCTION	EMP044	29	TRANSPORTATION
EMP045	41	MANUFACTURING	EMP046	39	CONSTRUCTION	EMP047	33	TRANSPORTATION	EMP048	53	WATER SUPPLY
EMP049	45	AGRICULTURE	EMP050	36	MINING	EMP051	62	CONSTRUCTION	EMP052	27	TRANSPORTATION
EMP053	48	MANUFACTURING	EMP054	43	CONSTRUCTION	EMP055	35	TRANSPORTATION	EMP056	51	WATER SUPPLY
EMP057	40	AGRICULTURE	EMP058	47	MINING	EMP059	60	CONSTRUCTION	EMP060	26	TRANSPORTATION
EMP061	44	MANUFACTURING	EMP062	38	CONSTRUCTION	EMP063	32	TRANSPORTATION	EMP064	54	WATER SUPPLY
EMP065	46	AGRICULTURE	EMP066	41	MINING	EMP067	61	CONSTRUCTION	EMP068	28	TRANSPORTATION
EMP069	42	MANUFACTURING	EMP070	45	CONSTRUCTION	EMP071	34	TRANSPORTATION	EMP072	52	WATER SUPPLY
EMP073	49	AGRICULTURE	EMP074	37	MINING	EMP075	63	CONSTRUCTION	EMP076	29	TRANSPORTATION
EMP077	43	MANUFACTURING	EMP078	44	CONSTRUCTION	EMP079	31	TRANSPORTATION	EMP080	56	WATER SUPPLY
EMP081	38	AGRICULTURE	EMP082	46	MINING	EMP083	64	CONSTRUCTION	EMP084	27	TRANSPORTATION
EMP085	41	MANUFACTURING	EMP086	39	CONSTRUCTION	EMP087	33	TRANSPORTATION	EMP088	53	WATER SUPPLY
EMP089	45	AGRICULTURE	EMP090	36	MINING	EMP091	62	CONSTRUCTION	EMP092	26	TRANSPORTATION
EMP093	48	MANUFACTURING	EMP094	43	CONSTRUCTION	EMP095	35	TRANSPORTATION	EMP096	51	WATER SUPPLY
EMP097	40	AGRICULTURE	EMP098	47	MINING	EMP099	60	CONSTRUCTION	EMP100	28	TRANSPORTATION
EMP101	44	MANUFACTURING	EMP102	38	CONSTRUCTION	EMP103	32	TRANSPORTATION	EMP104	54	WATER SUPPLY
EMP105	46	AGRICULTURE	EMP106	41	MINING	EMP107	61	CONSTRUCTION	EMP108	29	TRANSPORTATION
EMP109	42	MANUFACTURING	EMP110	45	CONSTRUCTION	EMP111	34	TRANSPORTATION	EMP112	52	WATER SUPPLY
EMP113	49	AGRICULTURE	EMP114	37	MINING	EMP115	63	CONSTRUCTION	EMP116	27	TRANSPORTATION
EMP117	43	MANUFACTURING	EMP118	44	CONSTRUCTION	EMP119	31	TRANSPORTATION	EMP120	56	WATER SUPPLY
EMP121	38	AGRICULTURE	EMP122	46	MINING	EMP123	64	CONSTRUCTION	EMP124	27	TRANSPORTATION
EMP125	41	MANUFACTURING	EMP126	39	CONSTRUCTION	EMP127	33	TRANSPORTATION	EMP128	53	WATER SUPPLY
EMP129	45	AGRICULTURE	EMP130	36	MINING	EMP131	62	CONSTRUCTION	EMP132	26	TRANSPORTATION
EMP133	48	MANUFACTURING	EMP134	43	CONSTRUCTION	EMP135	35	TRANSPORTATION	EMP136	51	WATER SUPPLY
EMP137	40	AGRICULTURE	EMP138	47	MINING	EMP139	60	CONSTRUCTION	EMP140	28	TRANSPORTATION
EMP141	44	MANUFACTURING	EMP142	38	CONSTRUCTION	EMP143	32	TRANSPORTATION	EMP144	54	WATER SUPPLY
EMP145	46	AGRICULTURE	EMP146	41	MINING	EMP147	61	CONSTRUCTION	EMP148	29	TRANSPORTATION
EMP149	42	MANUFACTURING	EMP150	45	CONSTRUCTION	EMP151	34	TRANSPORTATION	EMP152	52	WATER SUPPLY
EMP153	49	AGRICULTURE	EMP154	37	MINING	EMP155	63	CONSTRUCTION	EMP156	27	TRANSPORTATION
EMP157	43	MANUFACTURING	EMP158	44	CONSTRUCTION	EMP159	31	TRANSPORTATION	EMP160	56	WATER SUPPLY
EMP161	38	AGRICULTURE	EMP162	46	MINING	EMP163	64	CONSTRUCTION	EMP164	27	TRANSPORTATION
EMP165	41	MANUFACTURING	EMP166	39	CONSTRUCTION	EMP167	33	TRANSPORTATION	EMP168	53	WATER SUPPLY
EMP169	45	AGRICULTURE	EMP170	36	MINING	EMP171	62	CONSTRUCTION	EMP172	26	TRANSPORTATION
EMP173	48	MANUFACTURING	EMP174	43	CONSTRUCTION	EMP175	35	TRANSPORTATION	EMP176	51	WATER SUPPLY
EMP177	40	AGRICULTURE	EMP178	47	MINING	EMP179	60	CONSTRUCTION	EMP180	28	TRANSPORTATION
EMP181	44	MANUFACTURING	EMP182	38	CONSTRUCTION	EMP183	32	TRANSPORTATION	EMP184	54	WATER SUPPLY
EMP185	46	AGRICULTURE	EMP186	41	MINING	EMP187	61	CONSTRUCTION	EMP188	29	TRANSPORTATION
EMP189	42	MANUFACTURING	EMP190	45	CONSTRUCTION	EMP191	34	TRANSPORTATION	EMP192	52	WATER SUPPLY
EMP193	49	AGRICULTURE	EMP194	37	MINING	EMP195	63	CONSTRUCTION	EMP196	27	TRANSPORTATION
EMP197	43	MANUFACTURING	EMP198	44	CONSTRUCTION	EMP199	31	TRANSPORTATION	EMP200	56	WATER SUPPLY

- IT centric
- Architecture driven
- Static and slow

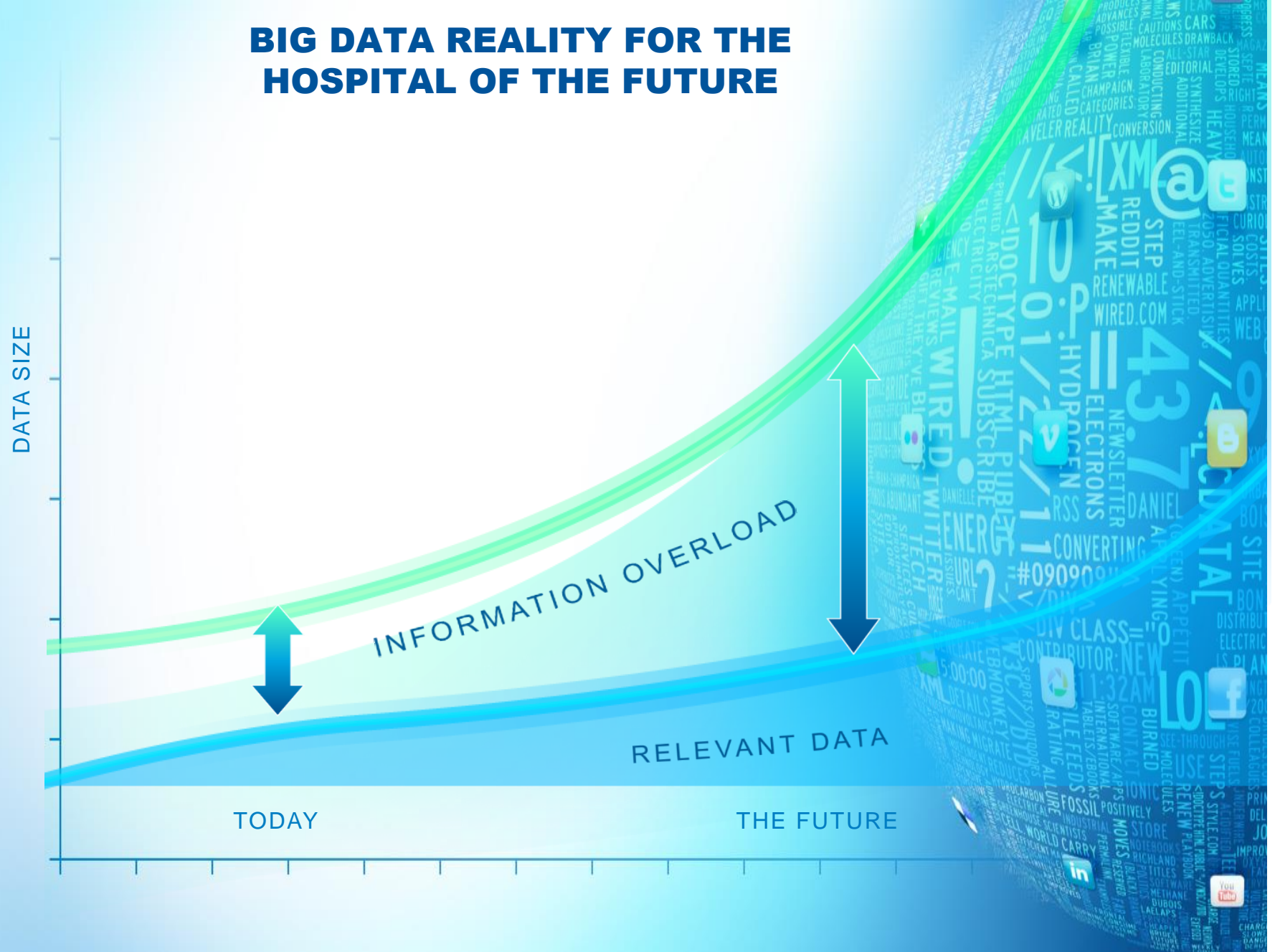
Data Visualization



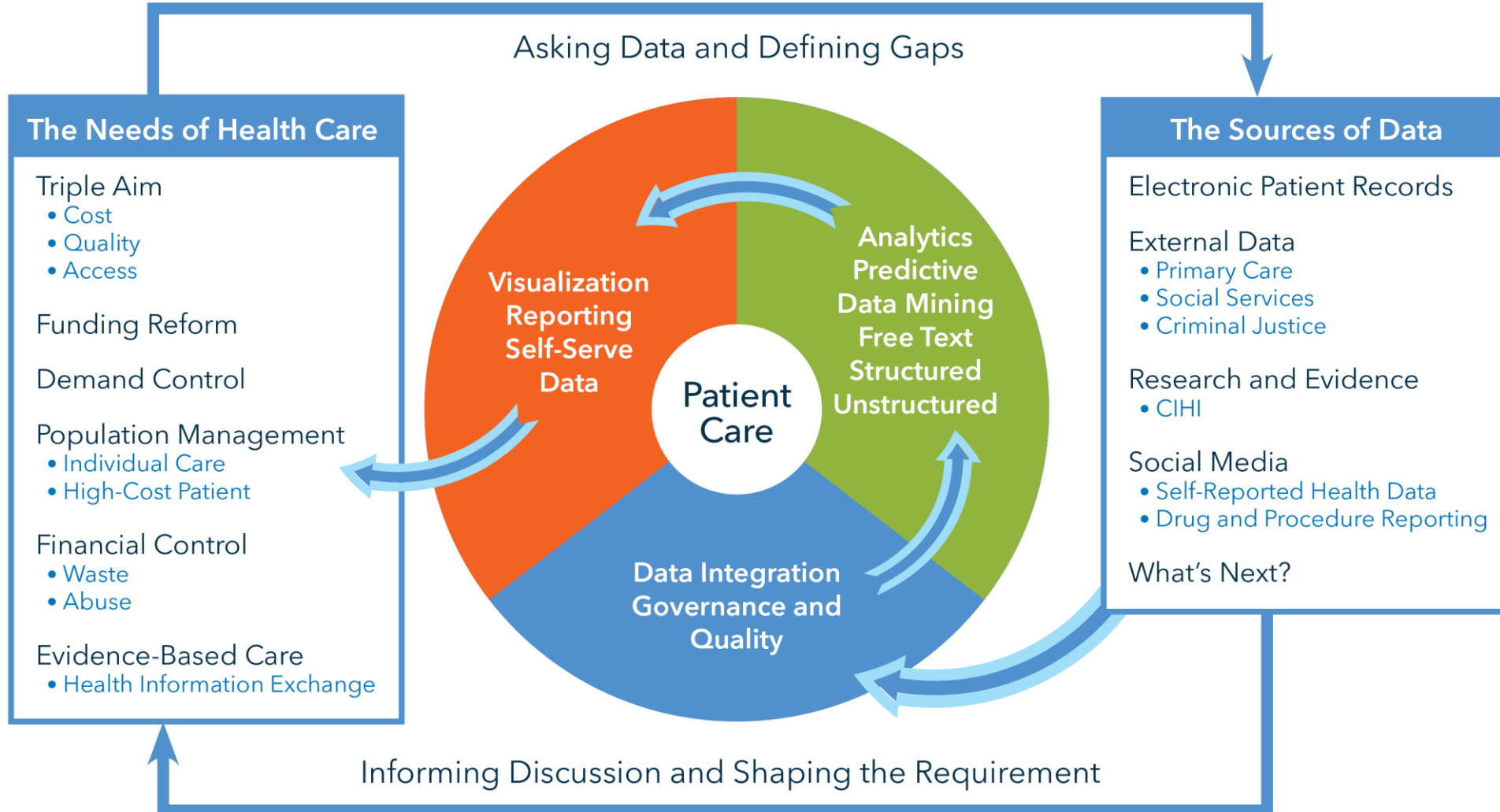
- User centric
- Exploration
- Dynamic

BIG DATA REALITY FOR THE HOSPITAL OF THE FUTURE

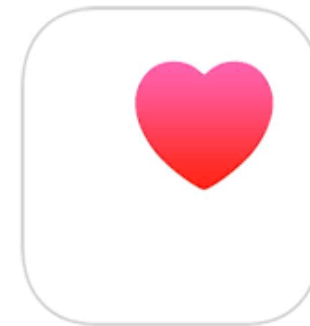
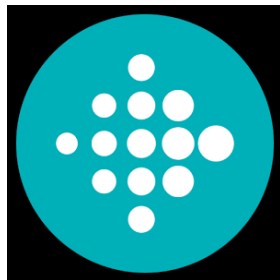
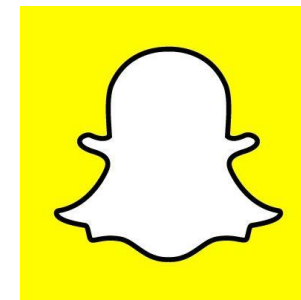
- EMR & HIS
- Unstructured Text & Voice
- Device Generated
- Finance
- Operational
- Environmental
- Social Media
- Research
- Staffing



The Who, What, Where and Why of Health Analytics



ANALYTICS DATA SOURCES



COLLABORATE ENABLING ANALYTICS ACROSS HEALTH CARE STAKEHOLDERS



MOBILE



DESKTOP APPLICATIONS



COLLABORATION APPLICATIONS



WEB

VISUALIZATIONS

REPORTS

DASHBOARDS

COMMENTS

INSIGHTS

- What proportion of Canadian youth (13-17) post about their mental health, and describe experiencing bullying or suicidal thoughts in the past 12 months on social media?

SOME RESULTS

- 1.1M tweets by teens (13-17 years old) in English language
- In Canada between Jan 6 – Jan 13, 2016
- Topics
 - **Feelings**: 106k \approx 10%
 - **Bullying**: 11k \approx 1%
 - **Suicide**: 382 \approx 0.035% of tweets
- 39% correlation between **suicide** and **bullying**
- 12.7% of authors have tweeted at least twice about **suicide**

PROCESS

1. Download tweets made in Canada within a specific time frame
2. Detect language and filter English (SAS Contextual Analysis)
3. Score tweets by age and filter 13-17 year olds (SAS Enterprise Miner)
4. Tag tweets by topic (family, bullying, school, suicidal thoughts, feelings, ...) (SAS Contextual Analysis)
5. Visualization (SAS Visual Analytics)
 - Relationship between topics
 - Contextual analysis of specific topics

can someone with my number please text me because i have nobody to talk to and i just want to hurt myself so badly and i'm scared

my own father told me I deserved being killed. I hate him

there's not one day I don't think of killing myself

Tweeted about depression

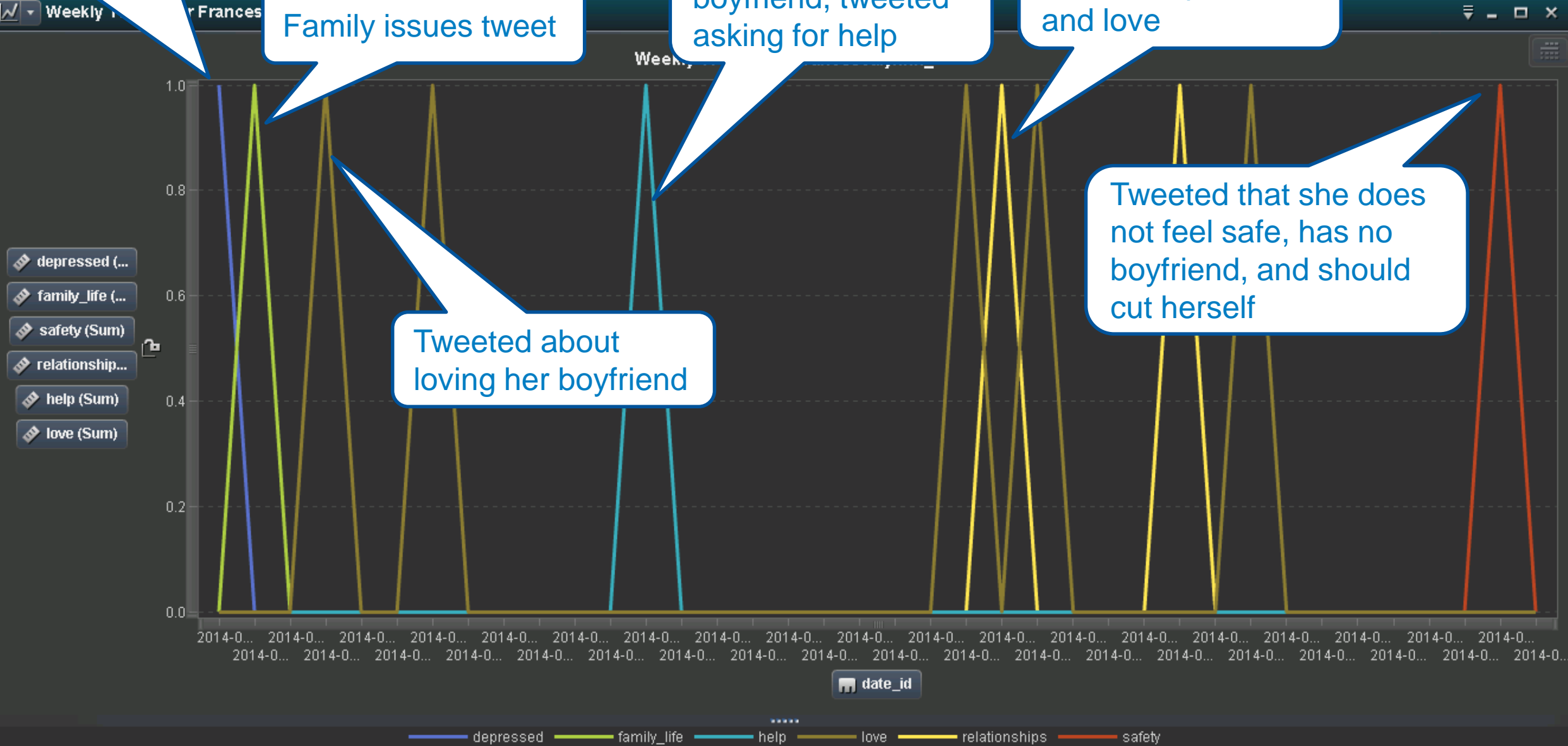
Family issues tweet

Had a fight with her boyfriend, tweeted asking for help

Tweeted about relationship trouble and love

Tweeted about loving her boyfriend

Tweeted that she does not feel safe, has no boyfriend, and should cut herself



- depressed (...)
- family_life (...)
- safety (Sum)
- relationship...
- help (Sum)
- love (Sum)

date_id

depressed family_life help love relationships safety



Visualization 4

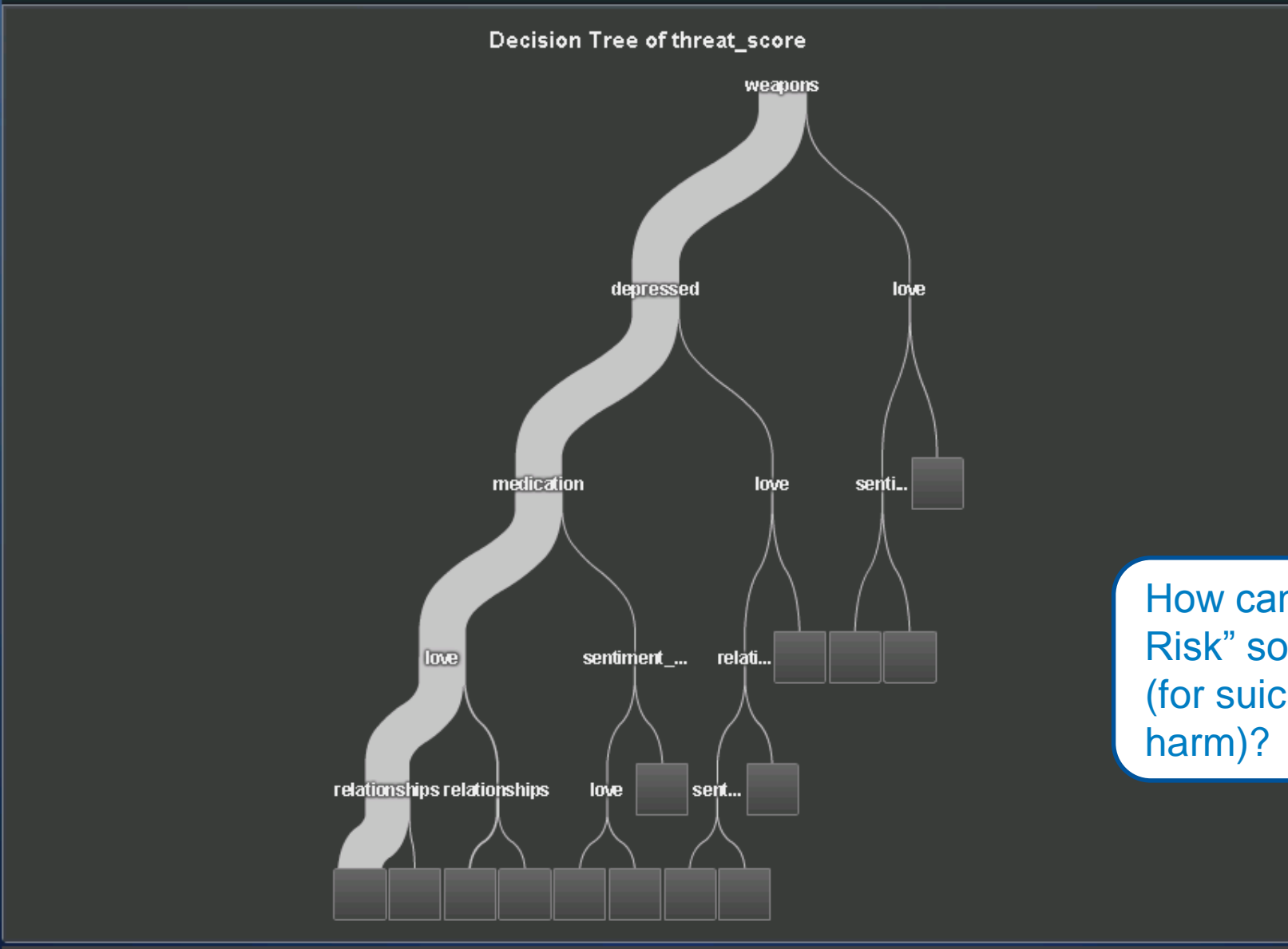
HAPPY2

Search data

Category	Count
author	4,731
date_id	174
datetime	174
hashtags	201
location	2,371
sentiment	3
source	1
text	5,000+
tweeted.at	1,350
weekday	7

Measure
alcohol
day
depressed
family_life
followers
friends
help

Property	Value
Name	
Role	
Model type	
Format	
Aggregation	



Roles * Filter... Ra... Pr... Co... ▾

Decision Tree

Use Automatic Chart

Target

threat_score ▾

Predictors

- weapons
- sentiment_score
- relationships
- medication
- love
- depressed

How can we predict "At-Risk" social media users (for suicide and self-harm)?

DATA SOURCES & METHODOLOGY

Web-Based Media Sources

- Forums
- Chat Rooms
- Face Book
- Twitter
- Documents/Narrative

Data Aggregation and Analysis Methods

- Web/Document Crawling
- Bag of Words Approach
 - Statistical/Mathematical
- Natural Language Processing
 - Dictionary/Taxonomy/Ontology
- Sentiment Analysis
- Predictive Model Building

FURTHER METHODOLOGY

Clustering

- Let maths decide topics
- Best for exploring data
- Manual refinement and topic labeling

Classification

- Humans decide topics of interest
- Automatic refinement using algorithms
- Best if you have a starting point on what to look for

Get more data

- Partnering with Twitter/Facebook/Snapchat
- Partnering with Demographical institutions/police reports
- Building predictive models for analyzing risk

SOLUTION FRAMEWORK



HOW DO WE MANAGE DATA IN THE PHYSICAL WORLD?

1. Acquire



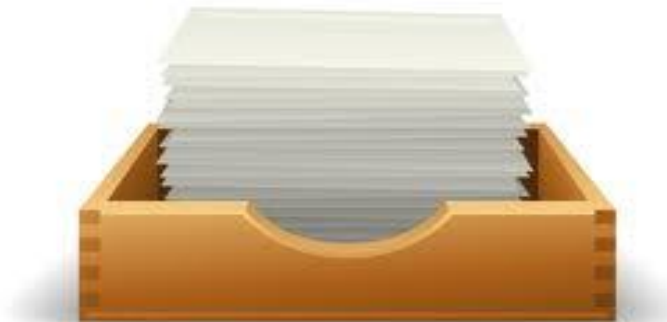
2. Determine Relevance



3. Store



Trash



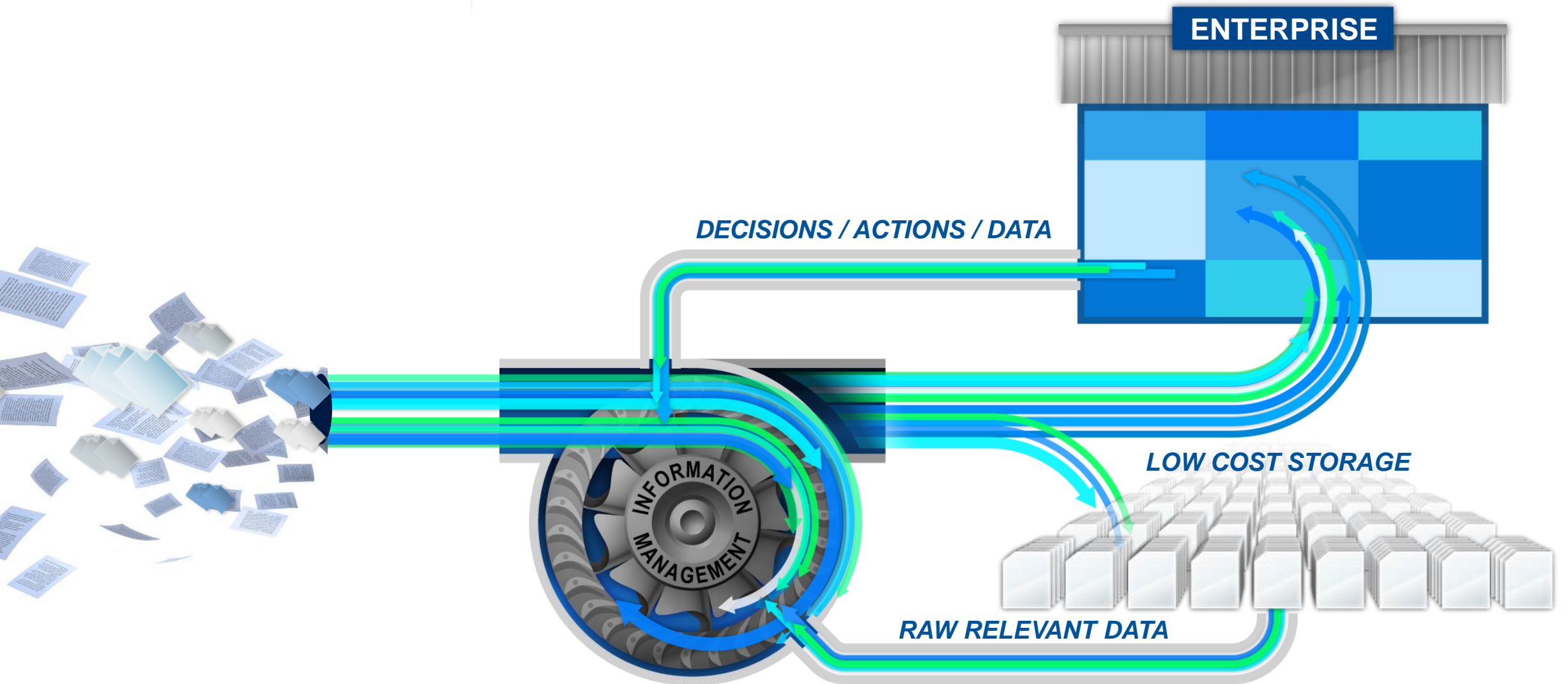
Cache



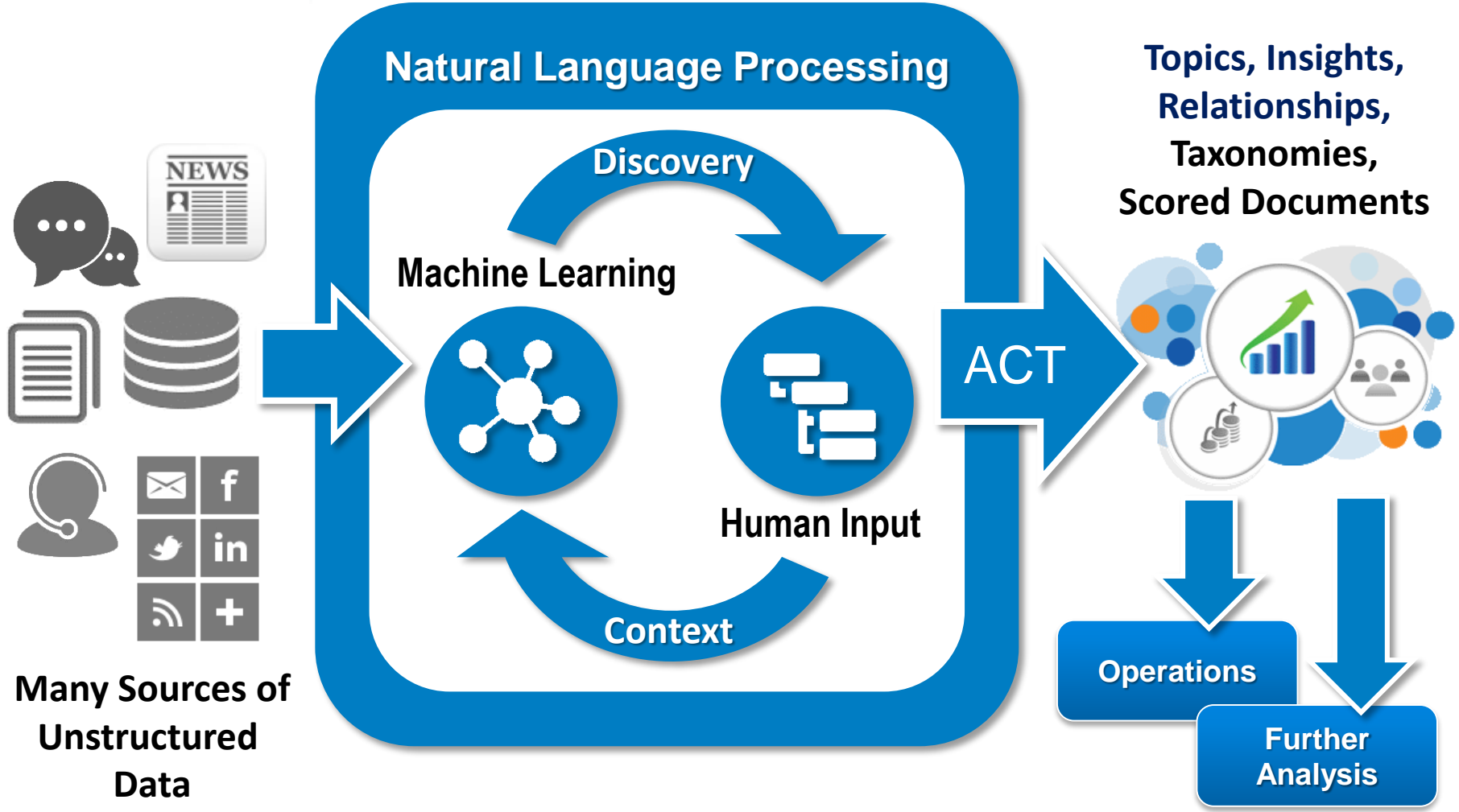
Storage

How to ingest TONS of DATA

Stream It, Score It, Store It



The SAS Approach



Questions



**THE
POWER
TO KNOW®**

SAS.com