



PLATFORM
ReVolution



CANADIAN
INSURANCE
CONFERENCE

Alexis Iglauer

Head of Analytics & Technology (L&H)

PartnerRe

Big Data, Advanced Analytics, and Data Privacy

Disclaimer

The following presentation is for general information, education and discussion purposes only, in connection with the Canadian Reinsurance Conference 2019. Any views or opinions expressed are those of the presenters alone. They do not constitute legal or professional advice; and do not necessarily reflect, in whole or in part, any corporate position, opinion or view of PartnerRe or its affiliates, or a corporate endorsement, position or preference with respect to any issue or area covered in the presentation.

GDPR - Data Subject Rights

Know

Access

Rectify

Withdraw

Object

Object to
automated
processing

Be Forgotten

Portability

It's not “just in Europe”

sg FEAT principles

- Fairness, Ethics, Accountability, Transparency

us New York Commissioner's letter

us California Consumer Privacy Act of 2018

ca PIPEDA fair information principles

-

Are you going to keep that?

- “Big data is what happened when the cost of storing information became less than the cost of making the decision to throw it away.” – *George Dyson, 2013*
- Data, the new asbestos?



PLATFORM
ReVolution



CANADIAN
INSURANCE
CONFERENCE

Puneet Bakshi

Vice President and Chief Transformation Officer

Munich RE 

Structuring a data science team's setup, processes and tools to be naturally privacy-friendly

Agenda

- The Privacy Landscape
- The Privacy Challenge
- Data Privacy – Effect on the Insurance Industry
- Data Privacy vs Data Science
- Planning for the future of Privacy
- Key Takeaways

The Privacy Landscape

- Privacy laws are intended to protect the individuals right to privacy while at the same time allowing for collection and use of personal information.
- This is a delicate balance to be maintained but legislation always favours the individuals right to privacy
- The laws are complex and vary depending on region
- Penalties can significantly impact a companies bottom line.

The Privacy Landscape

Although the laws vary by region, they generally agree on the following:

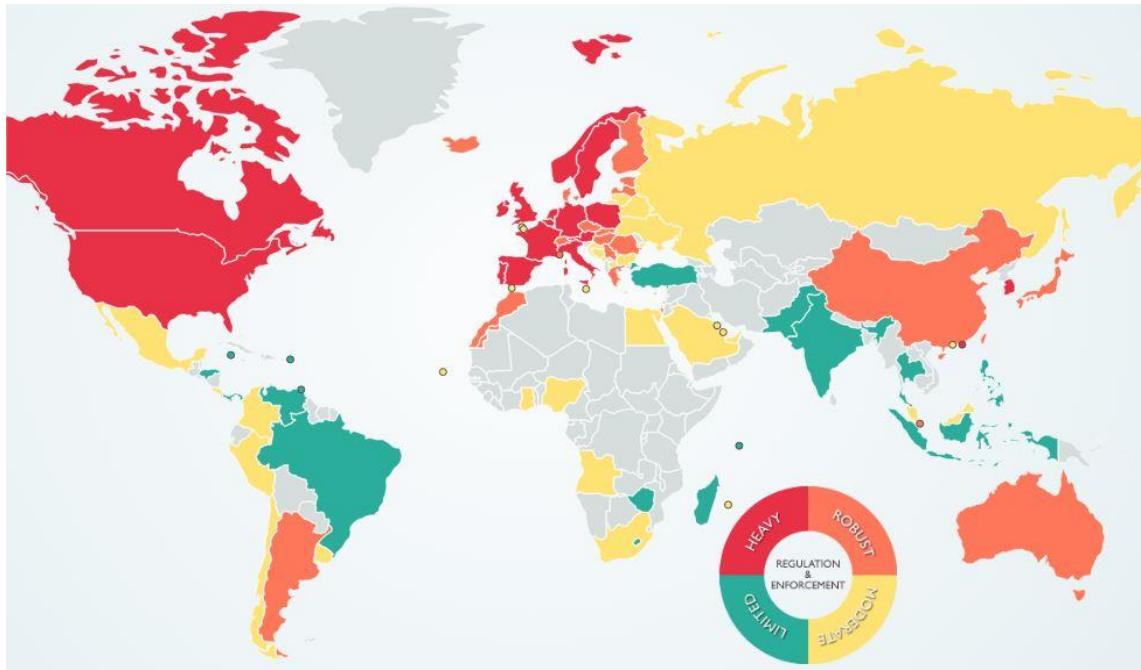
- The individual must be informed and agree that their data will be collected
- There is a legal requirement to process the data
- Data processing is required for vital reasons
- Data processing is necessary for the public interest

And.....

- Data regarding race or ethnicity, sexual orientation, political views, religious or health **can only be processed in special cases** (e.g. if you have expressly agreed to it or if the processing of this data is necessary due to national laws)

The Privacy Challenge

Global insurers and reinsurers have to deal with a growing number of data protection and privacy laws and around the world



Source: www.dlapiperdataprotection.com



Personal Information Protection and Electronic Documents Act



The Health Insurance Portability and Accountability Act



General Data Protection Regulation

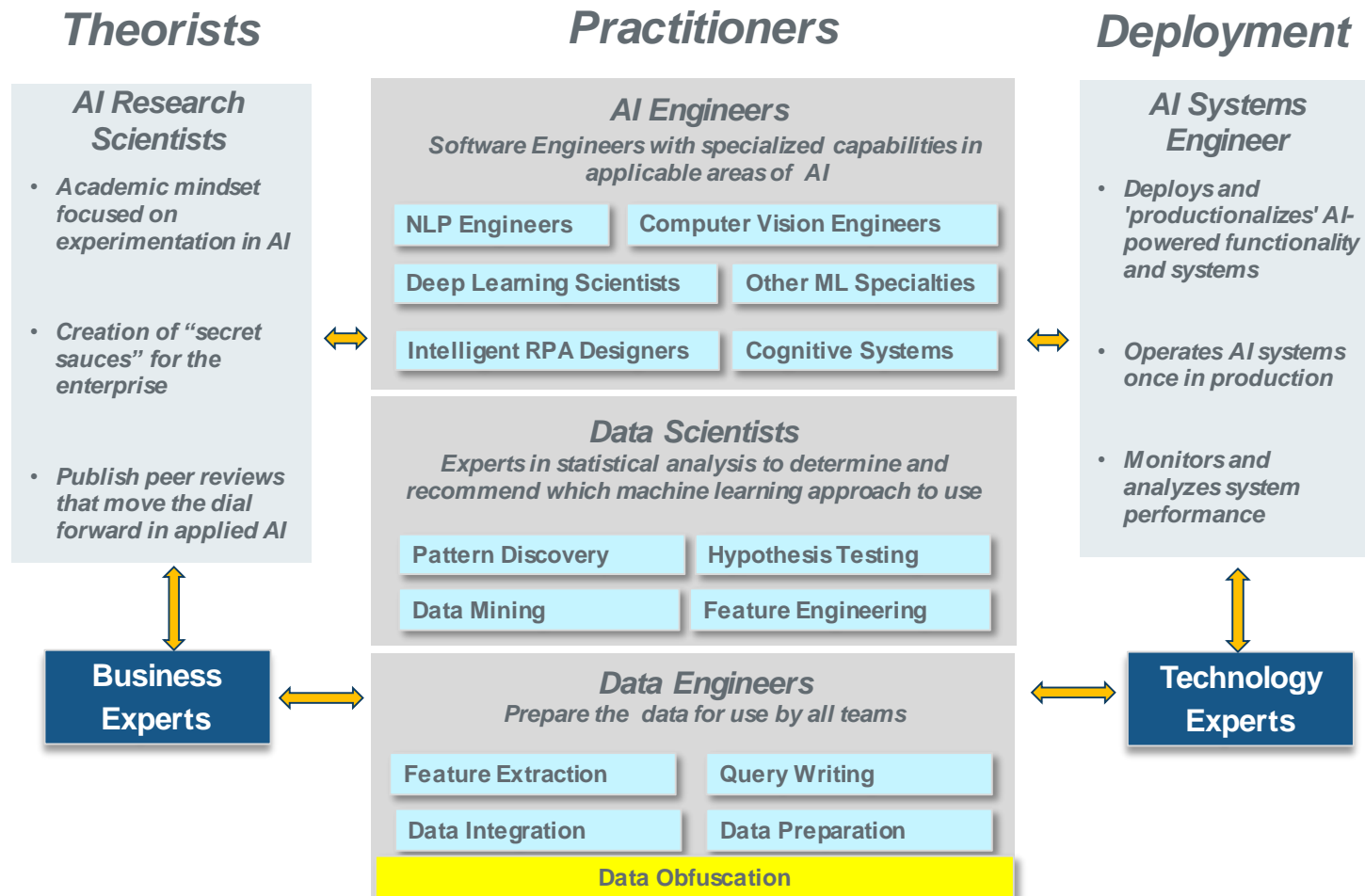
Data Privacy - Effect on the Insurance Industry

- **Diluted Data:** Insurers are opting for anonymizing data that is aggregated and used for analytics
- **Minimized New Data Capture:** The industry has been slow to collect new data for furthering analytic and AI capabilities
- **Diminished Underwriting Quality:** Due to anonymized data as well as limited new data, the development of new underwriting criteria is slowed.
- **Slowdown in New Product Development:** New products require new data sources for projection, Valuation and Risk assessments
- **New Focus on Data Protection:** Significant capital expenditure on the protection of data to meet Legislative standards. (e.g. GDPR 72 Hour Breach Reporting)
- **Claimants using Privacy legislation** as a tool to access information held by an insurer prior to a claim being litigated

Data Privacy vs Data Science

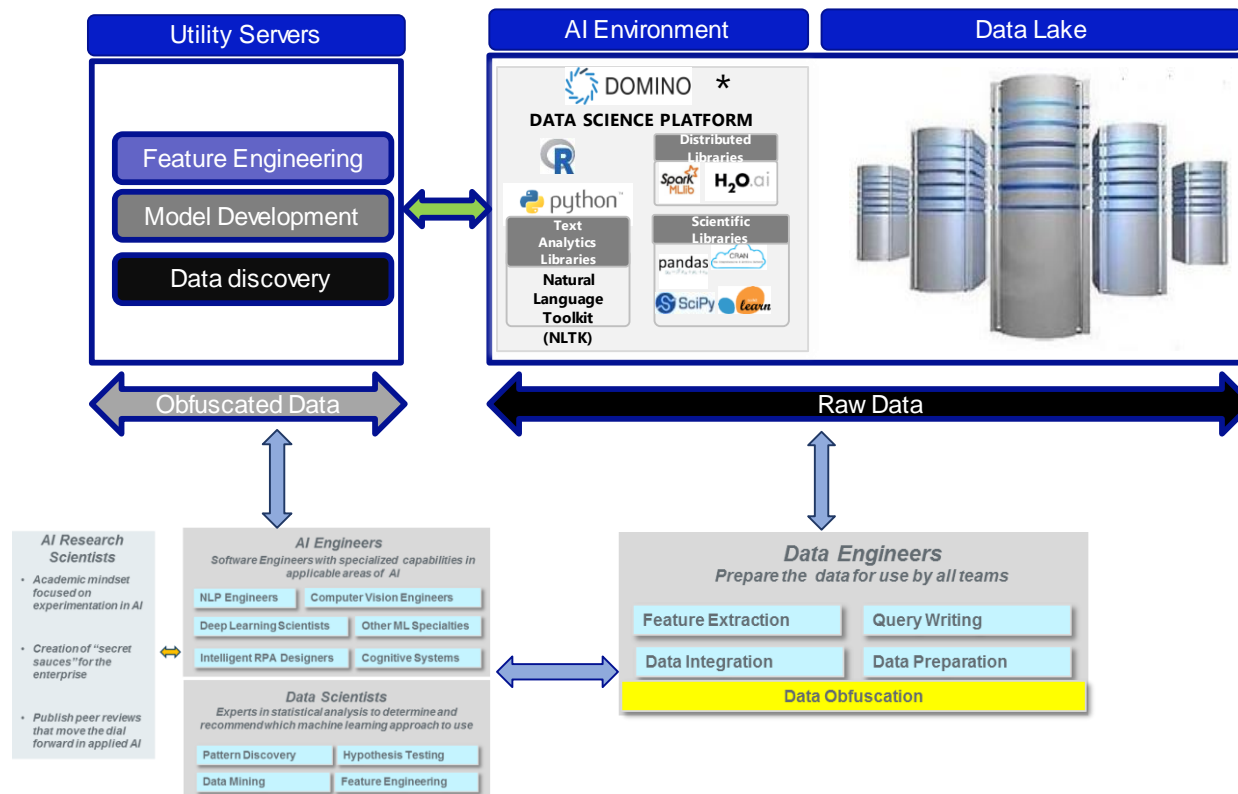
- The data science practice requires a consistent and reliable source of existing and new data to be effective.
- Access to data must be carefully considered when building a data science team.
- Typically not all team members need access to the same data and there are opportunities to segregate the access appropriately
- Data Science Infrastructure must also support the data protection and privacy policies and guidelines

Data Privacy vs Data Science



- A Data Science Team consists of several roles
- **Data Engineers and Business Experts** can work with unobfuscated data
- Data Engineers can carry out the data obfuscation in support of all the other functions within the data science team
- For data privacy concerns all other roles can work with obfuscated data

Data Privacy vs Data Science



- The data within the infrastructure should also be separated to ensure privacy policies are met
- Strict access controls and auditing to production environments are necessary
- External data is usually obfuscated but for data exchanged between carriers and re-insurers the data obfuscation step may be necessary before handing off to the data science team

Planning for the Future of Data Privacy

To meet the challenges imposed by the privacy legislation, Insurance companies and financial institutions are looking at novel ways to satisfy their reliance on data.

- **Purchasing Anonymized Data** – Using anonymized data from external sources to augment the data they collect for underwriting and claims
- **Logically Separating Data** – Building Logical data warehouses to control access by individuals and teams
- **Publishing Clear Enterprise Data Privacy Guidelines** and principles and making these available to customers and soliciting consent for use of data

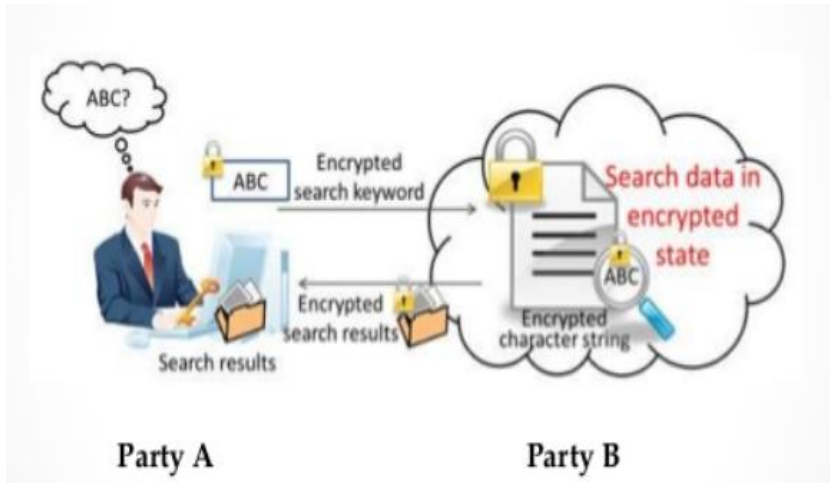
New Uses of Available Data

The variety and richness of available data supports leading edge capabilities within the insurance industry both now and in the future

- **Homomorphic Encryption**
- **IOT sensory Information for Risk Evaluation**
- **Cell Phone Data for accident detection**
- **Automated repair estimation for cars**

New Uses of Data

Homomorphic Encryption – Used to building models that operate on encrypted data as opposed to raw data, coupled with ZK-SNARK for exchange of information



- Zero-Knowledge Succinct Non-Interactive Argument of Knowledge, (ZK-SNARK) allows a sender and receiver to exchange and process information and return a response without ever revealing the content of the information being exchanged
- Consumers can apply for insurance to an automated underwriting system without revealing their actual information to the insurance company.

New Uses of Data

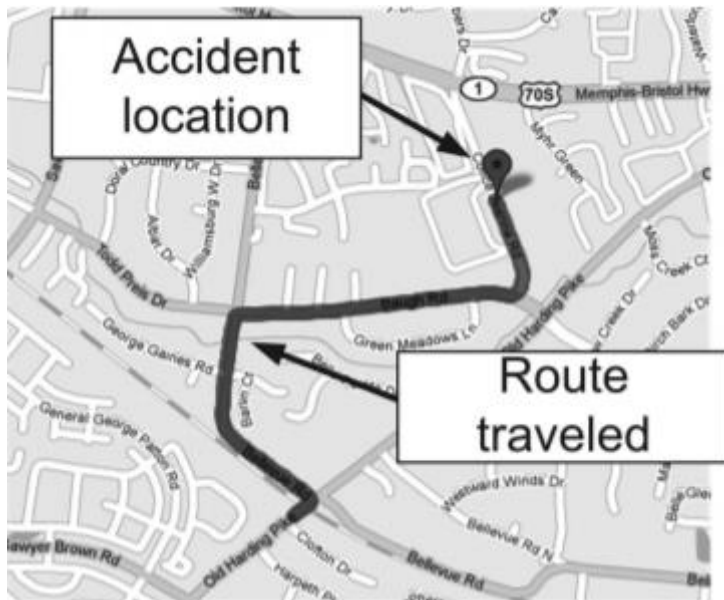
IOT sensory Information for Risk Evaluation – Smart watches can now register EKG and other fitness data that can be sent to insurance companies with the consent of the consumer



With EKG data combined with other data, AI algorithms can calculate risk on an individual level for life and health policies and recommend premium discounts for consumers

New Uses of Data

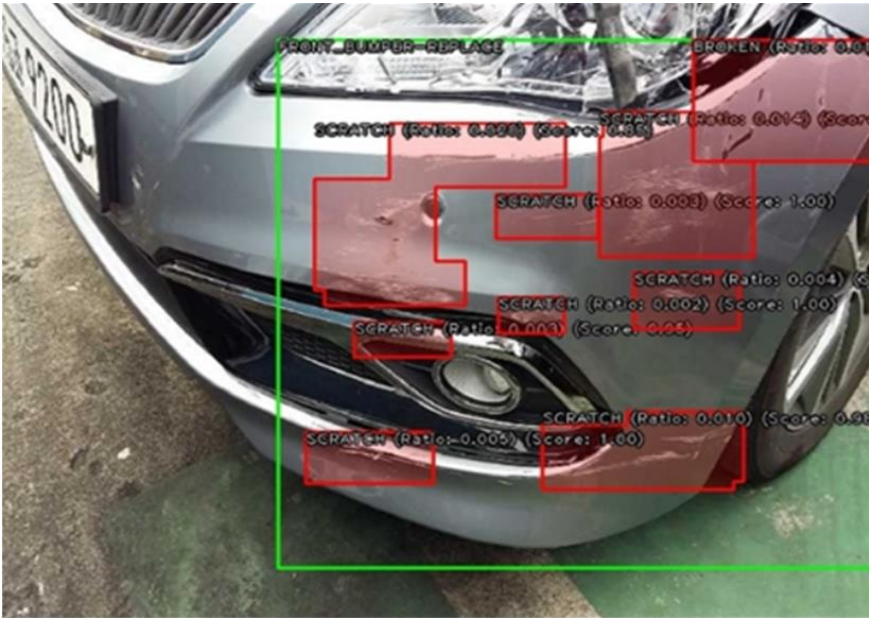
Cell Phone Data for Accident Detection – With customer consent, cellphones can be used to detect accidents and record driving patterns.



- With accident detection, a claim notification can be sent directly to the insurance company and specific instructions on claim submission can be sent to the consumer.
- Allows the insurance company to proactively call the consumer at the time of the accident and can also assist with emergency dispatch.

New Uses of Data

Automated Repair Estimation – With customer consent, cellphones can be used to detect accidents and record driving patterns.



Source of image: Korea Insurance Development Institute

- Insurance companies will use images from accident scenes to automatically estimate the cost of repairs to an automobile
- Allows insurance companies to instantly process claims, provide automated quotes and route the damaged car to the best auto repair provider.

Key Takeaways

- The variety of privacy laws around the world can make it difficult for a multinational insurance company to develop an effective data science practice
- Segregation of data through access controls, processes and infrastructure plays a key part in remaining compliant to privacy laws
- Companies have leveraged different technologies to help in overcoming the challenges of privacy laws (e.g. Homomorphic Encryption, Data Obfuscation)
- Legislation has forced the industry to develop clear privacy and data lifecycle management policies that govern the collection, use, protection and destruction of data
- New uses of data that is currently available will force insurers to strengthen internal data governance and privacy procedures if they wish to effectively leverage the variety of data available from consumers.



PLATFORM
ReVolution



CANADIAN
INSURANCE
CONFERENCE

Thomas D. Fletcher

VP Data Analytics, North America

PartnerRe

Big Data, Advanced Analytics, and Data Privacy
The Pros and Cons of Various Privacy Methods

The following presentation is for general information, education and discussion purposes only, in connection with the Canadian Reinsurance Conference. Any views or opinions expressed are those of the presenters alone. They do not constitute legal or professional advice; and do not necessarily reflect, in whole or in part, any corporate position, opinion or view of PartnerRe or its affiliates, or a corporate endorsement, position or preference with respect to any issue or area covered in the presentation.

Analyzing data is beneficial ...

- Proliferation of data makes joining data for analytic purposes appealing.
- Linking internal & external sources can yield many analytic opportunities

For decision-making, leads to:

- Consistency
- Efficiencies
- Opportunities



Improved outcomes

- Business –
 - Reduced costs
 - Increased opportunities
- Customers –
 - Relevant and efficient customer experience
 - Minimized hassle

Yet, risks can be great if data are misused

Reputational Risk

- Many campaigns against companies following breaches
- Unsavory reputation leads to undesirable outcomes

Business Loss

- Customers and corporations often reduce business with companies involved in privacy breaches

Investor Confidence

- Damaging reputations often yield retreat in stock
- Can be difficult to recover or seek new funding

Class Action Suits

- Customer disappointment goes beyond business loss
- Discrimination can and often will lead to suits

Regulatory Fines

- Many high profile cases of hearings before congress
- Fines can be severe

THE WALL STREET JOURNAL.



Forbes

The Washington Post
Democracy Dies in Darkness

Marriott faces massive data breach expenses even with cybersecurity insurance

Equifax Inc.

NYSE: EFX



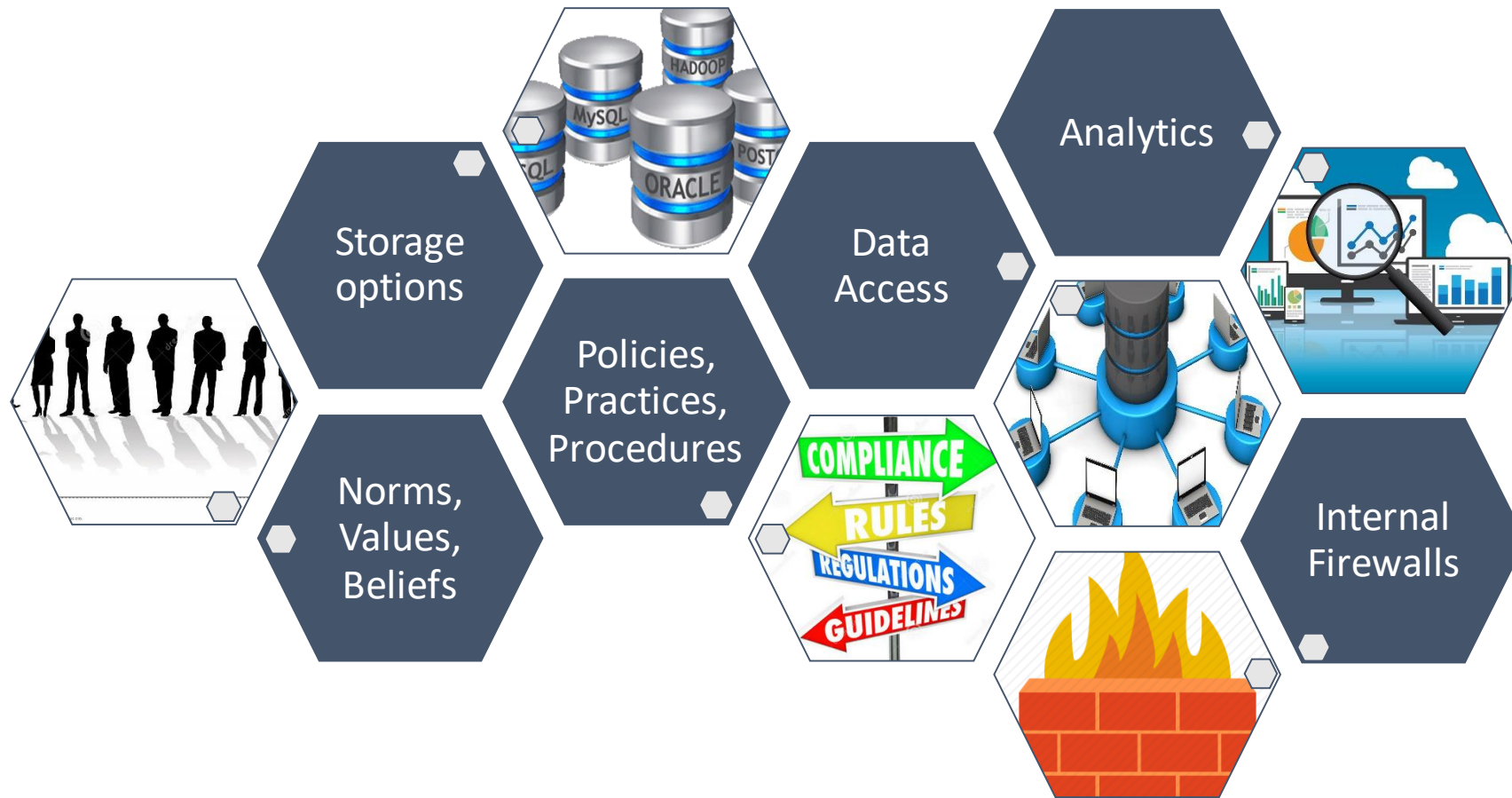
UCLA Health Settles Class Action Data Breach Lawsuit for \$7.5 Million

Facebook Faces Potential \$1.63 Billion Fine in Europe Over Data Breach

Uber will pay \$148 million in connection with a 2016 data breach and cover-up

Facebook's Zuckerberg just survived 10 hours of questioning by Congress

Ultimately, to be ethical, privacy-friendly, etc.
... some things need to change



While we have to leverage data in a privacy-friendly way, that's not always easy

Aggregation of Census

- Counts, averages, and the like are often reported to mask census information
- Different reporting periods coupled with thin slices of data can result in identification through deduction

Deletion often fails

- Data persists. Copies often exist
- AOL deleted data once NYT id'd a person; same data continue to exist on the internet (650k users search history)

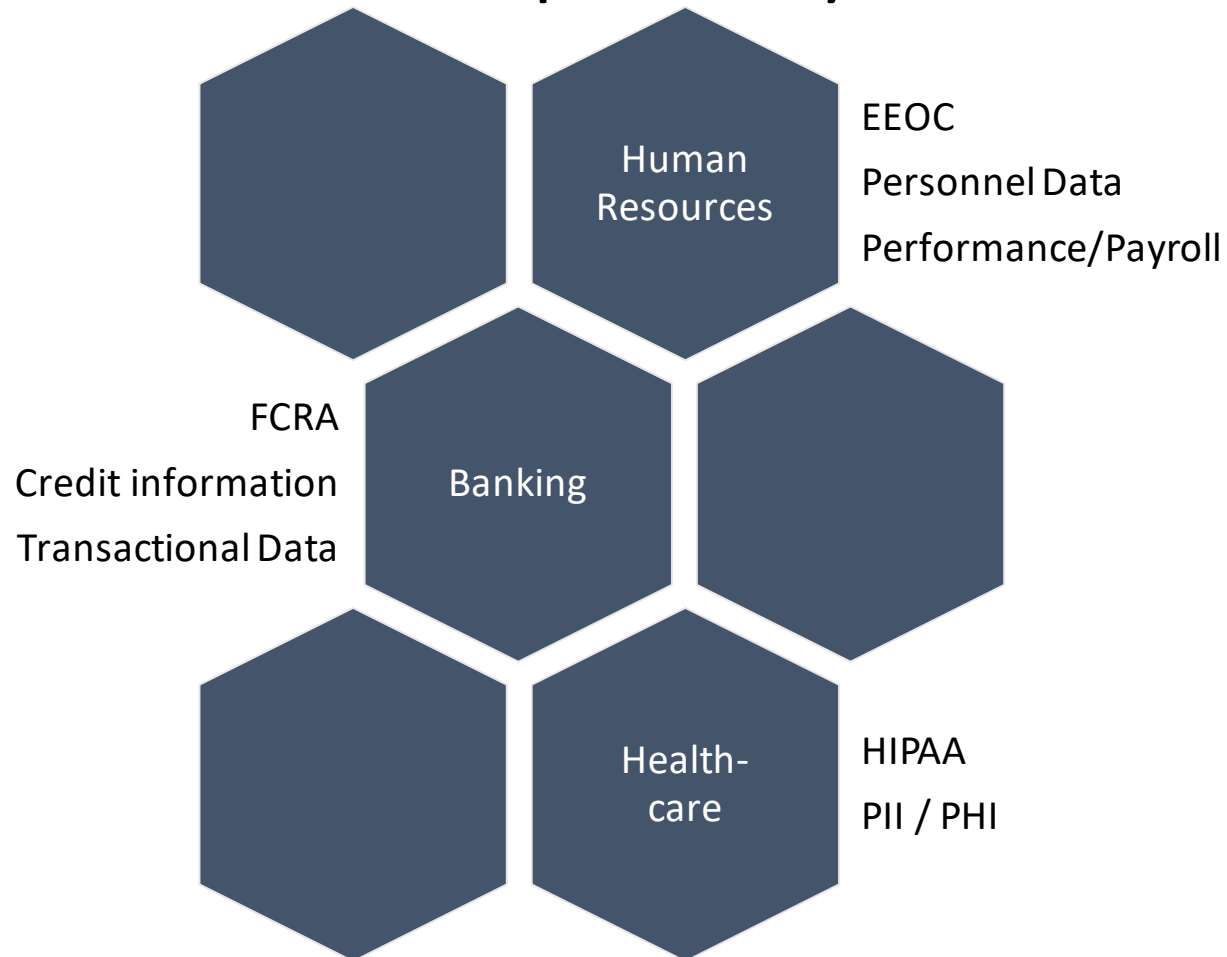
Hello Gov. – Your records

- Research is necessary, and data are required for research, but ...
- MIT student (now professor) identified the Gov and sent him his personal medical records from a research database that he assured the public were anonymous. - Dept of Homeland Security testimony

Only 3 Fields Needed

- It is estimated that 87% of US can be re-identified with Gender, DOB, Postal code via public data
- Bank transactional data can also render unique rows with a limited set of fields.

We are not alone: Other disciplines have tackled privacy issues with varying success



- There has been an evolution of techniques to address privacy dating back hundreds of years (e.g., census data). Four categories:
 - Aggregation and statistical synthesis
 - Deletion of data (or not collecting)
 - Internal firewalls, policies, procedures
 - Masking, encryption and differential privacy
- No one method ensures privacy friendly in all situations – nor do they all allow for same interpretation, analysis, precision, etc

Aggregation and statistical synthesis

Description: As a direct means to mask individual information, statistical techniques such as reporting aggregate counts by factors such as geography, demographic groups and the like. Among the oldest and rudimentary mechanisms to protect privacy

Pros:

- Individuals masked due to counting
- Simple in design (frequencies)
- Can be applied across many situations (but, ... care not to over slice data)

Cons:

- Information loss
- Requires larger data to make meaningful summaries
- With thin data, either can't report certain groups, or risk re-identifying
- Limiting further analysis

Example:

Suppose a census database exists with all relevant data. To protect against identification of individuals, repeated queries are required (e.g., counts by demographic groups). With enough queries, the data are no longer anonymized

Don't ask, don't know – or delete after use

Description: Frequently, it is recommended to not collect data – then there is limited or no risk of leakage. Similarly, if the data is created then have practices to delete (either after an analysis – or without any such analysis)

Pros:

- *In theory*, non existence protects against such leakage
- Effective if the data are not relevant or required for business use

Cons:

- Limiting if such data are required at some point (e.g., discrimination)
- Deletion is often precarious at best (copies can persist)

Example:

Data are collected to predict human potential on the job. EEOC and the courts require demonstration that the *test* is valid and does not create adverse impact. To conduct adverse impact analysis, demographics of protected class had to be collected (e.g., gender, race, age)

Internal firewalls, policies, procedures

Description: In addition to statistical techniques to ensure privacy, cultural, rules oriented, and mechanical techniques can exist. Storing data in one location with limited access. Ensuring practices exist to minimize access as well as create a value for protecting data

Pros:

- Enhances ability to comply with regulations and expectations
- Limits accidental leakage
- If strong culture emerges, most everyone values protection of data

Cons:

- Can be difficult to develop and enforce
- May provide challenges for some analyses (data is not as freely accessible)

Example:

A vendor supplies data for *scores* related to all individuals in a book of business. Analysts in the organization do not need to know which scores belong to which individuals. Mechanisms can exist to store the anonymized data on separate servers with limited access to servers containing the PII data

Masking, encryption and differential privacy

Description: This grouping of techniques are related but different. Masking is simply a form of recoding data to a format that isn't recognizable without a key. Encryption involves a more technical form of masking. Differential privacy is the result of analysis whereby random noise is injected

Pros:

- PII is limited, masked or removed
- Data cannot be recovered unless unencrypted
- Use of privacy loss parameter guarantees privacy via aggregate reporting in an opt out manner

Cons:

- Requires an extra step, or thought
- Tools are emerging to facilitate (especially for differential privacy)
- Specificity could be lost in some scenarios due to privacy parameter
- Tools not developed for individual data

Example:

Suppose a census database exists with all relevant data. To protect against identification of individuals, repeated queries are required (e.g., counts by demographic groups). If differential privacy is leveraged, one cannot re-anonymize the data due to random noise in the data

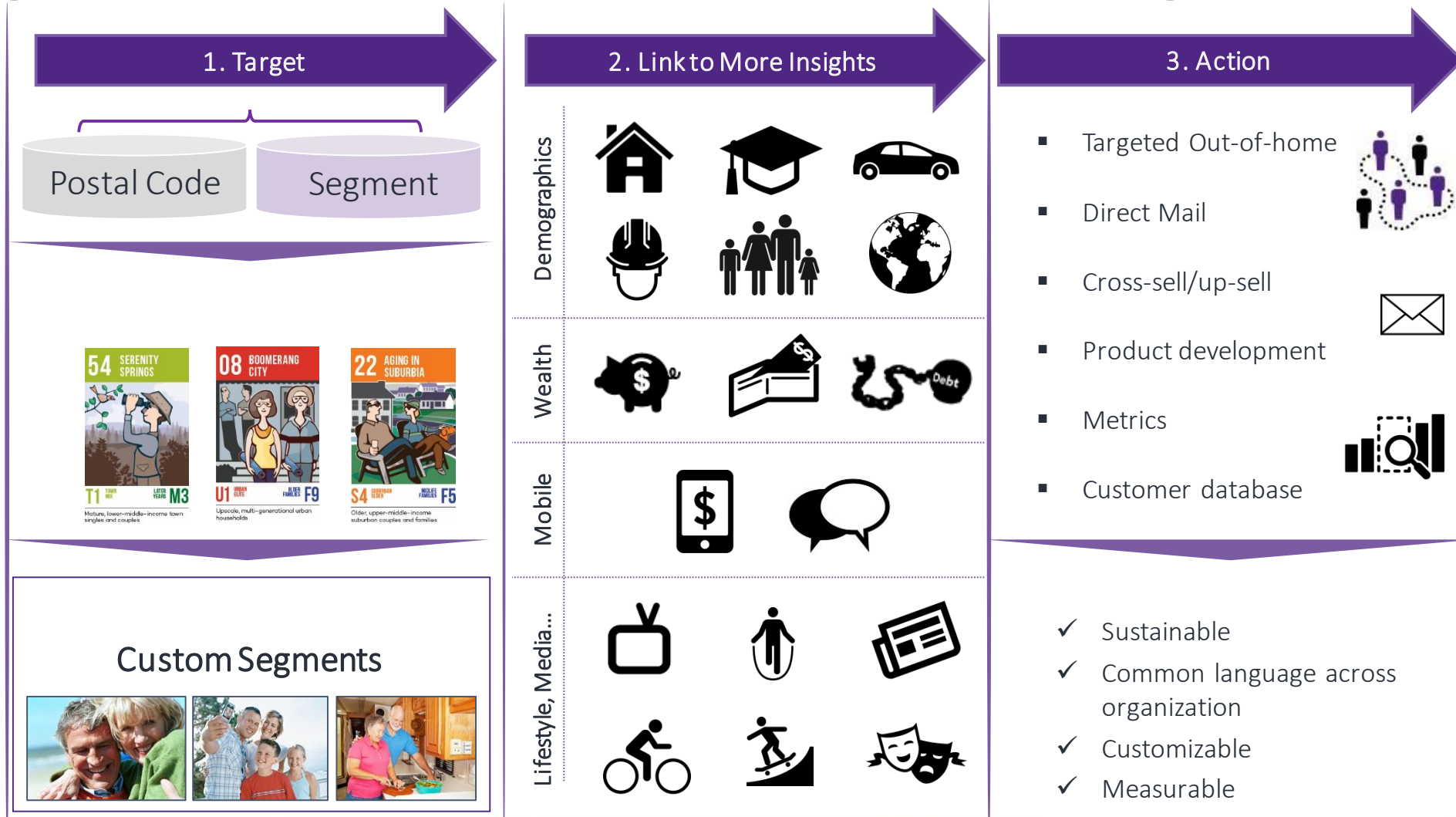
Leveraging Privacy Friendly Third Party Data
for Insurance

Environics Analytics






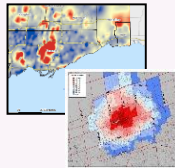





Stan Ivankovic
Director, Business Development, Insurance

Align Products, Services and Messages

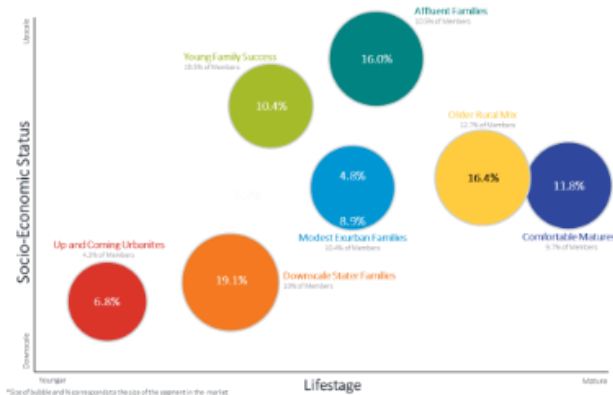


Analytics in Insurance

Goals	INCREASE RELEVANCY	BECOME MORE EFFICIENT	IMPROVE PREFORMANCE	ACCELERATE ONBOARDING	OPTIMIZE CHANNEL STRATEGY	NETWORK OPTIMIZATION	INCREASE ENGAGEMENT	REDUCE ATTRITION	IMPROVE COMMERCIAL INITIATIVES
Tools	Customer Segments & Personas 	Economic Segments 	Gap Analysis 	Onboarding Analysis 	Channel Analysis 	Site Modeling 	Next Best Product 	At-Risk Model 	B2B Analysis 
Actions	Define your segments and identify hidden opportunities	Match services levels to member needs and value	Identify and target the gaps in product penetration, market share and wallet share	Enhance and accelerate member engagement with relevant products & offers	Influence digital adoption and track use of traditional channels	Optimize current branch network and identify areas for new locations	Prioritize offers based on member behaviors	Proactively identify and engage at-risk member	Locate potential B2B business opportunities

Increase Awareness and Drive Traffic

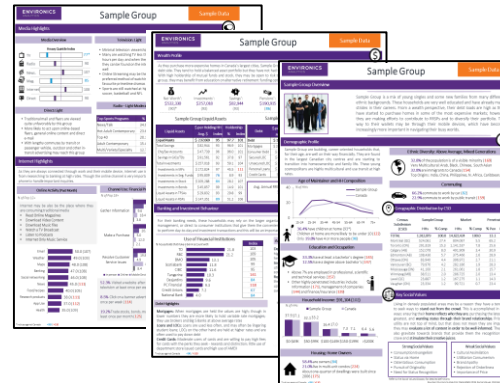
Understand Landscape & Identify Target Consumers



Through the lens of Segmentation

- Understand which consumers buy life insurance today, and how/why they are buying
- Prioritize segments by potential
- Group segments into meaningful targets for marketing purposes

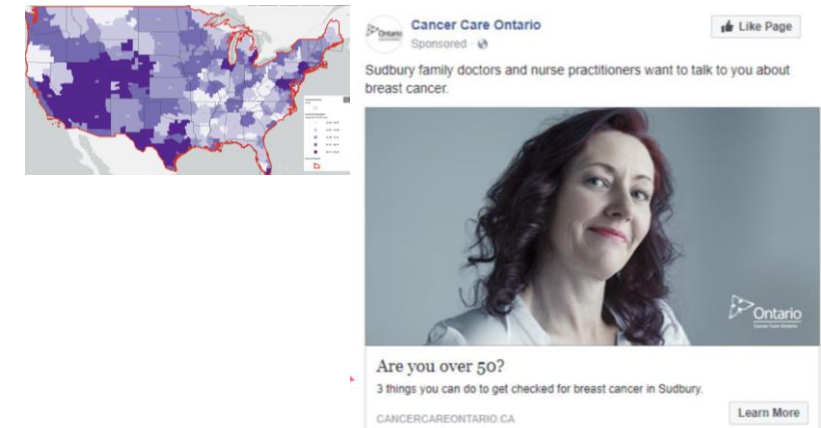
Describe Target Consumers



Create a common view of your consumers'

- Demographic composition
- Financial status and coverage needs
- Social Values - what makes them tick
- Lifestyle & Media habits
- Channel preferences

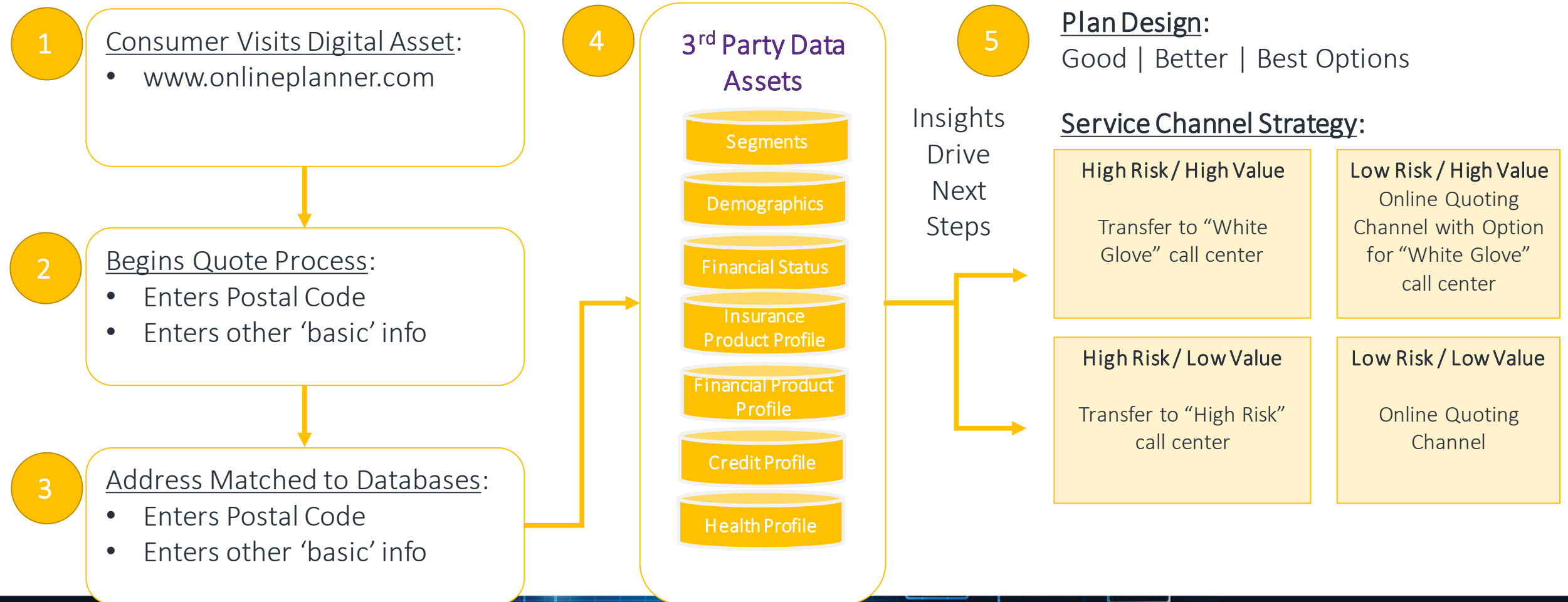
Develop & Implement Acquisition Strategy



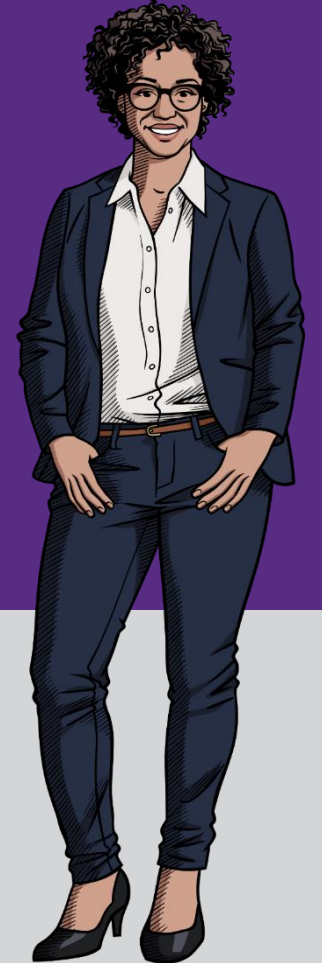
Build out segment-focused strategies

- Markets: Identify markets with highest potential
- Media: Digital, Offline, or Cross-channel
- Message: Appeal to their values and what drives their decisions

Guide the Process and Remove Friction



Case Study: Cancer Care Ontario



Ontario's Average Risk Cancer Screening Programs

The Ontario Cervical Screening Program (OCSP)

4.0 Million



Women 24-69*

2017
% Overdue

38.0%

Gap = 1.6M

The Ontario ColonCancerCheck (CCC)

4.2 Million



Women and Men
50-74*

2016
% Overdue

38.6%

Gap = 1.6M

The Ontario Breast Screening Program (OBSP)

2.0 Million



Women 52-74*

2017
% Overdue

37.2%

Gap = 0.75M

* Age range used to derive the overdue population

Framing our Facebook Campaign Test Case

Objective:

To assess the comparative effectiveness of Environics Analytics' SocialValues informed tailored and non-tailored ad messaging on social engagement among women aged 50-59 years

Scientific Method Used:

To control for bias, the FSAs targeted in the Facebook campaign were randomized; half were assigned a location specific tailored ad and the other half were assigned the non-tailored ad.

Target population counts were also checked for even distribution and confirm highest concentrations of women overdue for a mammogram

Privacy Note:

All FSA's with counts less than six individuals were suppressed and were not included in this campaign

Tailored Ad - Ottawa


VarCode	Social Value	Ontario		Target			
		Base Count	Base %	Count	%	% Pen	Index
SV00012	Community Involvement	2,773,721	24.13	222,551	29.43	8.02	122
SV00060	Personal Control	2,636,588	22.94	215,915	28.55	8.19	124

Community Involvement




Personal Control




Cancer Care Ontario
 Sponsored · 🌐

Like Page

You spend time helping others in your Ottawa community. Do something for yourself today.

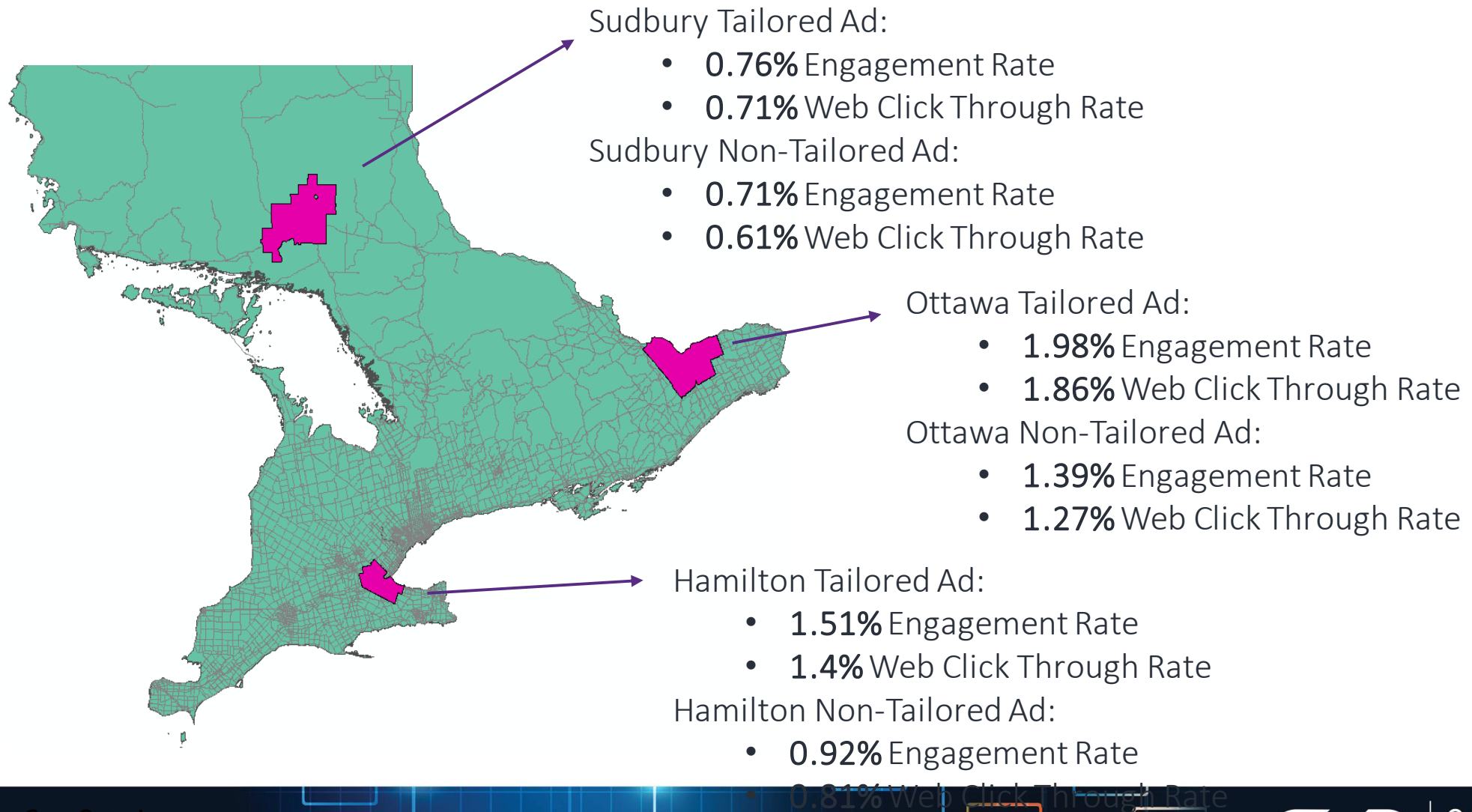


Are you over 50?
3 things you can do to get checked for breast cancer in Ottawa.

CANCERCAREONTARIO.CA

Learn More

Tailored vs Non-Tailored Ad Stats by Location



Overall Performance of Our Geo Tailored Ads

When comparing the performance of our Tailored Ads vs. the Non-Tailored Ad for all three sites:

The **click rate** (# of clicks/# of reached) on average was

1.7 times higher in FSAs receiving tailored messaging!

The **engagement rate** (# of likes, shares, comments/# of reached) on average was

1.6 times higher in FSAs receiving tailored messaging!

The **link clicks** (# of clicks to the campaign link/# of reached) on average was

1.7 times higher in FSAs receiving tailored messaging!

All comparison stats were found to be statistically significant ($p < 0.05$)

Questions?



Stan Ivankovic

Director, Business Development, Insurance

Environics Analytics

Stan.Ivankovic@environicsanalytics.com

