

Technology

Big Data Analytics

Jim Ferrell, Insurance Technologies

Keith Hoeffner, LexisNexis



16th Annual Intercompany Long Term Care Insurance Conference

What is “Big Data”



- Not a new DB design
- Data aggregation does not provide context
- Language (actual data) is created by people and therefore requires human context to interpret



“Big Data” is defined as the aggregation of the digital traces of everything we do.

We are just beginning to understand the value of this data, how “things” are connected and the potential value provided from mining, storing and reporting/studying this data.



If you were able to gather all the data created from the beginning of time to the year 2000, it would be less than we now create in one minute. – Bernard Marr – (Big Data explained to anyone In less than 2 minutes – To Absolutely anyone)

Over 90% of the data available in the world was created in the past 3 years



Big Data is the type of change that will have implications for everyone, regardless of awareness or desire.

Big Data is completely transforming portions of our business model.

Big Data impacts everyone.....



This innovation brings with it legitimate concern for privacy.

Think about the first time you realized your web browser was tracking your usage and using it to make targeted suggestions...

Or the first time Facebook presented an advertisement for something you just searched for....

Examples of value based usage



Retailers are leveraging what's trending on social media or search trends to customize their supply chains.

Credit card companies are leveraging big data to auto-detect fraudulent activities.

Cities are leveraging big data to minimize traffic by enabling “smart” traffic signals.



Hospitals are analyzing medical data and patient records to predict patient discharge readiness.

Red Roof Inn leveraged weather, and flight cancellation data and built an algorithm which considered these and other variables to generate targeted mobile marketing campaigns to stranded travelers.

Big data addresses business challenges across the insurance continuum



Insurers and distributors face many challenges related to acquiring, growing and retaining their customers



Marketing data to enhance and inform target market identification



Consumer Characteristics

Age and DOB
Gender
Marital Status
Income
Education
Occupation
Presence of Children
Household Composition
Homeownership
Census Characteristics

Lifestyle / Behavioral Data

Direct Mail Responder / Buyer
Online Product Responder / Buyer
Contributors to Charities
Crafts/ Hobbies
Sports Enthusiasts
Pet Owners
Opportunity Seekers
Motorcycle Owners
Boat Owners
RV Owners
Snowmobile Owners



Financial Data/Scores

Mutual Funds / Annuity Investors
Household Deposit Score
Investment Balance Score
Investible Assets Score
Net Assets Score
Wealth Indicator
Home Purchase Amount
Home Current Value
Mortgage Amount & Lender
Investment Property

Segmentation Data

LexisNexis Insurance Personas
- Insurance Derived Segmentation
- Coverage Propensity Profiles
Spanish Language Probability Indicator

Hotline Event Data

New Movers
New Homeowners

The Facts

- National Coverage
- 200 million individuals
- 110 million households
- 70 million homeowners

The Strengths

- Multi-sourced for enhanced accuracy
- Depth of data for directing messages and offers
- Attribute selections for expanded targeting and segmentation

Marketing propensity scores to enhance targeting and results performance



Target Prospects who meet your preferred profile criteria

Use Case	LexisNexis Custom Life Industry Solution
Risk Prediction	Marketing Risk Classification Models: Public record and demographic derived score that predicts the likely risk class of a consumer (e.g., standard, preferred, etc.)
Shopping Propensity	Channel Preference Models: Predicts how likely a consumer will shop a specific channel (direct, independent or exclusive)
	Shopping Activity Score: Predicts the propensity and implied frequency of insurance shopping (<i>on development roadmap</i>)
Cross-Sell Response Prediction	Product specific Cross-sell Scores: Predicts the likelihood of response to Annuity, Mortgage Protection, Medicare Supplement, Final Expense, and Long Term Care offers
Retention Improvement	Direct Payment Score: Predicts the likelihood of a consumer to elect direct payment of premium (EFT payment)
	Life Attrition Model: Predicts the likelihood of policyholder attrition
Wealth Identification	Investment and Deposits Score: Estimate of total household bank and investment balances
	Investable and Net Assets Score: Est. of all household investable assets: deposits, stocks, mutual funds, retirement accts, etc. , minus all liabilities
Lifetime Value Scores	Standard or Custom combination of above models into a Marketing Optimization score

The choices are abundant

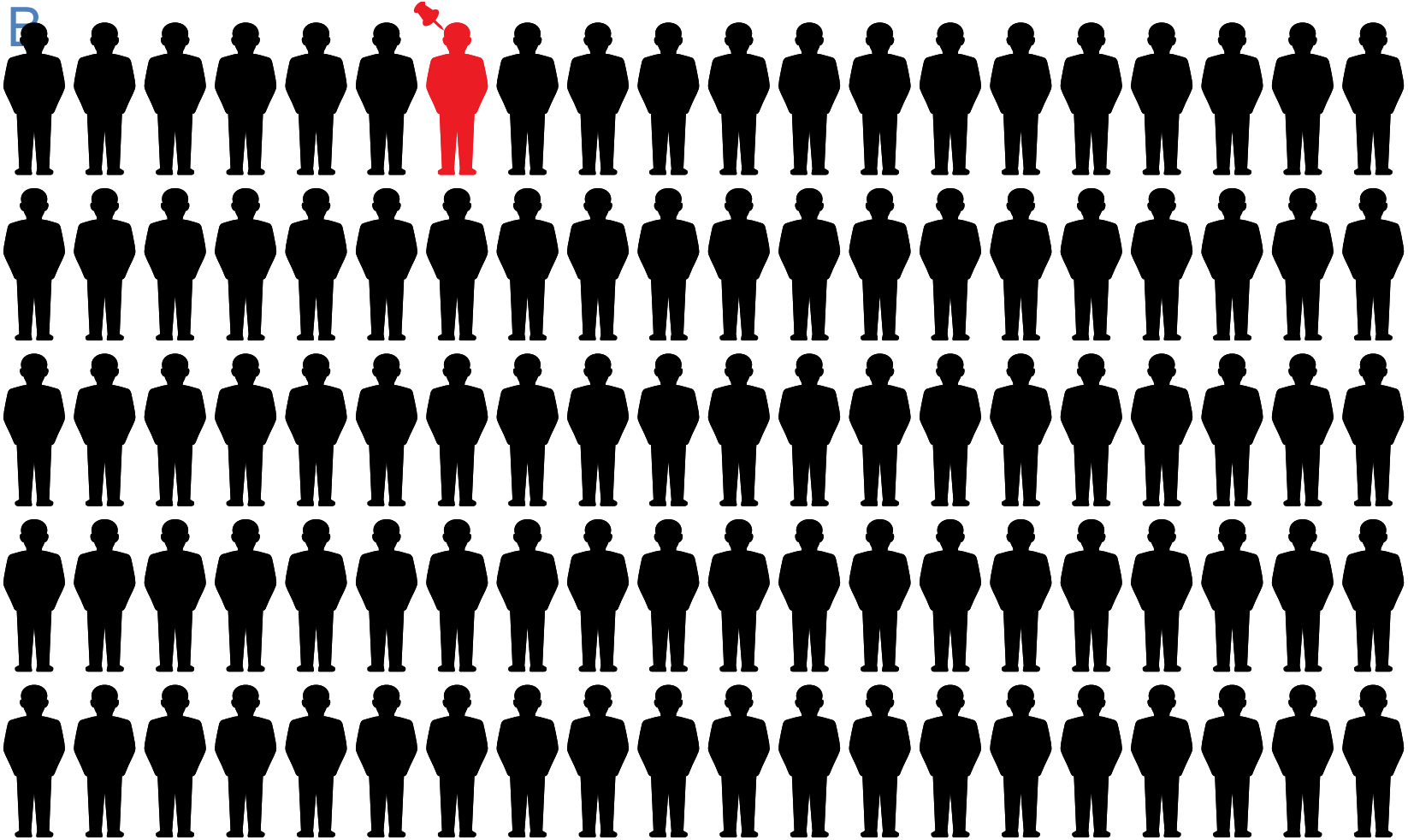


Succession planning
Scarce internal resources
Legacy systems
Low Interest Rates
Competition
NIGO FRAUD
Regulation

What's so scary about fraud, anyway?



\$80B > \$12



Changing consumer perceptions about fraud



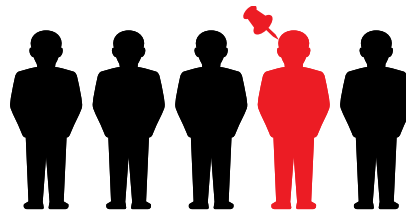
Percent of people who say they know someone who has committed insurance fraud?

31%

Percent of those respondents who say they have **reported** someone for committing fraud?

17%

Four of five adult Americans think a variety of insurance crimes are unethical.



Source: CAIF Study, *Four Faces: Why some Americans do—and do not—tolerate insurance fraud*

The prevalence of identity fraud: what's really out there



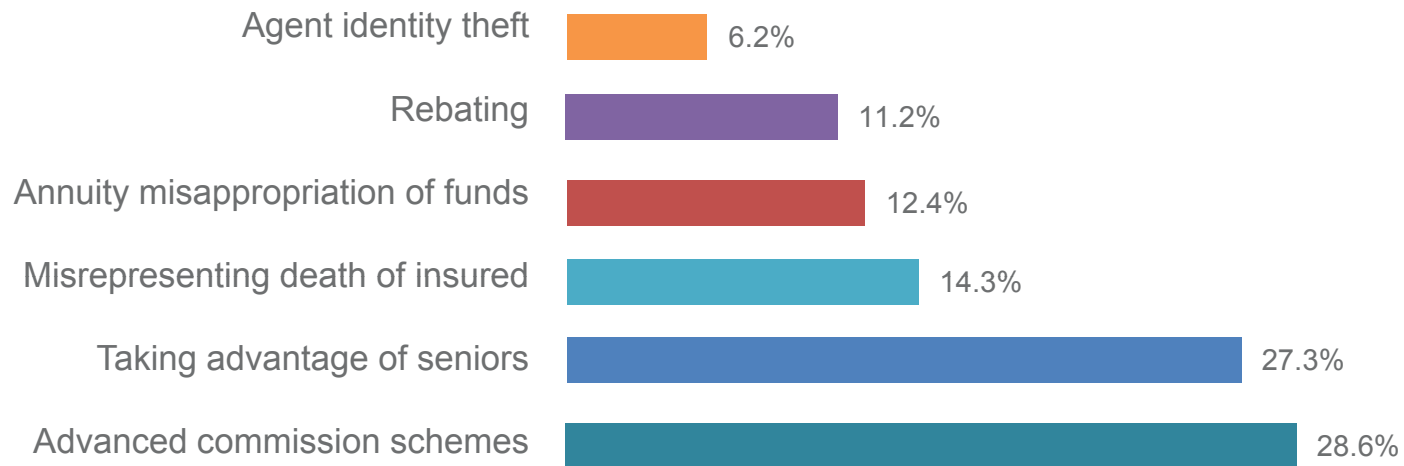
Insights from transactions submitted to LexisNexis® for identity verification in December 2011

- **657,092** identities linked to 3 or more SSNs
- **84,093** identities linked to 4 or more dates of birth
- **77,458** identities using a SSN belonging to a different name and address
- **53,317** applications from a vacant address
- **43,672** identities using a SSN belonging to a deceased individual

Focus on fraud: CEFLI Poll conducted June 2013



What type of fraud schemes is your company most concerned with?



Source: An Advanced Discussion on Fraudulent Schemes and Deterrents. *CEFLI Educational Webinar Series June 18, 2013*

How's that working for you?



Historical approach

- Most fraud caught at time of claim
- Reactive approach
- Relies heavily on manual processes, experienced claim reps
- Automated exception processing—only catches outliers and obvious low hanging fruit

88% currently employing anti-fraud
technology

BUT

**Less than half are using
technology for non-claims
functions**

“The State of Insurance Fraud Technology”
Sept 2012, CAIF and SAS

Real world examples: Identity fraud



The Coach



Narcissistic Nephew



Resurrected Rita



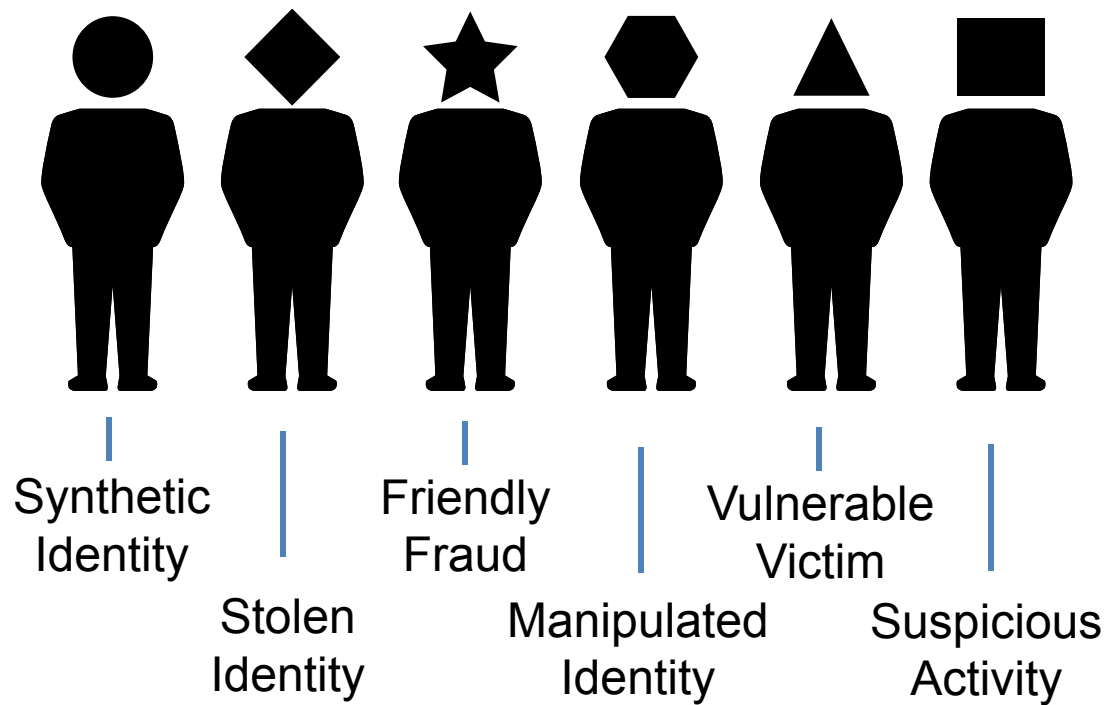
Today's Challenges:

- Labor intensive and reactive
- Difficult due to inability to see connections outside your own data
- Don't know about it until it's too late

The Future:

- Combining internal and external data
- Using relationship analytics to link and analyze vast quantities of data, exposing hidden collusion and revealing networks
- Create predictive models to help instantly detect fraud in real-time

Detecting identity fraud is especially challenging because not all identity fraud schemes are the same and they are constantly evolving



While a great deal of identity fraud is the result of a stolen identity, there are growing trends in identity fraud schemes

A look inside the lab: Social network analytics



The background features a complex social network graph with nodes of various sizes and colors (blue, green, yellow) connected by thin lines. Three large, semi-transparent circles are overlaid on the graph, each highlighting a different cluster of nodes. The top circle is orange and labeled 'Shared Historical Addresses'. The bottom-left circle is blue and labeled 'Shared Assets (Property, Vehicles etc.)'. The bottom-right circle is purple and labeled 'Shared Business Ownership'.

Shared
Historical
Addresses

Shared
Assets
(Property,
Vehicles
etc.)

Shared
Business
Ownership

Identity clustering (hospital/nursing facility)



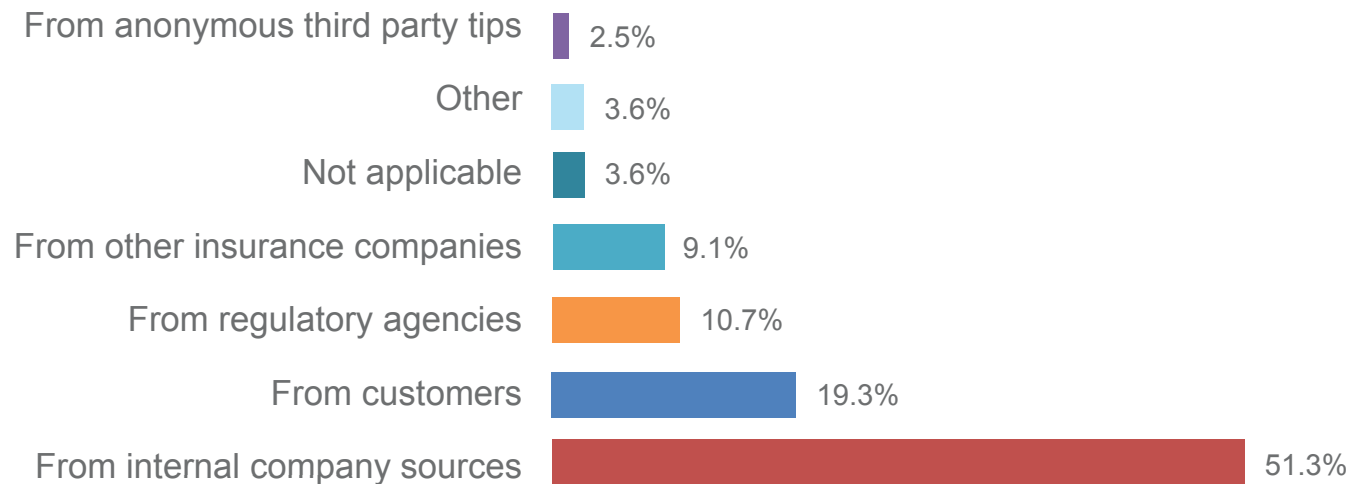
**1500 E. MEDICAL
CENTER DR.
ANN ARBOR,
MI 48109**



First, get help



What is your company's primary means for obtaining information leading to life insurance fraud-related cases?



Source: An Advanced Discussion on Fraudulent Schemes and Deterrents. CEFLI *Educational Webinar Series* June 18, 2013

Five questions for mitigating fraud risk



1. **Resolve:** “Who are you?”
2. **Verify:** “Do you exist?”
3. **Authenticate:** “Are you who you say you are?”
4. **Evaluate:** “Can I do business with you?”
5. **Alert:** “Are you exhibiting high-risk behavior?”



Be proactive in fighting fraud now...



Getting Started – Using Big Data



Develop a Data Strategy

Use external sources of "Big Data" integrate into existing practices

Conduct data Tests

Consider Predictive Modeling

What do you do?



In the LTC space, “Big Data” has been viewed as the CRM and marketing data we aggregate.

What if we leveraged medical data or prescription history which allowed us to weight targeted age groups with specific medical trends as possible leads?

What do you do?



What about the organizational and technical obstacles related with gathering and analyzing this data?

The largest obstacle is access to and aggregating the data in a way that makes it useful.

What do you do?



You don't need to have all your analytics defined prior to beginning to store the data.

However, you do need at least one specific goal in mind...

Much of the value will be unknown, and uncovered once aggregated and walking thru the analytics correlating the data.



It's defined as;

“the systematic computational analysis of data or statistics...”

Or

“information resulting from the systematic analysis of data or statistics....”



The key is not in identifying what has happened, but in identifying patterns or narrowing down the possible future scenarios.

- Uncover hidden patterns
- Unknown correlations
- Trends
- Customer preferences
- Unknown useful business information



But Predictive Analytics is **EXTREMELY DIFFICULT**

1. Data Science not effective without the **RIGHT** data
2. Poorly defined data
3. Goals not in alignment with data
4. Reluctance to begin without perfect data
5. Shortage of data analytic skills



Thank You