Business Forecasting and Analytics Forum
September 19-20 • Chicago, IL

Using Big Data to Enhance Demand-Driven Forecasting & Planning

Using real-time information to sense demand signals and respond quickly to changes in demand

September 20, 8:30am

Gary Angel – Ernst & Young

Widely considered one of the leading digital measurement experts in the world, Gary leads EY’s Digital Analytics Practice. EY acquired Gary’s previous company – Semphonic – in March of 2013. As Semphonic’s President and co-Founder, Gary led Semphonic’s growth over a 15 year period from a 2-person practice to the one of the leading digital analytics practices in the United States. Voted the most Influential Industry Contributor by the Digital Analytics Association in 2012, Gary writes an influential blog (http://semphonic.blogs.com/semangel), has published more than twenty whitepapers on advanced digital analytics practice and is a frequent speaker at industry events.

View presentation online at:
https://jpkgroupsummits.com/attendee5
Forecasting & Big Data

with Gary Angel
You know what forecasting is…

But do you know what big data is?
It’s not this…

<table>
<thead>
<tr>
<th>Big data</th>
<th>“Not so big” data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Volume</strong></td>
<td>Megabytes/gigabytes</td>
</tr>
<tr>
<td>Terabytes/petabytes</td>
<td></td>
</tr>
<tr>
<td>Variety</td>
<td>Structured/relational</td>
</tr>
<tr>
<td>Unstructured (text, voice, video)</td>
<td></td>
</tr>
<tr>
<td>Velocity</td>
<td>Data at rest</td>
</tr>
<tr>
<td>Data in motion (streaming)</td>
<td></td>
</tr>
<tr>
<td>Veracity</td>
<td>Trusted/cleansed</td>
</tr>
<tr>
<td>Untrusted/uncleansed</td>
<td></td>
</tr>
</tbody>
</table>

- It sounds plausible, but doesn’t it make big data sound like just more of the same?
Traditional data and the Four V’s

Traditional data (however large, fast, inaccurate or various)

- Account
- Date
- Address
- Status
- Opt-in
- Name
- Spend
- Purchases
- Last Activity

It doesn’t matter how large this gets
It doesn’t matter how many sources you have
It doesn’t matter how fast these change

Traditional database, reporting and analysis tools can handle it
But knock things down a level …

Big data
(however small, at rest, accurate or singular)

And you introduce three critical new dimensions

Order
The sequence of events matters

Time
The time between events matters

Pattern
The pattern of events matters
And the order of events matters

This set of data

Saw ad  Web visit  Call  Forecast Sale

is completely different than this

Sale  Web visit  Saw ad  Call  Forecast an Annoyed Customer
The time between events matters

These events in time

Saw ad  | Web visit  | Call  | Forecast
May 3rd | May 4th   | May 7th | Sale

Yield a different forecast than these events in time

Saw ad  | Web visit  | Call  | Forecast
Jan. 2nd | Mar. 3rd   | July 12th | Unclear
And the pattern of events matters

This consumption pattern

Weekend | Weekday | Weekend | Weekday
---|---|---|---
High Use | Low Use | High Use | Low Use

Has the same average but a different forecast

Weekend | Weekday | Weekend | Weekday
---|---|---|---

Weekend Forecast | Weekend Forecast
Traditionally, we’ve handled this by creating sub-forecasts

This consumption pattern

Weekend: High Use, Weekday: Low Use, Weekend: High Use, Weekday: Low Use

Weekend Forecast, Weekday Forecast
It’s true in every paradigm case of big data

► Big data shifts analytics to patterns in detail-level data.

**Utilities**
- Old: Read meter once a quarter
- New: Read meter every 10 minutes
  → Aggregate stream
  → Analyze the flow of usage by day/time

**Digital**
- Old: Track total views by page
- New: Read pages
  → Aggregate stream
  → Analyze visitor intent and success by visit

**Sales**
- Old: Forecast using Aggregate Averages
- New: Track each event
  → Aggregate and Predict Account
  → Aggregate predictions
It matters because when order, time and pattern of events are important

- Relational data models break down
  - SQL is extremely cumbersome and almost impossible to use when querying order, time between and patterns

- Stats tools are limited
  - Traditional analysis methods (like regression and correlation) don’t work directly

- Integration strategies break
  - Traditional joins become multi-joins and no longer facilitate useful integration, even when keys are present
Which means that when it comes to big data…

Almost *everything* you took for granted in traditional forecasting is broken!
What is big data

So if you didn’t know… now you know
Three common big data scenarios

1. You have unstructured data (text, voice or video) from sources like Twitter, Call-Center Notes, or published video content.

2. You need to apply basic or even complex statistical analysis to identify known patterns (improving, seasonal, etc.) at the individual level before aggregating.

3. You need to identify patterns in the data before you can understand what it means.
Structuring unstructured data

1. You have unstructured data (text, voice or video) from sources like Twitter, Call-Center Notes, or published video content

- Survey Research
- Social media
- Call-Center
- Feedback
- Offline research

Once text actions are categorized they can be used with traditional forecasting techniques.
But that doesn’t work when patterns are complex or individual

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>$130</td>
<td>$142</td>
<td>$145</td>
<td>$158</td>
<td>$167</td>
<td>$187</td>
</tr>
<tr>
<td>C2</td>
<td>$130</td>
<td>$82</td>
<td>$145</td>
<td>$25</td>
<td>$135</td>
<td>$15</td>
</tr>
<tr>
<td>C3</td>
<td>$130</td>
<td>$150</td>
<td>$145</td>
<td>$18</td>
<td>$25</td>
<td>$20</td>
</tr>
</tbody>
</table>

ACCELERATING

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>$130</td>
<td>$142</td>
<td>$145</td>
<td>$158</td>
<td>$167</td>
<td>$187</td>
</tr>
<tr>
<td>C2</td>
<td>$130</td>
<td>$82</td>
<td>$145</td>
<td>$25</td>
<td>$135</td>
<td>$15</td>
</tr>
<tr>
<td>C3</td>
<td>$130</td>
<td>$150</td>
<td>$145</td>
<td>$18</td>
<td>$25</td>
<td>$20</td>
</tr>
</tbody>
</table>

OSCILLATING

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>$130</td>
<td>$142</td>
<td>$145</td>
<td>$158</td>
<td>$167</td>
<td>$187</td>
</tr>
<tr>
<td>C2</td>
<td>$130</td>
<td>$82</td>
<td>$145</td>
<td>$25</td>
<td>$135</td>
<td>$15</td>
</tr>
<tr>
<td>C3</td>
<td>$130</td>
<td>$150</td>
<td>$145</td>
<td>$18</td>
<td>$25</td>
<td>$20</td>
</tr>
</tbody>
</table>

SEASONAL

NO SINGLE AVERAGING STRATEGY WILL YIELD THE RIGHT ANSWER
Applying analytics to each record

2. You have structured data from sources. You need to apply basic or even complex statistical analysis to identify known patterns (improving, seasonal, etc.) at the individual level before aggregating.

Create an individual forecast for each record. Then aggregate traditionally.
Finding potential patterns in the data

You need to identify patterns in the data before you can understand what it means.

Use pattern matching techniques to spot and identify behavioral signatures and then categorize them for aggregation.
Summing up

1. Big data isn’t just having a lot of data – big data problems occur when the level of meaning resides above the detail level in the sequence, order or pattern of the events.

2. When that happens, traditional IT techniques break down.

3. And so do forecasting techniques.

4. To make big data useful, you need to add an extra step to your process where you categorize the data in ways that allow it to be aggregated.

5. There are three common ways to do that: applying text analytics to structure unstructured text, applying statistical analysis to identify individual-level trends, and applying pattern-matching to categorize behavior.

When you’ve done that, you can forecast with big data.
Thank you

Gary Angel

► gary.angel@ey.com
► LinkedIn: www.linkedin.com/pub/gary-angel/0/176/a43/
► +1 415 894 8255
► @garyangel

Get it on Amazon:
About EY
EY is a global leader in assurance, tax, transaction and advisory services. The insights and quality services we deliver help build trust and confidence in the capital markets and in economies the world over. We develop outstanding leaders who team to deliver on our promises to all of our stakeholders. In so doing, we play a critical role in building a better working world for our people, for our clients and for our communities.

EY refers to the global organization, and may refer to one or more, of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. For more information about our organization, please visit ey.com.

Ernst & Young LLP is a client-serving member firm of Ernst & Young Global Limited operating in the US.

© 2013 Ernst & Young LLP.
All Rights Reserved.