



An Agent-Based Framework for the Active Multi-Level Modeling of Organizations

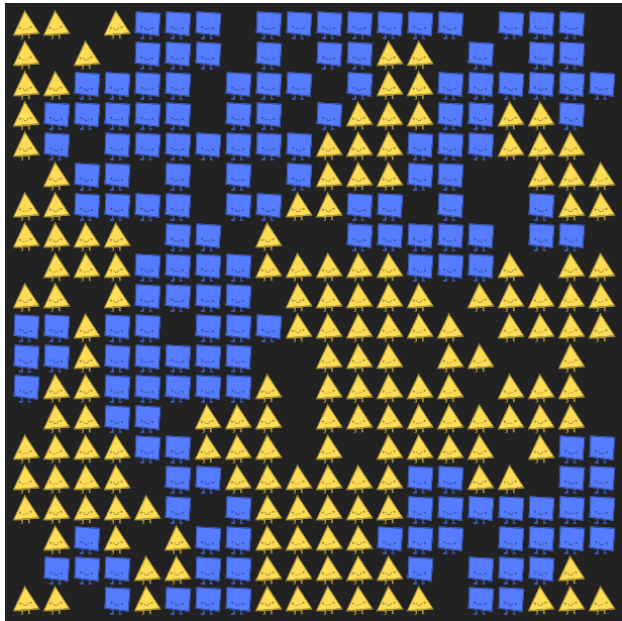
Geoffrey P. Morgan
Kathleen M. Carley

Key Points

- Multi-Level Phenomena exist, and are of significant interest to organizational researchers
- Multi-Level Phenomena require Active Multi-Level Modeling to create good/useful models
 - Active Multi-Level Modeling is ABM at multiple levels of granularity
 - A specific sub-family of multi-modeling
- Data-Centric Modeling is key to rapid development of topical/applied models : data must support instancing at different granularities



Agent-Based Models and Emergence



Parable of the Polygons
(<http://ncase.me/polygons/>)
– An interactive version of
the Schelling Model



A swarm of Blue Jack Mackerel form a "bait ball", which confuses predators

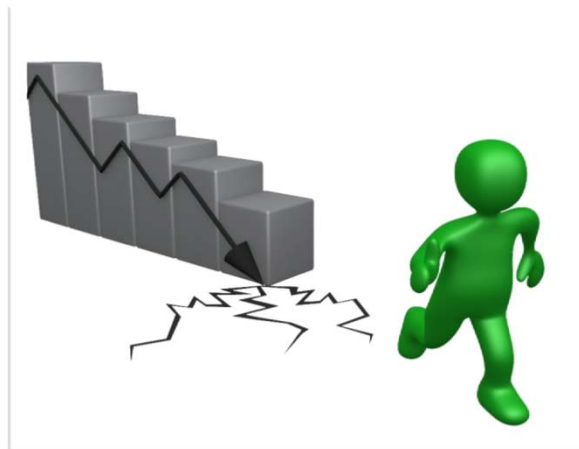
Beautiful elegant useful models with a
single level of agency

Multi-Level Phenomena

Definition: *Phenomena that occur with actors at multiple levels of granularity*



Change Resistance



Turnover



Organizational Resilience

Change Resistance

Definition: Overt or covert resistance to a introduced organizational change

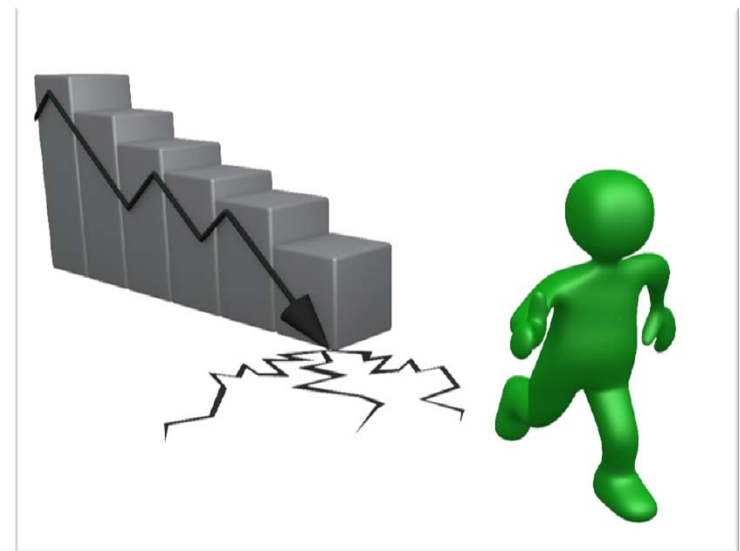
- Noted Risk Factors:
 - Individuals feel they are harmed by the change
 - Individuals don't understand the need for the change
 - Organization has gone through multiple changes in recent memory
 - Organizational culture is not transparent and trusted
 - Organization introduces change poorly



Turnover

Definition: The percentage of workers leaving an organization over a given unit of time

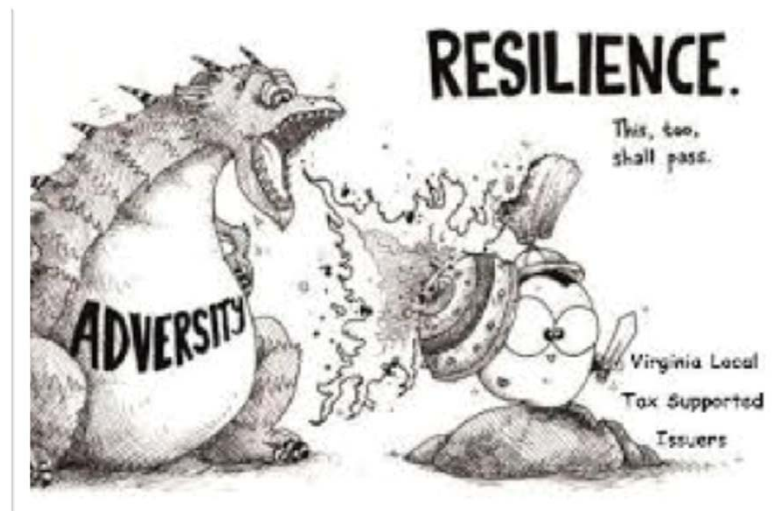
- Noted Factors:
 - Organizations that are going through significant challenges experience higher turnover
 - Individuals which are highly embedded in the organization are much less likely to leave
 - Organizational socialization procedures can reduce chance of turnover



Organizational Resilience

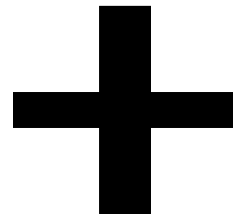
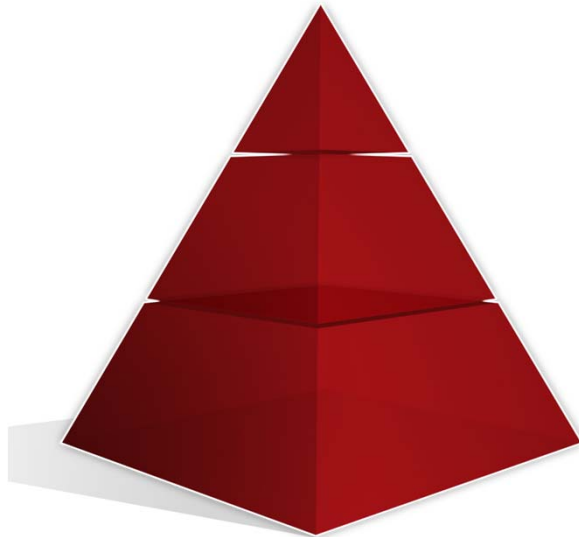
Definition: The ability of an organization to anticipate, prepare for, and respond to organizational crises

- Noted Factors:
 - Groups are resilient if they are more heterogeneous
 - Individuals are more resilient if they have resources they need
 - Resilience response is shaped by the organization's ideology
 - Organizational guidance which ignores individual practice harms resilience

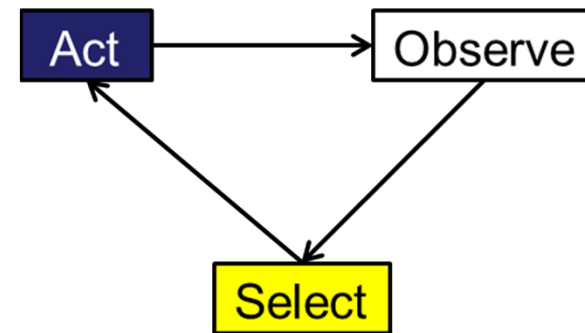


Multi-Modeling for Multi-Level Phenomena

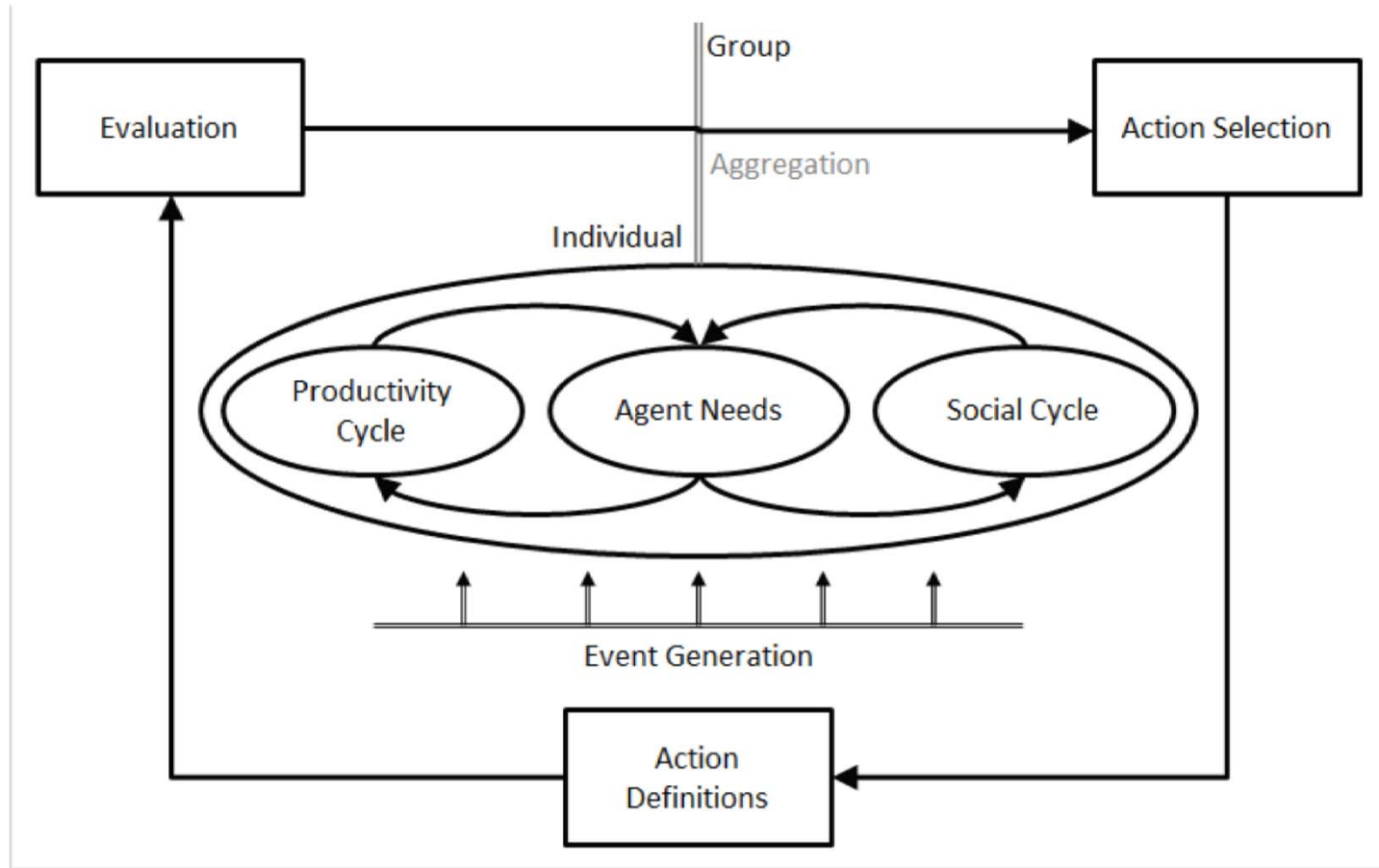
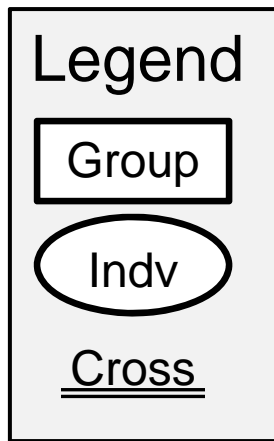
Org Model



ABM Decision Loop



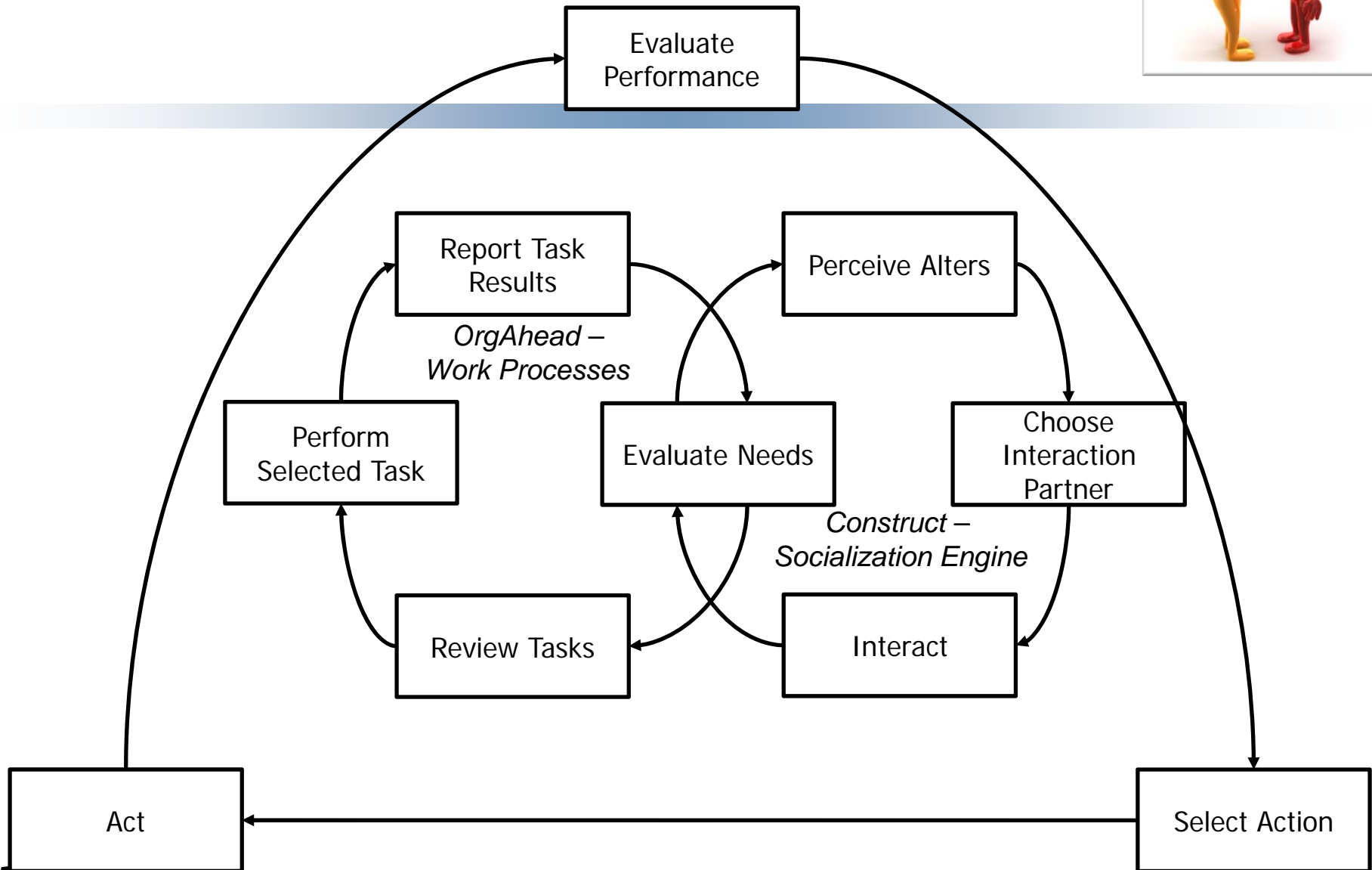
Proposed Framework



Three+ models with intentional inter-operation!

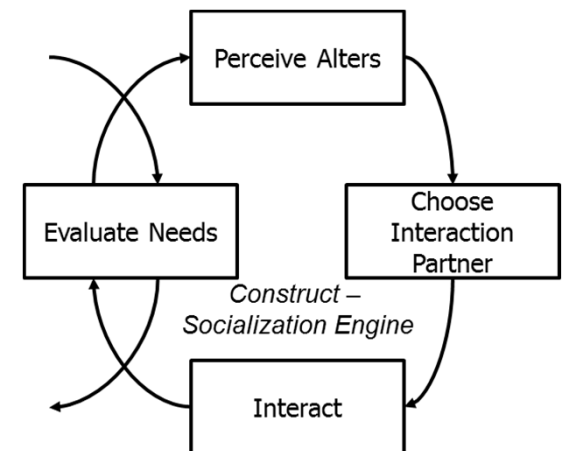


MultiOrg Simulation Cycle



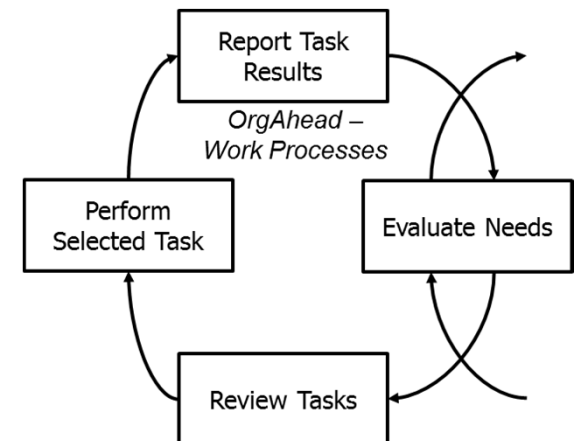
Construct Mechanisms

- Construct is a simulation which focuses on interactions between actors and the diffusion of knowledge between actors
- Construct mechanisms use:
 - Agent x Agent: “who knows who?”
 - Agent x Knowledge: “who knows what?”
 - Agent x Group: “who belongs to what?”
- Construct agents socialize 1 to 1 and may exchange information about:
 - Themselves
 - Other agents
 - Knowledge they have



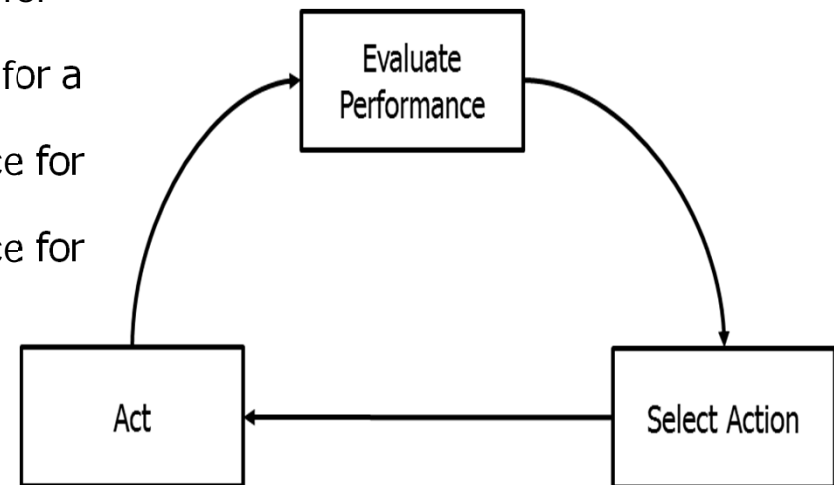
OrgAhead Mechanisms

- OrgAhead is an organization simulation focuses on agents doing tasks – predicting what tasks the agent should do
- OrgAhead agents prioritize doing work they think will be rewarded, so they need to perceive the priorities of groups to which they belong:
 - Agent x Group – “who belongs to what?”
 - Agent x Knowledge – “What can I do?”
 - Infers Group x Knowledge – “What does the group I’m part of care about?”
- OrgAhead Agents choose and perform tasks



Group Mechanisms

- Mutual Learning Simulation
Indoctrination
 - Agents are socialized to learn and prioritize things related to their groups
 - Group priorities can conflict!
- Active Organizational Performance:
 - The group can:
 - Inhibit specific task performance for all individuals
 - Inhibit specific task performance for a single individual
 - Promote specific task performance for all individuals
 - Promote specific task performance for all individuals.
 - Detach group members



Data-Centric Modeling



James Hardy/AltoPress/Maxppp

Informing the Model with Data

- Groups:
 - Structural: Informed via Clustering
 - Functional: Informed via Survey Data
- Knowledge:
 - Word Network (Agent x Word)
 - Words selected by polarization score



Horizontal Merger of a Multinational

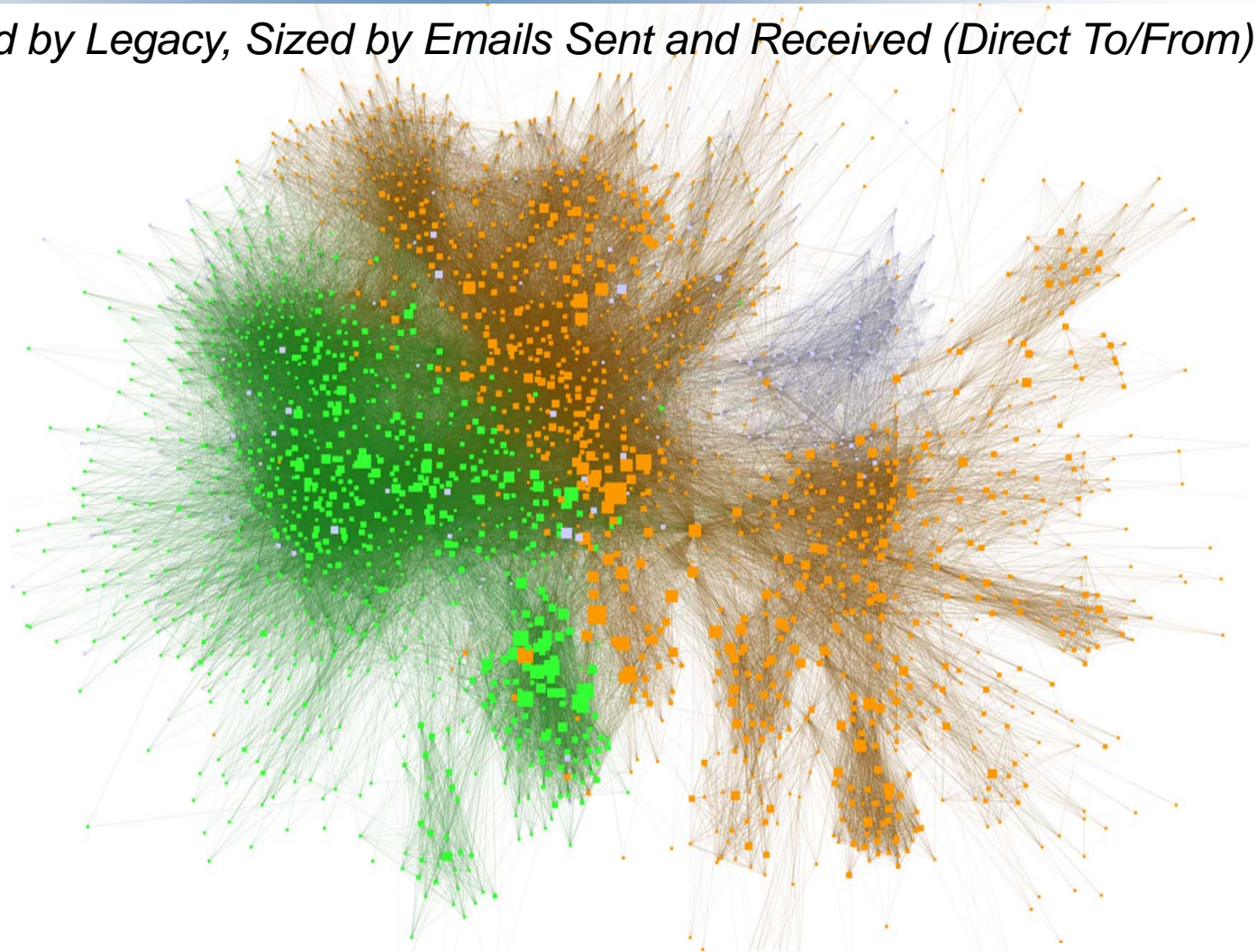
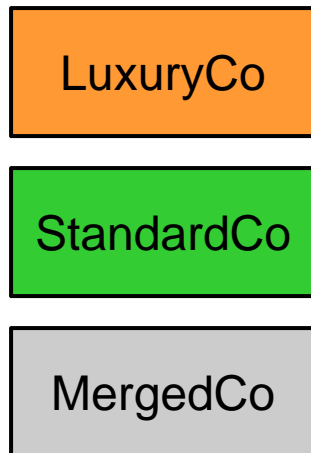


- Large Multinational, purchased another large company
 - Wants to understand the integration process
 - Asked academic researchers if they wanted to help
- Allowed collection of email-server data for multiple months at multiple points in time
 - Collection Period 1: Right after merger announcement
 - Collection Period 2: A year later
 - Collection Period 3: Another year later
- Encouraged employees to participate in org surveys administered by research team

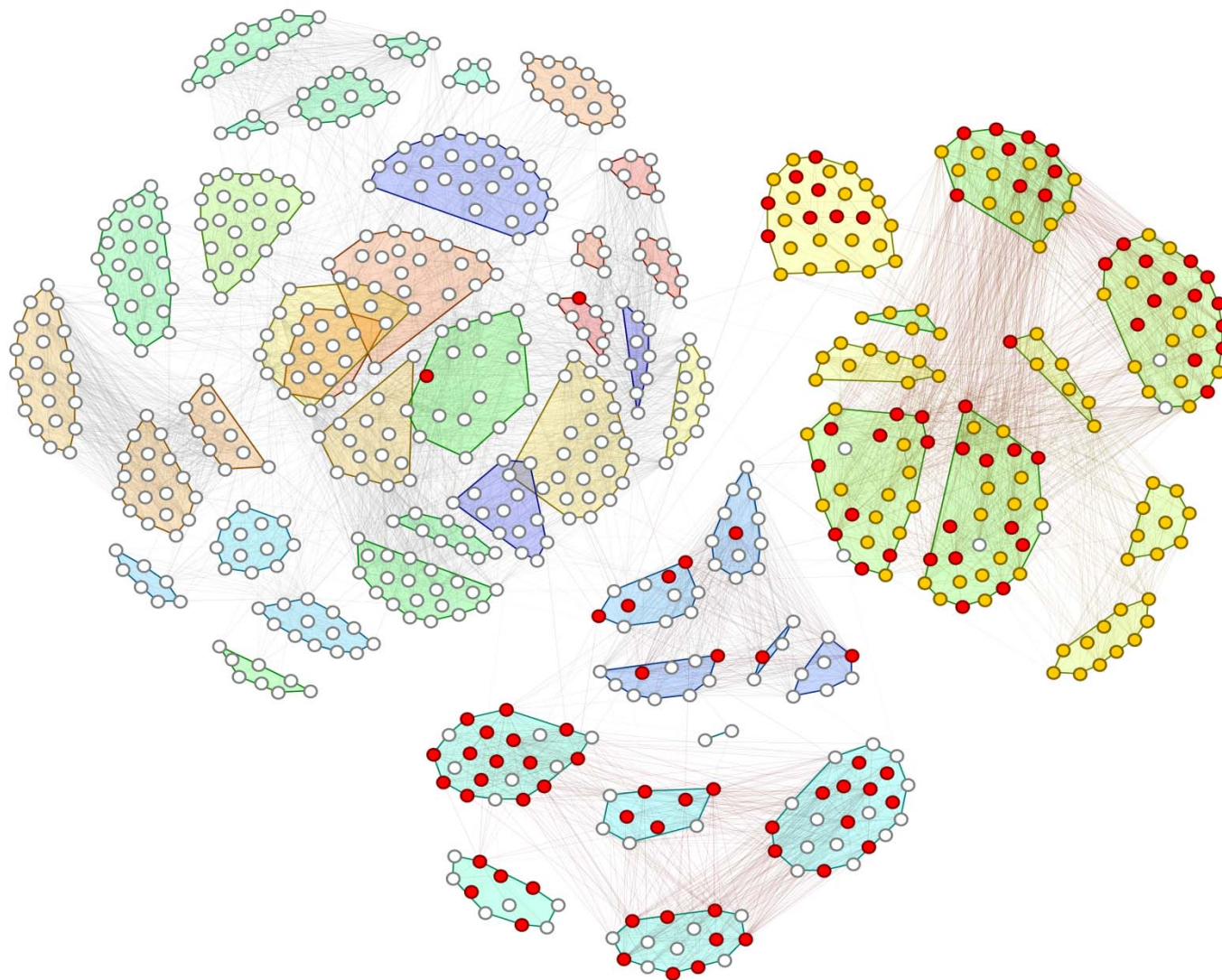
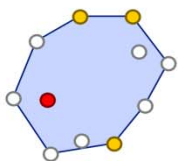


Internal Email Interactions

Employees - Colored by Legacy, Sized by Emails Sent and Received (Direct To/From)



Group Structures via Clustering



IDENTIFYING KNOWLEDGE VIA CONTENT PROCESSING



De-Identification

- Legal Requirement!
- Used Stanford NER (Named Entity Recognizer) to identify and then de-identify:
 - People
 - Locations
 - Organizations



De-Identifying Entities Consistently in Unstructured Content

- First, identify and create anonymous mappings for all NER tokens
 - Replace proper names with tokens:
 - “Jean Paul” = “Name_1”
 - “Abe Lincoln” = “Name_2”
 - Replace locations with tokens:
 - “San Francisco” = “Location_1”
 - “New York” = “Location_2”
 - Replace organizations with tokens
 - “Bank of Omaha” = “Org_1”
 - “IKEA” = “Org_2”
- Replace all numeric characters with ‘#’
 - ###-###-#####
 - ##-###
 - ##,###



Using Content as a Proxy for Knowledge

Every organization has its unique jargon, informed by the collective backgrounds and contributions of all members.

1. Can we identify words or tokens that are consistently and regularly associated with LuxuryCo and StandardCo?
2. Is the overall language of LuxuryCo and StandardCo becoming more or less similar?



Token Score

- For token t of all Tokens T , we have group A , G , and a Prior P
- We have two terms:
 - the token's odds score based on percentage appearance in the A and G 's documents, but we flatten out marginal cases
 - the token's appearance in A or G (depending on the odds ratio outcome) subtracted against the percentage appearance of the token in Prior P

$$S(t) = flattenedOdds(t) * freq(t)$$

$$flattenedOdds(t) = \begin{cases} abs(odds(t)) > .1, odds(t) \\ else, 0 \end{cases}$$

$$odds(t) = \left(1 - \left(\frac{1}{\left(\frac{|t_A|}{|T_A|} / \frac{|t_G|}{|T_G|} \right)} \right) \right) - .5$$

$$freq(t) = \begin{cases} odds(t) \geq 0, \max\left(\frac{|t_A|}{|T_A|} - \frac{|t_P|}{|T_P|}, 0\right) \\ odds(t) < 0, \max\left(\frac{|t_G|}{|T_G|} - \frac{|t_P|}{|T_P|}, 0\right) \end{cases}$$



Example, "relax"

Group A uses "relax"
 100 times in a corpus
 of 10,000 total word
 instances. Group B
 uses it 10 times in a
 corpus of 5,000
 instances. The Prior P
 has the word 30 times
 out of 40,000
 instances.

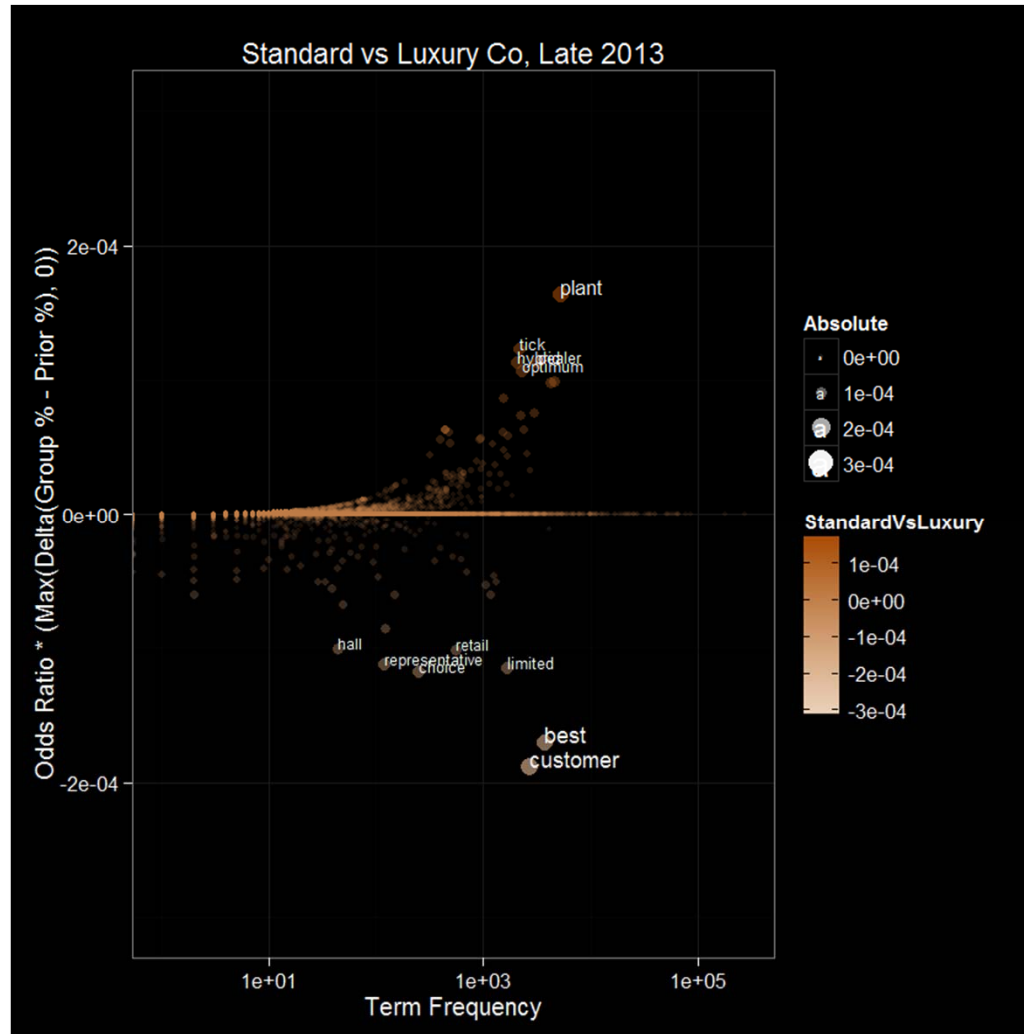
$$S(t) = .002775 = .3 * .00925$$

$$flattenedOdds(t) = .3 = abs(.3) > .1, .3 \\ else, 0$$

$$odds(t) = .3 = \left(1 - \left(\frac{1}{\left(\frac{100}{10000} / \frac{10}{5000} \right)} \right) \right) - .5$$

$$freq(t) = .00925 = .3 \geq 0, \max \left(\frac{100}{10000} - \frac{30}{40000}, 0 \right)$$

Illustrative Graphic, Late 2013



Difference Score

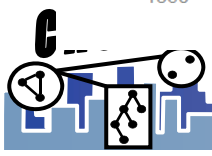
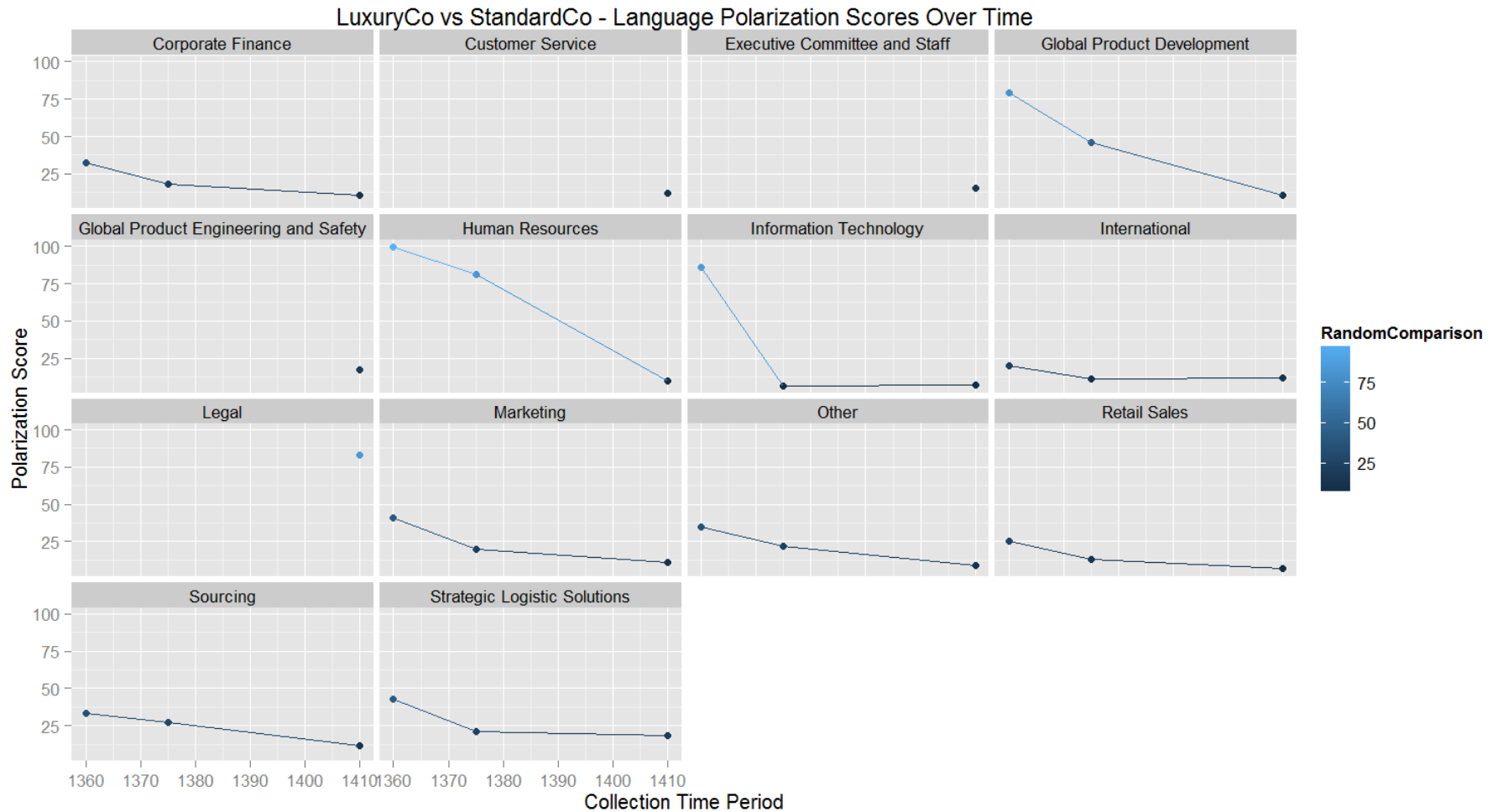
- We can sum the absolute value of the token scores to evaluate how different the two groups are in language after accounting for a prior

$$Score(T, A, G, P) = \sum_t abs(s(t))$$

- We normalize this by dividing the found score by an polarization of the corpus as a whole via multi-sampling.
- Numbers range from 0 to Infinite, where 1 means this is only as polarized as random (so, not very polarized), while 50 would mean this found is 50x stronger than random.



Polarization Over Time





Leveraging Data

Groups from structure

Knowledge from highly polarized words

Research Questions:

1. Are we better able to replicate multi-level phenomena than single-level models?
 1. Can we predict departure from the organization?
 2. Can we predict conflict between organizational units?

Key Points

- Multi-Level Phenomena exist, and are of significant interest to organizational researchers
- Multi-Level Phenomena require Active Multi-Level Modeling to create good/useful models
 - Active Multi-Level Modeling is ABM at multiple levels of granularity
 - A specific sub-family of multi-modeling
- Data-Centric Modeling is key to rapid development of topical model : data must support instancing at different granularities

