Leveraging Diversity
Outline

The Path of Inclusion
Identity and Cognitive Diversity
Prediction
Problems Solving
Case: The Netflix Prize
Takeaways
Framework

**Tools:** Diverse Perspectives, Heuristics, and Interpretations

**Tasks:** Problem Solving and Prediction
The Path of Inclusion
Hiring diverse people is the **right** thing to do.
Hiring diverse people is the **required by law**.
Seeking diversity enlarges the pool and results in better employees.
Diversity

Ability
Diversity is a strategic advantage. It makes organizations more productive and more innovative on cognitive tasks.
Diversity

Ability
Identity and Cognitive Diversity
Poison Rats
Se pa manje rat
Pwazon Rat

$ 5
Prediction
Iowa Electronic Markets

IEM Prices
Obama 0.535
McCain 0.464

Final Gallup Poll
Obama 0.55
McCain 0.44

Actual Outcome
Obama 0.531
McCain 0.469
Methods of Divination

Stars and Planets (astrology)
Rolling Dice
Tarot Cards
Palm Reading
Crystal Balls
Head Shape (Phrenology)
Atmospheric Conditions
Dreams
Animal Entrails
Moles on the body

Lightning
Smoke and Fire
Flight of Birds
Neighing of Horses
Tea Leaves and Coffee Grounds
Passages of Sacred Texts
Numbers
I Ching
Guessing
MODELS

David Orrell "The Future of Everything."
Interpretations: Pile Sort

Place the following food items in piles

<table>
<thead>
<tr>
<th>Broccoli</th>
<th>Carrots</th>
<th>Canned Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fresh Salmon</td>
<td>Arugula</td>
<td>Fennel</td>
</tr>
<tr>
<td>Spam</td>
<td>Ahi Tuna</td>
<td>Canned Posole</td>
</tr>
<tr>
<td>Niman Pork</td>
<td>Sea Bass</td>
<td>Canned Salmon</td>
</tr>
<tr>
<td>Veggie</td>
<td>Organic</td>
<td>Canned</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Broccoli</td>
<td>Fresh Salmon</td>
<td>Canned Beets</td>
</tr>
<tr>
<td>Arugula</td>
<td>Sea Bass</td>
<td>Spam</td>
</tr>
<tr>
<td>Carrots</td>
<td>Niman Pork</td>
<td>Canned Salmon</td>
</tr>
<tr>
<td>Fennel</td>
<td>Ahi Tuna</td>
<td>Canned Posole</td>
</tr>
</tbody>
</table>
## Airstream Sort

<table>
<thead>
<tr>
<th>Veggie</th>
<th>Meat/Fish</th>
<th>Weird?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broccoli</td>
<td>Fresh Salmon</td>
<td>Canned Posole</td>
</tr>
<tr>
<td>Fennel</td>
<td>Spam</td>
<td>Sea Bass</td>
</tr>
<tr>
<td>Carrots</td>
<td>Niman Pork</td>
<td>Arugula</td>
</tr>
<tr>
<td>Canned Beets</td>
<td>Canned Salmon</td>
<td>Ahi Tuna</td>
</tr>
</tbody>
</table>
Diversity Prediction Theorem

Crowd Error = Average Error - Diversity

\[(c - \theta)^2 = \frac{1}{n} \sum_{i=1}^{n} (s_i - \theta)^2 - \frac{1}{n} \sum_{i=1}^{n} (s_i - c)^2\]
Galton’s Steer

Crowd Error = Average Error – Diversity

0.6 = 2,956.0 - 2955.4
## 2005 NFL Draft

<table>
<thead>
<tr>
<th>Player</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>CROWD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alex Smith</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ronnie Brown</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>2.7</td>
</tr>
<tr>
<td>Braylon Edwards</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3.3</td>
</tr>
<tr>
<td>Cedric Benson</td>
<td>4</td>
<td>4</td>
<td>13</td>
<td>4</td>
<td>8</td>
<td>4</td>
<td>8</td>
<td>5.9</td>
</tr>
<tr>
<td>Carnell Williams</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>13</td>
<td>4</td>
<td>6.4</td>
</tr>
<tr>
<td>Adam Jones</td>
<td>16</td>
<td>9</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>8.1</td>
</tr>
</tbody>
</table>
2005 NFL Draft

<table>
<thead>
<tr>
<th>Predictor</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>CROWD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Squared Error</td>
<td>158</td>
<td>89</td>
<td>210</td>
<td>235</td>
<td>112</td>
<td>82</td>
<td>75</td>
<td>34.4</td>
</tr>
</tbody>
</table>
NFL Experts

<table>
<thead>
<tr>
<th>Predictor</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>CROWD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Squared Error</td>
<td>158</td>
<td>89</td>
<td>210</td>
<td>235</td>
<td>112</td>
<td>82</td>
<td>75</td>
<td>34.4</td>
</tr>
</tbody>
</table>

Average Error: 137.3
Diversity: 102.9
Crowd Error: 34.4
Problem Solving
Gunter Blobel: The exception
Perspectives

8
Fe
26
Iron
55.847
The Technocratic Ideal

Frederick Winslow Taylor
1856-1915

http://www.resourcesystemsconsulting
Simple: Shovel Landscape

- Efficiency
- Size
Caloric Landscape
Masticity Landscape
Ben and Jerry’s Perspective

chunk size

number of chunks
Consultant’s Perspective

caloric rank
Ben and Jerry’s Local Optima: Ave = 90

chunk size

number of chunks

chunk size

86

91

92

91
Consultant’s Local Optima: Ave = 80

caloric rank

78  92  76  74
Ben and Jerry’s Perspective

chunk size

number of chunks

chunk size

number of chunks
Consultant’s Perspective

caloric rank
Different Peaks
Heuristics
IQ Question:

Fill in the Blank: 1  2  3  5  _  13
$x_{i+2} - x_{i+1} = x$
IQ Question:

1  4  9  16  _  36
$x_i^2$

1  4  9  16  25  36
IQ Question:

1 2 6 _ 1806
\[
x_{i+1} - x_i = x_i^2
\]

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>6</td>
<td>42</td>
<td>1806</td>
<td></td>
</tr>
</tbody>
</table>

\[
2 - 1 = 1^2 \\
6 - 2 = 2^2 \\
42 - 6 = 6^2 \\
1806 - 42 = 42^2 \\
\]
$x_{i+1} - x_i = x_i^2$

A combination of the first two heuristics
Superadditivity

$1 + 1 = 3$
Network + Electrical Engineers

About This Map

Click on the links below to switch layers on and off.

EXISTING LINES
- 345-499 kV
- 500-899 kV
- 700-799 kV
- 1,000 kV (DC)

PROPOSED LINES
- New 765 kV
- AC-DC-AC Links

INTERCONNECTIONS
Major sectors of the U.S. electrical grid
- Eastern
- Western
- Texas (ERCOT)
A Test

• Create a bunch of agents with diverse perspectives and heuristics

• Rank them by their performance on a problem.

• Note: all of the agents must be “smart”
Experiment

Group 1: Best 20 agents

Group 2: Random 20 agents

Have each group work collectively - when one agent gets stuck at a point, another agent tries to find a further improvement. Group stops when no one can find a better solution.
The IQ View

Alpha Group

138  137  139
140  136  132

Diverse Group

121  84  111
75  135  9
The diverse group almost always outperforms the group of the best by a substantial margin.

See Lu Hong and Scott Page

*Proceedings of the National Academy of Sciences (2002)*
The Toolbox View

Alpha Group
- ABC
- ACD
- BDE
- BCD
- ADE

Diverse Group
- AHK
- FD
- AEG
- EZ
- BCD
- IL
What Must be True?

**Calculus Condition**: Problem solvers must all be smart—we must be able to list their local optima

**Diversity Condition**: Problem solvers must have diverse heuristics and perspectives

**Hard Problem Condition**: Problem itself must be difficult
Case: Netflix Prize
Outline

Netflix Prize: Background
Predictive Models
  Factor Models
Ensembles of Models
Ensembles of Teams
The Value of Diversity
Netflix Prize

November 2006, Netflix offers a prize of $1 million to anyone who can defeat their Cinematch recommender system by 10% of more.
Some Details

Netflix users rank movies from 1 to 5

Six years of data
Half million users
17,700 movies

Data divided into (training, testing)
Testing Data divided into (probe, quiz, test)
Interesting Asides

*Lost in Translation* and *The Royal Tenenbaums* had the highest variance

*Shawshank Redemption* had the highest rating

*Miss Congeniality* had the most ratings.
Singular Value Decomposition

Each movie represented by a vector:

\((p_1, p_2, p_3, p_4 \ldots p_n)\)

Each person represented by a vector:

\((q_1, q_2, q_3, q_4 \ldots q_n)\)

Rating: \(r_{ij} = m_i + a_j + p \cdot q\)

Training: choose \(p, q\) to minimize

\((\text{actual}_{ij} - r_{ij})^2 + c(\|p\|^2 + \|q\|^2)\)
BellKor’s Initial Models

Approximately 50 dimensions

Best Model: 6.8% improvement

Combination of Models: 8.4% improvement
Two Questions

Q1: Why more than one model?

Q2: Why do more work better than one?
Q1: Why More than one Model

This question has two answers.

A1: they used different variables

A2: their stochastic optimization technique got stuck in different places
Different Tuning Parameters and Initial Points Lead to Different Peaks on a Rugged Landscape
A2: Diversity Prediction Theorem

\[ \text{SqE}(c) = \text{SqE}(s) - \text{PDiv}(s) \]

\[ (c - \theta)^2 = \frac{1}{n} \sum_{i=1}^{n} (s_i - \theta)^2 - \frac{1}{n} \sum_{i=1}^{n} (s_i - c)^2 \]
BellKor’s Pragmatic Chaos

More is Better: Seven person team created combining top two teams. Now over 800 predictor sets (sets of variables).

Difficult be build a “grand” model but possible to build lots of “huge” models.
Ensemble Effects

Best Model 8.4%

Ensemble: 10.1%

Rules: Once someone breaks 10%, then the contest ends in 30 days.
Enter "The Ensemble"

23 teams from 30 countries who blended their predictive models who tried in the last moments to defeat BellKor’s Pragamatic Chaos
“The contest was almost a race to agglomerate as many teams as possible,” said David Weiss, a Ph.D. candidate in computer science at the University of Pennsylvania and a member of the Ensemble.

“The surprise was that the collaborative approach works so well, that trying all the algorithms, coding them up and putting them together far exceeded our expectations.”

New York Times 6/27/09
And The Winner is…

RMSE for The Ensemble: 0.856714
RMSE for Bellkor's Pragmatic Chaos: 0.856704

By the rules of the competition the scores are rounded to four decimal places so it was a tie.

However, BellKor’s Pragmatic Chaos submitted 20 minutes earlier so they won. (and they had the lower error)
Oh, by the way..

BellKor’s Pragmatic Chaos 10.06%
The Ensemble 10.06%
50/50 Blend 10.19%
Holedigging
Boosting
Collective Problem Solving
2. Create Oracles
3. Create Perspectives/Skills Spreadsheets

<table>
<thead>
<tr>
<th>name</th>
<th>engineer</th>
<th>sales</th>
<th>physics</th>
<th>statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>
4. Listen to Others But Avoid Group Think

"All those in favour say 'Aye'."

Learning

Average individual squared error of seven experts who made forecasts about the NBA draft from May 23rd through June 25th.

May 23rd : 213.17
May 30th : 86.33
June 13th: 114.5
June 18th : 139.67
June 22nd : 109
June 25th: 69.67
# Avoiding Group Think

<table>
<thead>
<tr>
<th>Date</th>
<th>Individual</th>
<th>Diversity</th>
<th>Collective Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>May 23(^{rd})</td>
<td>213.17</td>
<td>168.03</td>
<td>45.14</td>
</tr>
<tr>
<td>May 30(^{th})</td>
<td>86.33</td>
<td>81.41</td>
<td>28.57</td>
</tr>
<tr>
<td>June 13(^{th})</td>
<td>114.5</td>
<td>70.31</td>
<td>44.19</td>
</tr>
<tr>
<td>June 18(^{th})</td>
<td>139.67</td>
<td>113.3</td>
<td>26.34</td>
</tr>
<tr>
<td>June 22(^{nd})</td>
<td>109.0</td>
<td>84.0</td>
<td>25.0</td>
</tr>
<tr>
<td>June 25(^{th})</td>
<td>69.67</td>
<td>35.58</td>
<td>33.58</td>
</tr>
</tbody>
</table>
# Avoiding Group Think

<table>
<thead>
<tr>
<th>Date</th>
<th>Individual</th>
<th>Diversity</th>
<th>Collective Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>May 23rd</td>
<td>213.17</td>
<td>168.03</td>
<td>45.14</td>
</tr>
<tr>
<td>May 30th</td>
<td>86.33</td>
<td>81.41</td>
<td>28.57</td>
</tr>
<tr>
<td>June 13th</td>
<td>114.5</td>
<td>70.31</td>
<td>44.19</td>
</tr>
<tr>
<td>June 18th</td>
<td>139.67</td>
<td>113.3</td>
<td>26.34</td>
</tr>
<tr>
<td>June 22nd</td>
<td><strong>109.0</strong></td>
<td><strong>84.0</strong></td>
<td><strong>25.0</strong></td>
</tr>
<tr>
<td>June 25th</td>
<td>69.67</td>
<td>35.58</td>
<td>33.58</td>
</tr>
</tbody>
</table>
Encourage Dissent

If everyone agrees, then either the predictive task was easy and everyone has the correct forecast (in which case the meeting was a waste of time) or the task was challenging and everyone has the same, wrong forecast.
5. Technology Can Supplement Hierarchy
Goldcorp Challenge

March 6, 2000, Goldcorp offers $575k to participants who would help find gold at its Red Lake Mine in Ontario, Canada

110 targets identified, over 50% were new, over 80% were successful. Company value up from $100 Million to $9 Billion.
Prediction Markets
The Math Tells What’s Possible

\[ \overrightarrow{F} = m \overrightarrow{A} \]
The Parable of the Bike
The Need for Leadership

homogeneous

Cognitively diverse