Personality Research for NLP

Prof. Yair Neuman
Ben-Gurion University of the Negev
co-Director, BIRL, Univ. Toronto
Visiting Prof. Weizmann Institute of Science & Oxford University
yneuman@bgu.ac.il
https://bgu.academia.edu/YairNeuman

EMNLP 2015
Conference on Empirical Methods in Natural Language Processing
LISBON

A tutorial presented at the 2015 Conference on Empirical Methods on Natural Language Processing, September, Lisbon, Portugal
What is Personality and why is it Important for the NLP Community?
A psychological study once asked an interesting question:

Who divorce more?

Men who *try* to understand their wives?

or

Men who *don’t try* to understand their wives?
Those of you who are married can guess the answer...

Men who try to understand their wives divorce more than men who don’t try to understand their wives

The question is Why?
The answer that I gave elsewhere* is that Understanding other people is a tricky issue

• **Understanding** involves conscious theorization

• Human beings are better in applying unconscious **heuristics**

In sum, we want to understand other people but this is not an easy task

The lesson: Don’t try to understand your spouse, just adapt ...
However ...

Understanding other people is important than ever in our technological and consumer oriented society

In the past, and until quite recently in historical terms, it was more important to know **WHAT** you are (e.g. a Noblemen or a peasant) than **WHO** you are

Today the **WHO** you are is a central issue
Consumers’ Research

In consumers’ research “granular analytics” is an attempt to profile the client for high resolution tasks (e.g. marketing, advertising, prioritization)

Similarly to “Personalized Medicine”, we would like to find the best “match” for a specific individual
e-Health

In e-Health it is important to screen the mental state of patients for diagnosis, prevention and treatment.

Think for example about the importance of automatically screening for Depression.

The murderous pilot of German Wings.
Forensic Analytics

Can we screen for potential offenders such as pedophiles, school shooters or school bullies?
The Conclusion

Understanding other people is a must

Understanding other people is understanding their PERSONALITY
What is Personality?

Personality refers to the subject’s consistent patterns of thought, emotion, and behavior. Consistent = Stable across time and contexts.
The Borromean Rings of Personality

THOUGHTS

EMOTION

BEHAVIOR
Thoughts

Thoughts concern the **cognitive** aspect of personality

Mainly the **schemes** through which we represent our inner and outer world

**Beliefs** about self and others (I am .., They are ...)
Emotions

No need to explain but … a recent interesting neuro-imaging study suggests that all of our emotions can be mapped into four dimensions

1. Valence (whether the stimuli is positive or negative)
2. Arousal (the strength of the affect)
3. Social aspect (e.g. jealousy)
4. Sexual desire (e.g. lust)
Behavior

Behavior concerns the *actions* taken by the individual

Not necessarily actual actions but fantasies and intentions too
Let’s illustrate these dimensions of personality through the

“understanding” husbands

vs.

“non-understanding” husbands
Thoughts

The “understanding” husband: “I don’t understand my wife”

The “non-understanding husband”: “I think that my wife needs me”
Emotions

The “understanding” husband: “I’m frustrated”

The “non-understanding husband”: “I’m OK”
Behavior

The “understanding” husband: “I must get divorce”

The “non-understanding husband”: “I’m going to have a pint of beer”

or a bottle of Port Wine produced in Portugal …
**In sum**

Personality involves consistent patterns of thought, emotion and behavior.

Personality can be described in terms of:

**Types:** Discrete categories of personality or

**Traits:** Continuous dimensions of personality
How to scientifically identify a personality trait?

- Human languages use numerous words and phrases to describe people: Nerd, liberal, open-minded etc.
- Which of these tags can be used to describe a personality ‘trait’?
Not every descriptive term should be used as a personality tag (e.g. Liberal, Nerd, Funny)

However, it is difficult to decide ...
The Lexical Approach to Personality

- The naïve use of language is informative
- Through subjects’ self-reports examine whether certain linguistic tags cluster together
- If you identify clusters conclude that there are underlying personality factors explaining these clusters
Theories of Personality
The Five Factor Model of Personality (FFM) - The “Big Five”

- The dominant model in the current personality research


“The default model of personality structure”
“anyone’s personality can be characterized in terms of the five dimensions of the FFM”

The Five Factors of Personality

EXTRAVERSION

AGREEABLENESS

CONSCIOUSNESS

NEUROTICISM

OPENNESS
EXTRAVERSION

Being assertive, sociable, out-going, energetic, effective, and ambitious

The complementary dimension is

INTROVERSION
NEUROTICISM

Negative reaction to stress, anxious, insecure

Vs.

Calm and stable
AGREEABLENESS

Friendliness and cooperative behavior

Vs.

Selfishness and arrogance
CONSCIENTIOUSNESS: organized, ordered, responsible vs.
OPENNESS

Openness to experience, creativity and imagination

Vs.

Closeness and rigidity
Dogma under criticism

- The FFM is a dogma with many theoretical and empirical problems

For example

- The FFM is based on the relations between variables across individuals but personality lies within the individual

- So?
Molennar and Campbell (2009) analyzed data of 22 subjects who were measured on 90 consecutive days with equivalent versions of the Big Five questionnaire.

The results are surprising

- The analysis has shown that intra-individual variation doesn’t explain the repeated measurement scores!
- It means that the correlation you may find at the group level of analysis is invalid at the individual level of analysis
• For example, you may find a negative correlation between being ‘happy’ and ‘sad’ at the group level of analysis but when you analyze the subject’s scores, as gained through repeated measurements, the correlation vanishes …

• No consistency, no personality …
A simpler explanation?

- Neuroticism was found* to be strongly correlated with FEAR, SADNESS and ANGER
- So … Maybe Neuroticism is just a fancy name for negative emotions?

The dimensions are highly associated, for instance

- 60% of the Extraverts are Agreeable
- 62% of those who are not Neurotic are Agreeable and so on …

In sum …

- The FFM is the most popular model of personality, specifically in computational personality
- Theoretically simple
- Results are “guaranteed” (anything is associated with Neuroticism …)
- Extremely limited
The Psychodynamic Approach: Beyond the Straw-Man of Dr. Freud
- Personality as the adaptive ways of coping with the “exigencies of life”
- Variability of personality traits is therefore variability in mental coping strategies
- Personality types are presented in terms of the (1) cognitive (2) affective and (3) defensive functioning
Defensive Functioning

- The mechanisms we use to defend ourselves from threatening thoughts and emotions
- For example, **splitting**
- Positive and negative qualities of the self and others are separated
The Splitting Mind

[Image of a devil and an angel]
Mom didn’t give me a candy
Mom is BAD

Mom gave me a hug
Mom is GOOD

Too complex!!!

Maybe there are two different mothers!
Splitting explains why through History we Find Two Orthogonal Kind of Women!
And ...
Identifying Splitting

- Highly relevant for diagnosis
- Identification of potential offenders (e.g. civilian mass murderers, political extremists etc.)

The Spectrum of Personality Dimensions

- Personality dimensions range from Adaptive, flexible and normal to Non-adaptive, rigid and pathological patterns.
The Obsessive Personality
Positive Aspects

- Perfectionists, hard-workers, believe in work ethics, task-oriented and put their emotions aside when the work has to be done
- Such a prototypical personality may be a highly successful surgeon or engineer
The Obsessive Personality
Negative Aspects

- Nothing is ever good enough and therefore no task can be finished
- The hard-working person might be a workaholic who is "caged" in his office
- The emotional constriction might be experienced as the loss of emotions that drains the joy of life and a basic sense of living
Personality Themes

- Personality is organized around “themes” or conflicts

The Paranoid Personality
• Organizing theme: Attacking/Being attacked by humiliating others
• Central affect: Fear, rage
• Thoughts of others: The world is full of potential attackers and users
• Central ways of defense: Projection. The aggression that exists in his mind is attributed to others in the world
A Famous Paranoid
Suspicions

- In fact, the major conflict of the paranoid is the one of trust/suspicion.
- The paranoid invests energy in trust issues.
- Believes no one.
- On the other hand, fully and uncritically trusts and adopts ungrounded Beliefs.
- Conspiracy theories.
For example

- Someone told me that his mother-in-law calls him regularly to “inform” him about “new scientific” inventions, trusting all possible rubbish in the media.
Aaron, It is amazing! The Daily Mirror just reported that Kim Kardashian invented a new drug for healing Cancer!
My diagnosis: A conflict over trust

- I’ve asked him if his mother-in-law is paradoxically suspicious of well-grounded information …
- He was shocked when I explained that I don’t really know his mother-in-law …
And don’t ask me to believe that the Americans really landed on the Moon!
Personality Prototypes

- The Shelder-Westen Assessment Procedure (SWAP) recently identified two main personality clusters

The Internalizing Spectrum

Depressed, anxious, emotionally inhibited and socially avoidant, blame themselves, self-critical

- Depressive
- Anxious-Avoidant
- Dependent-Victimized
- Schizoid-Schizotypal
Depressive Personality

- Prone to feelings of sadness and inadequacy
- Self-critical/self-punitive
- Concerns about abandonment and loss
A Famous Depressed Author
Franz Kafka
Anxious–Avoidant Personality

- Socially anxious and avoidant
- Manage anxiety in a way that limit and constrict their lives
A Famous Avoidant Singer
Michael Jackson
Dependent-Victimized Personality

- Dependent and afraid of being alone
- Insufficient concern for their well-being
- Difficulty in expressing anger
The Schizoid-Schizotypal

- Peculiarities in interpersonal relationships and thought processes
- Appear to have little need for human company
- Seem detached or indifferent
- Lack social skills and look awkward
- Odd appearance
- Feel life has no meaning
Willy Wonka from Charlie and the Chocolate Factory
• We recently found that Schizoids function as “Firewalls” for the spread of rumors on Twitter
• When receiving a rumor they are significantly less inclined to re-Tweet it
The Externalizing Spectrum

Angry and hostile, suspicious of others, conflicted about authority, prone to getting into power struggles, lack empathy, feel privileged and entitled, blame others, lack psychological insight, feel mistreated

- Antisocial-Psychopathic
- Paranoid
- Narcissist
The Antisocial-Psychopathic

- Exploit others
- Experience little remorse and empathy
- Manipulative
- Lack fear
- Enjoy playing the role of the predator
Personality Analysis and Deterrence
Khaled Mashal – The Leader of Hamas
Narcissistic Personality

- Grandiose self
- Dismissive of others
- Vulnerability beneath a grandiose facade
Muhammad Ali – The Boxer

“I am the Greatest. I said that even before I knew I was.”
Borderline-Dysregulated Spectrum

- Impaired ability to regulate emotions
- Unstable perceptions of self and others
- Intense and chaotic relationships
- Impulsive to include self-destructive impulses
Neurotic Styles

- Obsessional personality
  Overly rational, emotionally constricted and rigid, critical of themselves and others

- Hysteric-Histrionic personality
  Emotionally dramatic and theatrical, sexually provocative, idealizing others, paradoxically both intensely and superficially attached
<table>
<thead>
<tr>
<th>Personality</th>
<th>Tension</th>
<th>Affect</th>
<th>Belief Self</th>
<th>Belief Others</th>
<th>Defense</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schizoid</td>
<td>Fear of/longing for closeness</td>
<td>Pain</td>
<td>Dependency and love are dangerous</td>
<td>Social interactions are painful</td>
<td>withdrawal</td>
</tr>
<tr>
<td>Paranoid</td>
<td>Attacking/being attacked</td>
<td>Fear/rage</td>
<td>Aggression is dangerous</td>
<td>World is full of offenders</td>
<td>projection</td>
</tr>
<tr>
<td>Psychopathic/Antisocial</td>
<td>Manipulating/being manipulated</td>
<td>Rage/envy</td>
<td>I can make anything happened</td>
<td>Everyone is manipulative</td>
<td>Omnipotent control</td>
</tr>
</tbody>
</table>
In sum

- The psychodynamic approach is theoretically grounded and rich in insights
- Mostly applied to clinical cases
- Difficult to apply in practice specifically in the context of NLP
The Cognitive Behavioral Approach

We have genetically determined strategies, or more accurately schemes, that aim to support our survival and reproduction.

Personality is about our schemes.

Beliefs about self, others, future

- Thoughts result from interpretation/information processing
- Thoughts lead to emotions & behavior
- Overdeveloped vs. underdeveloped schemes of interpretation
## Over vs. Under Developed Schemes

<table>
<thead>
<tr>
<th>Personality</th>
<th>Overdeveloped</th>
<th>Underdeveloped</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obsessive</td>
<td>Control</td>
<td>Spontaneity</td>
</tr>
<tr>
<td>Dependent</td>
<td>Help seeking</td>
<td>Self sufficiency</td>
</tr>
<tr>
<td>Paranoid</td>
<td>Mistrust</td>
<td>Trust</td>
</tr>
<tr>
<td>Schizoid</td>
<td>Withdrawal</td>
<td>Intimacy</td>
</tr>
<tr>
<td>Narcissistic</td>
<td>Self-aggrandizement</td>
<td>Empathy</td>
</tr>
</tbody>
</table>
Beliefs → Behavior

<table>
<thead>
<tr>
<th>Personality</th>
<th>Self</th>
<th>Others</th>
<th>Assumption</th>
<th>Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paranoid</td>
<td>Vulnerable</td>
<td>Vicious</td>
<td>If I trust then I may be attacked</td>
<td>Don’t trust</td>
</tr>
<tr>
<td>Obsessive</td>
<td>My world can get out of control</td>
<td>Irresponsible</td>
<td>If I’m out of control the world will fall into chaos</td>
<td>Control as much as possible, rigidity</td>
</tr>
<tr>
<td>Narcissistic</td>
<td>Inferior</td>
<td>Superior</td>
<td>If others conceive me as non-superior than I’m inferior</td>
<td>Show your superiority</td>
</tr>
</tbody>
</table>
If you want to understand others analyze their beliefs, assumptions and behavior.
How to identify a Narcissist?
• If someone is saying:

“\textcolor{red}{I am a} \textcolor{blue}{brilliant} \textcolor{green}{person who graduated from} \textcolor{red}{Harvard university} \textcolor{green}{not like my brothers who studied in a} \textcolor{red}{community college}”
The association between the “I” and “Brilliant” is indicative of the subject’s belief about the self.
A Practical Application

Personality sensitive advertisements

• For a narcissistic personality paraphrase the advertisement so that it includes words involving high status:

Harvard, Manhattan, Cartier, Ferrari, Remy Martin Cognac, etc.
The Affective-Neuroscience Approach

- Basically seeks to ground the human affective system in the context of mammalian brains

Core Emotions

- SEEKING, FEAR, RAGE, LUST, CARE, PANIC, and PLAY

Adapted to Personality

- ANPS – The Affective Neuroscience Personality Scales


The Positive Emotions
Playfulness

- Joy, social gaming, laugher, humor

Having fun vs. being serious, playing games with physical contact, humor, and laughter, and being generally happy and joyful.

“I am known as one who keeps work fun”
Seeking

- Appetitive motivation system that involves a goal-directed behavior

Feeling curious, feeling like exploring, striving for solutions to problems and puzzles, positively anticipating new experiences, and a sense of being able to accomplish almost anything

“Almost any little problem or puzzle stimulates my interest”
Caring

Nurturing, being drawn to young children and pets, feeling softhearted toward animals and people in need, feeling empathy, liking to care for the sick, feeling affection for and liking to care for others, as well as liking to be needed by others

“I often feel a strong need to take care of others”
The Negative Emotions
Fear

- Flight, escape, avoiding harm

Feelings of anxiety, feeling tense, worrying, struggling with decisions, ruminating about past decisions and statements, losing sleep, and not typically being courageous

“I often cannot fall right to sleep because something is troubling me”
Anger (Rage)

- Anger, aroused by thwarting and frustration, affective attack

Feeling hotheaded, being easily irritated and frustrated, experiencing frustration leading to anger, expressing anger verbally or physically, and remaining angry for long periods

“When I get angry, I often feel like swearing”
Sadness (Panic)

- A separation distress system, attachment to significant others

Feeling lonely, crying frequently, thinking about loved ones and past relationships, and feeling distress when not with loved ones

“I often have the feeling that I am going to cry”
Significant correlations between the ANPS and the Big Five

<table>
<thead>
<tr>
<th></th>
<th>E</th>
<th>A</th>
<th>C</th>
<th>-N</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAY</td>
<td>.46</td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEEK</td>
<td></td>
<td></td>
<td>.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CARE</td>
<td>.25</td>
<td>.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEAR</td>
<td>-.19</td>
<td>-.17</td>
<td>-.24</td>
<td>-.75</td>
<td></td>
</tr>
<tr>
<td>ANGER</td>
<td></td>
<td>-.48</td>
<td>-.30</td>
<td>-.65</td>
<td></td>
</tr>
<tr>
<td>SADNESS</td>
<td>.15</td>
<td>.26</td>
<td></td>
<td></td>
<td>.17</td>
</tr>
</tbody>
</table>
In sum

- Grounds personality in basic emotion systems
- Can it represent the complexity of HUMAN personality?
- What does it mean to be a Narcissist according to the Affective Neuroscience Approach?
The requisite variety of personality theories: Why should we have more ideas in our tool kit
• Higher level features may be highly relevant for sensitive and complex tasks
Bogdanova, Rosso and Solorio (2014) found that high-level feature were much more effective than lower-level features (e.g. n-grams) in differentiating between pedophilia chats and normative cybersex chats.

• Enriching our personality dimensions means enriching the high-level features that we may use for automatic personality analysis
NLP and Computational Personality
• Automatic analysis of the subject’s personality based on his texts
• You can use non-textual features for the analysis (e.g. voice) but I focus on texts only
• Texts, whether written or spoken, are the richest source of information
Workshop on Computational Personality Recognition: Shard Task (2013)

- The data: A corpus of 2468 stream-of-consciousness essays that was labeled with personality classes of the FFM.
- Each student got a categorical score on each of the FFM. For instance you can be a Neurotic or non-Neurotic.
Percentage of Subjects in each of the Personality Dimensions

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>E</th>
<th>O</th>
<th>A</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50</td>
<td>48</td>
<td>49</td>
<td>47</td>
<td>49</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>52</td>
<td>51</td>
<td>53</td>
<td>51</td>
</tr>
</tbody>
</table>
A Straightforward ML approach

- Get a tagged corpus
- Identify relevant features
- Classify

- The bread and butter of NLP researchers
Features

- N-grams
- LIWC
- Sentiwordnet
- Emotions

“top-down approaches, based on lexical resources (including the ones for sentiment analysis) and social info, in general seem to help personality recognition more than bottom-up approaches” (Celli et al. 2013)
Fine emotions even work better than general coarse categories

<table>
<thead>
<tr>
<th>E</th>
<th>N</th>
<th>A</th>
<th>C</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>possessive</td>
<td>guilt</td>
<td>happy</td>
<td>excited</td>
<td>anxious</td>
</tr>
<tr>
<td>apart</td>
<td>eager</td>
<td>anger</td>
<td>apprehensive</td>
<td>delighted</td>
</tr>
<tr>
<td>happy</td>
<td>interested</td>
<td>homesick</td>
<td>homesick</td>
<td>blah</td>
</tr>
</tbody>
</table>

Important emotions “loaded” on each of the personality dimensions

• There are limits in gaining access to a tagged corpus
• How representative is the corpus used for the First Comp. Pers. Workshop?
• A corpus of students …
• A sample of students is a non-representative sample
A Vectorial Semantics Approach to Personality Analysis

1. Identify words that are the best representatives of a certain personality trait
2. Represent them as a vector
3. Measure the distance between your personality vector and the target text

For instance

Extraversion: dominant, assertive, authoritarian, forceful, assured, confident, firm, persistent

Depressive: sad, depressed, hopeless, gloomy, fatalistic
A Top-Down Approach

- Pros: Works extremely well, currently the state of the art results

F1 scores of Neuman and Cohen (NC) vs. Mohammad and Kiritchenko (MK)

<table>
<thead>
<tr>
<th></th>
<th>EXT</th>
<th>NEU</th>
<th>AGR</th>
<th>OPN</th>
<th>CON</th>
</tr>
</thead>
<tbody>
<tr>
<td>MK</td>
<td>56.28</td>
<td>58.25</td>
<td>54.20</td>
<td>60.57</td>
<td>56.56</td>
</tr>
<tr>
<td>NC</td>
<td>66.30</td>
<td>64.60</td>
<td>68.91</td>
<td>61.55</td>
<td>60.74</td>
</tr>
</tbody>
</table>
Cons

As a Top-Down approach it demands the expert’s knowledge for identifying the relevant words/vectors

Can we exhaust the complexity of human personality by focusing only on “words” as units of analysis?

Personality is probably expressed better as a discursive dynamics that involves the pragmatics of human communication
From Words to Discourse
From Semantics to Pragmatics

- The discursive level: Propositions, rather than words, as the target of analysis

- Pragmatics: The use of language in human communication as determined by the conditions of society*

• Understanding the “Pragmatics of Personality” is a must
• Specifically if we would like to understand personalities beyond their limited Anglo-Saxon cultural context
This is Probably the Next Phase in Automatic Personality Analysis
In psychology, there has been an intensive work in developing valid psychological questionnaires for measuring psychological dimensions
We can use an NLP approach in which the quest/inventory is allegedly administrated to the subject or more accurately to his written text.
"I am often nervous, fearful, and anxious, and I worry that something might go wrong"
We can use a TE engine and measure the degree in which the hypothesis, or several hypotheses derived from the personality items, is entailed from the subject’s written essay.
text: *I am very sad.*

hypothesis: I am often nervous, fearful, and anxious, and I worry that something might go wrong

There are three possible answers to this text-hypothesis pair:

1. text entails hypothesis
2. text contradicts hypothesis
3. text does not entail or contradict
Given a text written by subject X, let each sentence in the text be $X_i$. For each $X_i$, construct a text-hypothesis pair:

text: $X_i$

hypothesis: $H_i$, 

where each $H_i$ is an item taken from the psychological questionnaire.
Doesn’t work well...

- But maybe a simpler method
Convert personality items into lexico-semantic patterns

Item: "I am often nervous, fearful, and anxious"

has been converted into the pattern:

nsubj(nervous, I)
nsubj(fearful, I)
nsubj(anxious, I)
• Parse the essay and find sentences with first person pronoun
• Identify matching patterns for those that have been extracted from the psychological questionnaire
• Measure the similarity between the essay and the predefined patterns
For example

- Target sentence: I am sad

- Personality item for measuring Neuroticism:

  "I am often nervous, fearful, and anxious"
I AM SAD

IF nsubj (X, I) AND IF X = JJ
THEN SIM(X, nervous/fearful/anxious)

IF advmod (X, RB = 'seldom' or 'rarely' or 'barely' or 'hardly' or 'infrequently' or 'never' or 'scarcely' or 'almost never' or 'not often')

OR

IF neg(X)
THEN SIM = SIM*(-1)
We tested this methodology on the identification of the FFM

Few pre-defined patterns based on the “Short Five” (S5) questionnaire (Muck et al., 2007) (e.g. only 8 patterns for NEU)

Results

On average, the F1 score gained in this study (63.58) isn't significantly different from the one gained by Neuman and Cohen (2014) (i.e., 64.42), but is higher that the average F1 score gained by Mohammad and Kiritchenko (2013) (i.e., 57.17)
Practical Applications
Homeland Security: How to Identify a Murderer in a Haystack

School Shooters

Seung-Hui Cho murdered 23 students and faculty members at Virginia Tech on April 16, 2007
• School shooters receive extensive media coverage and create social anxiety
• There is no consistent diagnosis of school shooters
• Can we profile school shooters?
• Can we use the profile for future screening?
• We selected six texts written by school shooters

• For gaining comparative insights, we used the Blogs Authorship Corpus (Schler et al., 2006) and selected blogs written by males from the age of 15 to 25, ages approximately overlapping those of the school shooters

• Overall, we analyzed the blogs written by 6056 subjects
• We have analyzed only three parts of speech categories: nouns, verbs, and adjective

• From each text, we selected the ten most frequent nouns, the ten most frequent verbs, and the ten most frequent adjectives

• Overall, we used 30 words to represent each text as a vector
We measured the semantic similarity between each of the texts and word vectors representing four personality disorder traits:

1. Paranoid Personality Disorder (PPD),
2. Narcissistic Personality Disorder (NPD)
3. Schizotypal Personality Disorder (ScPD) and
4. Depressivity (DEP)
In addition, factors of REVENGE

1. Hopeless: hopeless, desperate
2. Lonely: lonely, lonesome
3. Helpless: helpless, defenseless
4. Pain: pain, misery, agony
5. Revengeful: revengeful, vengeful, vindictive
6. Chaotic: chaotic, disordered
7. Unsafe: unsafe, insecure
8. Abandoned: abandoned, deserted
9. Humiliated: humiliated, shamed
For measuring similarity, we used Peter Turney’s matrix

Following the automatic identification of sexual predators (Inches & Crestani, 2012) can we produce a ranked list of suspects to prioritize the investigation?

Three Statistical Models

Binary Logistic Regression Analysis

Tree Classification with CHAID and ten-fold cross-validation procedure and

K Nearest Neighbors Analysis with ten-fold cross-validation
Averaging the ranks produced by the three classifiers, we gain the best results, which are identifying all the shooters’ texts among the top 210 ranked texts: approximately 3% of our corpus

Enormous reduction in work load for the human agent
Screening for Potential Offenders

- Ethical considerations due to the high rate of false positives
- Results can be improved by the fusion of several information sources:
  1. Medical records: Psychiatry illness
  3. Criminal records: Previous convictions?
Military Intelligence: Profiling Political Leaders

• Personality profiling

• The "impressionist" and "informal" approach used by Jerry Post (CIA)

• The “formal” approach uses informants' reports of others through the use of standard inventories (The profiling of Kim Jong-il)
Personality and Text-in-Context

The challenge:

- Identify themes that emerge from the text in a given socio-political context as well as personality features.
- The way in which these themes, that may have been preformed by the leader and probably by his close circle, intermingle with the leader's personality to produce a synergetic output.
Morsi’s Case

- The speech given by the new elected Egyptian President Mohammed Morsi
- A leader of the Islamic Brothers
- September 26, 2012, the General Assembly of the UN
Surface Themes
Three Major Themes

1. The **Palestinian – Israeli conflict** and Egypt’s support of the Palestinian side
2. Establishing Egypt as an Arabic – African **leader** of the developing world
3. Egypt’s efforts on helping to end the violence in **Syria**
Motifs Analysis - Results

- We applied MOTIFS ANALYSIS to the semantic network of the speech.

- By identifying the words populating most of the motifs' edges, three words emerge: **Rights**, **Principles** and **Vision**.

- The hypothesis: Morsi’s speech is indicative of Obsessive Compulsive Personality features.
Obsessive Compulsive Personality

This type of personality seeks opportunity to prove himself as selflessly committed to the "greater cause”

Rigid and dogmatic personality, occupied with details, and presents an over conscientious attitude toward matters of morality and values
Testing the Hypothesis

• The Obsessive vector: Rigid, stubborn, over-conscientious, inflexible, organized, perfectionist …

• The Narcissistic vector: special, unique, selfish, envious, arrogant, contemptuous, brilliant …
Results

The speech was much closer to the Obsessive vector (0.09) than to the Narcissistic vector (0.02).

It was found that the speech was mostly similar to the Bureaucratic and the Purtian subtypes.
Conclusions

- President Morsi was a Bureaucratic-Puritan leader
- Don’t expect him to lead a change …
- The American enthusiasm was too early …
- In retrospect we know that we were right
Customers' Service: Who is the Unhappy Customer
• Can we better understand who is a non-satisfied customer?
• The **** dataset
• 335,022 Reviews
• 70,817 users
• We categorized the reviews into satisfied (3/4 stars) and non satisfied customers (1/2 starts)
Most reviews were very positive (70% vs. 30%)

We profiled the customers by analyzing the reviews of each customer
Correlations between Rating and Personality Dimensions

- Neuroticism = -0.215
- Angry = -0.185
- Etc.

- We built an index from these scores
- The sum of the z-scores
- Titled “Neurotic”
<table>
<thead>
<tr>
<th>Neurotic?</th>
<th>NO</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>A satisfied customer?</td>
<td>NO</td>
<td>34%</td>
</tr>
<tr>
<td></td>
<td>YES</td>
<td>58%</td>
</tr>
</tbody>
</table>
• The probability of being an unsatisfied costumer is 0.32
• The probability of being an unsatisfied costumer given that one is Neurotic is 0.43
Implications for Customers’ Service Priorization

- Some customers are not satisfied as a result of the service/purchase
- Some customers are not satisfied because their PERSONALITY cannot be satisfied
- We found two types of unsatisfied customers: The NEUROTICS and the PERFECTIONISTS (i.e. Obsessive)
The Rorschach Test
It Looks Like Darth Vader from Star Wars!
BUT DON’T POINT TO THIS SIMILARITY

The diagnosis of those who see monsters in the ink spots isn’t positive …
Human Personality is like a Rorschach Ink Spot
Sometimes we see our own projections rather than a real structure …

- Automatic personality analysis
- The challenge of understanding other people
- From naïve impressions to scientific analysis
- A pragmatic motivation – real world applications
Thank you for Attending this Tutorial