



**A P G**

# **Automatic Persona Generation (APG)**

**News & Social Media Analytics Team**

**Social Computing Group**

**Qatar Computing Research Institute**

Hamad Bin Khalifa University

# The APG Team

**Prof. Jim Jansen**  
Leader

**Dr. Joni Salminen**  
Post-doctoral researcher

**Dr. Haewoon Kwak**  
Scientist

**Dr. Jisun An**  
Scientist

**Soon-Gyo Jung**  
Engineer



# What is a Persona?

- **A ‘persona’ is a fictitious person describing an important user group (= a market segment with a name) (Cooper, 1999).**
- Simplifies numerical data into an easy format: a person.
- Personas help 1) communicate user information inside an organization, so that 2) decisions can be made keeping the end user in mind (Nielsen, 2013).
- Widely used in companies (Microsoft, HubSpot, Google) among software developers, designers and marketers (Pruitt & Grudin, 2003; Chapman et al., 2015).

# What is a Persona?

- A 'persona' is a fictitious person describing an important user group (= a market segment with a name) (Cooper, 1999).
- **Simplifies numerical data into an easy format: a person.**
- Personas help 1) communicate user information inside an organization, so that 2) decisions can be made keeping the end user in mind (Nielsen, 2013).
- Widely used in companies (Microsoft, HubSpot, Google) among software developers, designers and marketers (Pruitt & Grudin, 2003; Chapman et al., 2015).

# What is a Persona?

- A ‘persona’ is a fictitious person describing an important user group (= a market segment with a name) (Cooper, 1999).
- Simplifies numerical data into an easy format: a person.
- **Personas help 1) communicate user information inside an organization, so that 2) decisions can be made keeping the end user in mind (Nielsen, 2013).**
- Widely used in companies (Microsoft, HubSpot, Google) among software developers, designers and marketers (Pruitt & Grudin, 2003; Chapman et al., 2015).

# What is a Persona?


- A 'persona' is a fictitious person describing an important user group (= a market segment with a name) (Cooper, 1999).
- Simplifies numerical data into an easy format: a person.
- Personas help 1) communicate user information inside an organization, so that 2) decisions can be made keeping the end user in mind (Nielsen, 2013).
- **Widely used in companies (Microsoft, HubSpot, Google) among software developers, designers and marketers (Pruitt & Grudin, 2003; Chapman et al., 2015).**

# An example

## Phil

Male | 20 | United States

0 Favorites 0 Comments 0 Shares



*Phil is a 20 year old male living in United States. He likes to read about International Affairs, US Politics, and Health and Wellness on his Mobile. He usually watches about 1.6 minutes of video.*

<b>Name</b>	Phil
<b>Gender</b>	Male
<b>Age</b>	20
<b>Country</b>	United States
<b>Education Level</b>	College grad 28%
<b>Marital Status</b>	Not_married 93%

### Topics Of Interest

More Interested

- International Affairs
- US Politics
- Health and Wellness

Less Interested

- Elections
- News
- US Social Issues

### Quotes

"I was in Charleston when the tragedy occurred, puts things into perspective..."

"Getting recruits at a young age is the key to taking away their future."


"Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"

### Most Viewed Contents

- Health and Wellness** Is The South Racist? We Asked South Carolinians  
30. June 2015
- US Politics** 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump  
28. October 2015
- Health and Wellness** Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+  
12. September 2017
- Police** Did These Swedish Cops School The NYPD?  
24. April 2015
- Turkey** Confederate Flag Debate - We Ask People In South Carolina What They Think  
23. June 2015

### Audience Size

7,600,000 people

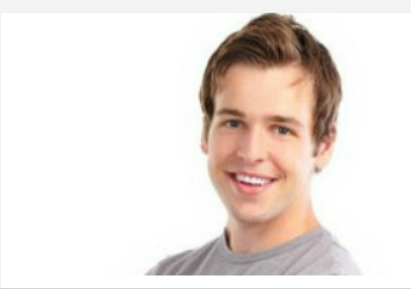


1/10, 28. February 2018, YouTube

# An example

## Phil

Male | 20 | United States



*Phil is a 20 year old male living in United States. He likes to read about International Affairs, US Politics, and Health and Wellness on his Mobile. He usually watches about 1.6 minutes of video.*

<b>Name</b>	Phil
<b>Gender</b>	Male
<b>Age</b>	20
<b>Country</b>	United States
<b>Education Level</b>	College grad 28%
<b>Marital Status</b>	Not_married 93%

### Topics Of Interest

More Interested

- International Affairs
- US Politics
- Health and Wellness

Less Interested

- Elections
- News
- US Social Issues

### Quotes

"I was in Charleston when the tragedy occurred, puts things into perspective..."

"Getting recruits at a young age is the key to taking away their future."

"Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"

### Most Viewed Contents

- Health and Wellness** Is The South Racist? We Asked South Carolinians  
30. June 2015
- US Politics** 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump  
28. October 2015
- Health and Wellness** Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+  
12. September 2017
- Police** Did These Swedish Cops School The NYPD?  
24. April 2015
- Turkey** Confederate Flag Debate - We Ask People In South Carolina What They Think  
23. June 2015

### Audience Size

7,600,000 people

0 Favorites 0 Comments 0 Shares

1/10, 28. February 2018, YouTube




# An example

## Phil

Male | 20 | United States

0 Favorites 0 Comments 0 Shares



*Phil is a 20 year old male living in United States. He likes to read about International Affairs, US Politics, and Health and Wellness on his Mobile. He usually watches about 1.6 minutes of video.*

<b>Name</b>	Phil
<b>Gender</b>	Male
<b>Age</b>	20
<b>Country</b>	United States
<b>Education Level</b>	College grad 28%
<b>Marital Status</b>	Not_married 93%

### Topics Of Interest

More Interested

- International Affairs
- US Politics
- Health and Wellness

Less Interested

- Elections
- News
- US Social Issues

### Quotes

"I was in Charleston when the tragedy occurred, puts things into perspective..."

"Getting recruits at a young age is the key to taking away their future."

"Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"

### Most Viewed Contents

- Health and Wellness** Is The South Racist? We Asked South Carolinians  
30. June 2015
- US Politics** 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump  
28. October 2015
- Health and Wellness** Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+  
12. September 2017
- Police** Did These Swedish Cops School The NYPD?  
24. April 2015
- Turkey** Confederate Flag Debate - We Ask People In South Carolina What They Think  
23. June 2015

### Audience Size

7,600,000 people


1/10, 28. February 2018, YouTube

# An example

## Phil

Male | 20 | United States

0 Favorites 0 Comments 0 Shares



*Phil is a 20 year old male living in United States. He likes to read about International Affairs, US Politics, and Health and Wellness on his Mobile. He usually watches about 1.6 minutes of video.*

Name	Phil
Gender	Male
Age	20
Country	United States

**Education Level** ?

College grad 28%

**Marital Status** ?

Not\_married 93%

### Topics Of Interest

More Interested

- International Affairs
- US Politics
- Health and Wellness

Less Interested

- Elections
- News
- US Social Issues

### Quotes

"I was in Charleston when the tragedy occurred, puts things into perspective..."

"Getting recruits at a young age is the key to taking away their future."


"Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"

### Most Viewed Contents

- Health and Wellness** Is The South Racist? We Asked South Carolinians  
30. June 2015
- US Politics** 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump  
28. October 2015
- Health and Wellness** Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+  
12. September 2017
- Police** Did These Swedish Cops School The NYPD?  
24. April 2015
- Turkey** Confederate Flag Debate - We Ask People In South Carolina What They Think  
23. June 2015

### Audience Size

7,600,000 people




1/10, 28. February 2018, YouTube

# An example

## Phil

Male | 20 | United States

0 Favorites 0 Comments 0 Shares



*Phil is a 20 year old male living in United States. He likes to read about International Affairs, US Politics, and Health and Wellness on his Mobile. He usually watches about 1.6 minutes of video.*

<b>Name</b>	Phil
<b>Gender</b>	Male
<b>Age</b>	20
<b>Country</b>	United States
<b>Education Level</b>	College grad 28%
<b>Marital Status</b>	Not_married 93%

### Topics Of Interest

More Interested

- International Affairs
- US Politics
- Health and Wellness

Less Interested

- Elections
- News
- US Social Issues

### Quotes

"I was in Charleston when the tragedy occurred, puts things into perspective..."

"Getting recruits at a young age is the key to taking away their future."

"Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"

### Most Viewed Contents

- Health and Wellness** Is The South Racist? We Asked South Carolinians  
30. June 2015
- US Politics** 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump  
28. October 2015
- Health and Wellness** Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+  
12. September 2017
- Police** Did These Swedish Cops School The NYPD?  
24. April 2015
- Turkey** Confederate Flag Debate - We Ask People In South Carolina What They Think  
23. June 2015

### Audience Size

7,600,000 people


1/10, 28. February 2018, YouTube

# An example

## Phil

Male | 20 | United States

0 Favorites 0 Comments 0 Shares



*Phil is a 20 year old male living in United States. He likes to read about International Affairs, US Politics, and Health and Wellness on his Mobile. He usually watches about 1.6 minutes of video.*

<b>Name</b>	Phil
<b>Gender</b>	Male
<b>Age</b>	20
<b>Country</b>	United States
<b>Education Level</b>	College grad 28%
<b>Marital Status</b>	Not_married 93%

### Topics Of Interest

More Interested

- International Affairs
- US Politics
- Health and Wellness

Less Interested

- Elections
- News
- US Social Issues

### Quotes

"I was in Charleston when the tragedy occurred, puts things into perspective..."

"Getting recruits at a young age is the key to taking away their future."

"Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"

### Most Viewed Contents

- Health and Wellness** Is The South Racist? We Asked South Carolinians  
30. June 2015
- US Politics** 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump  
28. October 2015
- Health and Wellness** Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+  
12. September 2017
- Police** Did These Swedish Cops School The NYPD?  
24. April 2015
- Turkey** Confederate Flag Debate - We Ask People In South Carolina What They Think  
23. June 2015

### Audience Size

7,600,000 people


1/10, 28. February 2018, YouTube

# An example

## Phil

Male | 20 | United States

0 Favorites 0 Comments 0 Shares



*Phil is a 20 year old male living in United States. He likes to read about International Affairs, US Politics, and Health and Wellness on his Mobile. He usually watches about 1.6 minutes of video.*

Name	Phil
Gender	Male
Age	20
Country	United States
Education Level	College grad 28%
Marital Status	Not_married 93%

### Topics Of Interest

More Interested

- International Affairs
- US Politics
- Health and Wellness

Less Interested

- Elections
- News
- US Social Issues

### Quotes

"I was in Charleston when the tragedy occurred, puts things into perspective..."

"Getting recruits at a young age is the key to taking away their future."


"Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"

### Most Viewed Contents

- Health and Wellness** Is The South Racist? We Asked South Carolinians  
30. June 2015
- US Politics** 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump  
28. October 2015
- Health and Wellness** Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+  
12. September 2017
- Police** Did These Swedish Cops School The NYPD?  
24. April 2015
- Turkey** Confederate Flag Debate - We Ask People In South Carolina What They Think  
23. June 2015

### Audience Size

7,600,000 people




1/10, 28. February 2018, YouTube

# An example

## Phil

Male | 20 | United States

0 Favorites 0 Comments 0 Shares



*Phil is a 20 year old male living in United States. He likes to read about International Affairs, US Politics, and Health and Wellness on his Mobile. He usually watches about 1.6 minutes of video.*

<b>Name</b>	Phil
<b>Gender</b>	Male
<b>Age</b>	20
<b>Country</b>	United States
<b>Education Level</b>	College grad 28%
<b>Marital Status</b>	Not_married 93%

### Topics Of Interest

More Interested

- International Affairs
- US Politics
- Health and Wellness

Less Interested

- Elections
- News
- US Social Issues

### Quotes

"I was in Charleston when the tragedy occurred, puts things into perspective..."

"Getting recruits at a young age is the key to taking away their future."


"Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"

### Most Viewed Contents

- Health and Wellness** Is The South Racist? We Asked South Carolinians  
30. June 2015
- US Politics** 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump  
28. October 2015
- Health and Wellness** Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+  
12. September 2017
- Police** Did These Swedish Cops School The NYPD?  
24. April 2015
- Turkey** Confederate Flag Debate - We Ask People In South Carolina What They Think  
23. June 2015

### Audience Size

7,600,000 people



1/10, 28. February 2018, YouTube

# Which one do you prefer?


A lot of numbers...

30	AJ+	success		~28/02/2017	facebook	7554 / 7555	1/03/2017 9:32:20	1/03/2017 15:03:56
29	AJE	success		~28/02/2017	facebook	1647 / 1647	1/03/2017 9:32:15	2/03/2017 10:29:07
28	AJ+ arabic	success		~28/02/2017	facebook	2760 / 2760	1/03/2017 9:32:08	1/03/2017 11:58:48
27	QF	success		~28/02/2017	facebook	158 / 158	1/03/2017 9:32:04	1/03/2017 10:06:52
26	AJ+	success		~28/02/2017	youtube	4729 / 4729	1/03/2017 9:31:55	7/03/2017 20:27:36
25	AJE	success		01/01/2016~28/02/2017	youtube	9156 / 9156	1/03/2017 9:31:45	12/03/2017 22:48:46
24	AJ+ arabic	success		~28/02/2017	youtube	2443 / 2443	1/03/2017 9:31:41	2/03/2017 2:51:22
23	QF	success		~28/02/2017	youtube	113 / 113	1/03/2017 9:31:34	1/03/2017 10:43:45
22	QF	success		~31/01/2017	youtube	109 / 109	1/02/2017 9:49:09	1/02/2017 10:48:37
21	AJ+ arabic	success		~31/01/2017	youtube	2187 / 2248	1/02/2017 9:49:06	5/02/2017 10:35:26
20	AJE	success		01/01/2016~31/01/2017	youtube	8501 / 8501	1/02/2017 9:49:02	13/02/2017 10:24:04
19	AJ+	success		~31/01/2017	youtube	4719 / 4719	1/02/2017 9:48:59	6/02/2017 19:04:27
18	AJ+	success		~31/01/2017	facebook	7186 / 7187	1/02/2017 3:10:02	20/02/2017 16:48:59
17	QF	success		~31/01/2017	facebook	149 / 149	1/02/2017 3:10:02	20/02/2017 16:34:07
16	AJE	success		~31/01/2017	facebook	1507 / 1507	1/02/2017 3:10:02	20/02/2017 16:41:09
15	AJ+ arabic	success		~31/01/2017	facebook	2520 / 2520	1/02/2017 3:10:00	20/02/2017 17:15:27

Austin, a 35-year-old diving enthusiast.

vs.

**Persona Profile**



Name: Austin  
Gender: Male  
Age: 38  
Country: United States

**About Persona**

Austin is a 38 year old male living in the United States and works in the Food Preparation and Services field. He likes to read about Racism, US-affairs, and US-politics on his Mobile. He usually watches about 14 minutes of video.

**Topics of Interest**

More Interested Topics

- Racism
- US-affairs
- US-politics

Less Interested Topics

- Technology & Science
- Refugees
- Entertainment

**Quotes** All Selected Comments

"These people are mentally unstable to function in society so I do think that they should be kept in these facilities but shouldn't be treated differently"

"You forgot one important thing. If the police have PC all that other stuff is irrelevant."

"this is somewhat misleading. NOT ALL TERRORISTS are linked to ISIS. We have caught lots of potential terrorists crossing illegally, as well as lots of murderers, rapists, child molesters, and gang members. THAT alone makes us much safer. The wall does work. When they built it in Yuma, it drastically reduced the number of crossings. Can you stop it all together, no. Can you slow it, absolutely. This video is also misleading about crossing through the desert for miles and miles. Yes thats true in some areas, but in others, you cross the river and you can disappear in town within 15 seconds."

**Most Viewed Videos**

- US-politics** Is The South Racist? We Asked South Carolinians  
2015-05-30 18:39:50
- US-affairs** Did These Swedish Cops School The NYPD?  
2015-04-24 01:16:25
- Others** Americans Show Why USA Is The #1 Country In The World  
2015-07-03 20:32:11
- US-politics** Donald Trump? In Mexico, We Asked Mexicans What They Thought About Trump  
2015-10-28 19:30:33
- US-affairs** Why Walls Won't Secure The U.S.-Mexico Border  
2016-05-25 20:00:00
- International Affairs** How To Stop An Oil And Gas Pipeline: The Unis'ofen Camp Resistance  
2014-11-05 19:48:44
- US-affairs** 6 Things To Know If Police Stop You In The US  
2014-08-16 12:00:09
- US-politics** Are Immigrants Stealing American Jobs?

***"Personas give faces to data."***

**—Dr. Jim Jansen**

# What is APG?

**A methodology and a system for automatically creating personas from online analytics data** (An et al., 2017; Jung et al., 2017; Salminen et al., 2017).

## Current status:

- a. processing hundreds of millions of user interactions from YouTube, Facebook and Google Analytics.
- b. stable and robust system using Flask Web framework, PostgreSQL database, own server, and Pandas/scikit-learn data analysis libraries.
- c. deployed with Al Jazeera Media Network, Qatar Foundation and Qatar Airways.



# The Bottom Line

Automatically generated personas fix many problems of persona creation reported in the literature (e.g., Chapman & Milham, 2006; Salminen et al., 2017).

**Automatic persona generation system developed at QCRI is, to our knowledge, the most advanced system in the world for this purpose** (see An et al., 2017 for review).

# How does APG work?

APG methodology has been reported in detail in our papers (see e.g. An et al., 2016; An et al., 2017; Kwak et al., 2017; Jung et al., 2017; Salminen et al., 2017).

We use non-negative matrix factorization, LDA, and other techniques. Read the papers for details.

**Summary: We are confident in the basic approach. Now, we want to improve each section of the persona profile, and the profile as a whole.**

The rest of this presentation focuses on the research roadmap.

# Research Roadmap

**Image:** How to retrieve / generate and choose correct persona profile pictures?

**Attributes:** How to infer attributes, such as psychographics, needs and wants, political orientation and brand affinities.

**Quotes:** How to find representative, contextually relevant, and non-distracting comments describing the persona.

The screenshot displays a user profile for 'Phil', a 20-year-old male from the United States. The profile includes a profile picture, a bio, and a list of attributes: Name (Phil), Gender (Male), Age (20), Country (United States), Education Level (College grad 28%), and Marital Status (Not\_married 93%). The 'Topics Of Interest' section lists 'More Interested' topics like International Affairs, US Politics, and Health and Wellness, and 'Less Interested' topics like Elections, News, and US Social Issues. The 'Quotes' section features three representative comments. The 'Most Viewed Contents' section lists several videos, including 'Is The South Racist? We Asked South Carolinians' and 'Explaining The DMZ: The Militarized Border Between The Koreas'. The 'Audience Size' section shows 7,600,000 people. The profile is dated 1/10, 28. February 2018, on YouTube.

**Temporal analysis:** How to analyze change and stability of personas over time?

**Information architecture:** How to choose the correct information elements and layouts for a given user, use case or industry?

**Evaluation:** 1) How to ensure personas are complete, clear, consistent and credible?  
2) How to measure usefulness of personas for individuals and organizations?

Finding better ways to automatically process and choose useful user information from vast amounts of online data.  
"Giving faces to data"

# Research Roadmap

**Image:** How to retrieve / generate and choose correct persona profile pictures?

**Attributes:** How to infer attributes, such as psychographics, needs and wants, political orientation and brand affinities.

**Quotes:** How to find representative, contextually relevant, and non-distracting comments describing the persona.

The screenshot displays a user profile for 'Phil' (Male, 20, United States) with the following details:

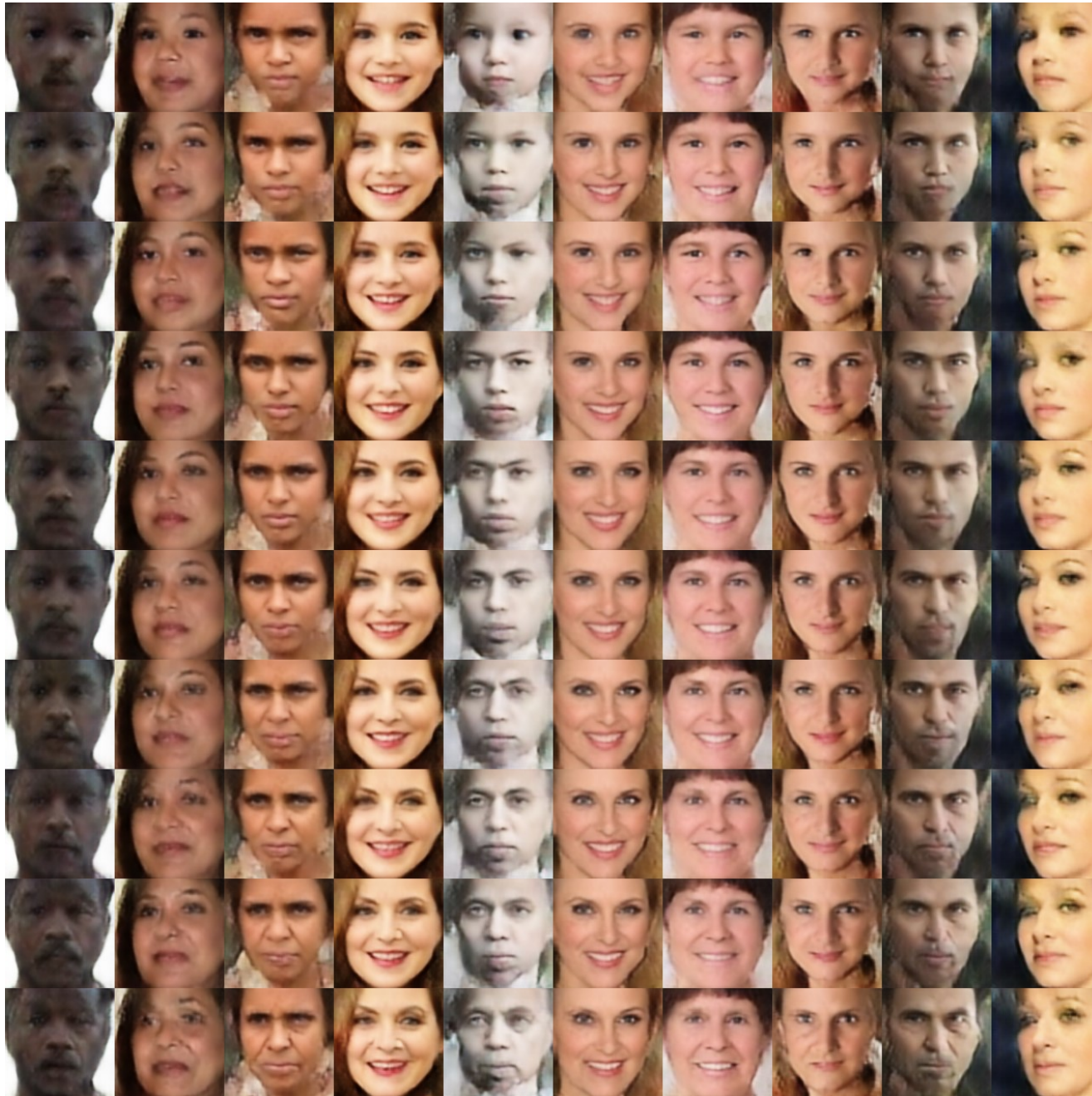
- Topics Of Interest:**
  - More Interested: International Affairs, US Politics, Health and Wellness
  - Less Interested: Elections, News, US Social Issues
- Quotes:**
  - "I was in Charleston when the tragedy occurred, puts things into perspective..."
  - "Getting recruits at a young age is the key to taking away their future."
  - "Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"
- Most Viewed Contents:**
  - Health and Wellness: Is The South Racist? We Asked South Carolinians (30. June 2015)
  - US Politics: 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump (28. October 2015)
  - Health and Wellness: Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+ (12. September 2017)
  - Police: Did These Swedish Cops School The NYPD? (24. April 2015)
  - Turkey: Confederate Flag Debate - We Ask People In South Carolina What They Think (23. June 2015)
- Audience Size:** 7,600,000 people

**Temporal analysis:** How to analyze change and stability of personas over time?

**Information architecture:** How to choose the correct information elements and layouts for a given user, use case or industry?

**Evaluation:** 1) How to ensure personas are complete, clear, consistent and credible?  
2) How to measure usefulness of personas for individuals and organizations?

Finding better ways to automatically process and choose useful user information from vast amounts of online data.  
"Giving faces to data"



- Used GAN to alter age of 200 base pictures
- Results not impressive
- Tougher problem than advertised

**The new goal: Generate facial images from input [age, gender, country]**

# Research Roadmap

**Image:** How to retrieve / generate and choose correct persona profile pictures?

**Attributes:** How to infer attributes, such as psychographics, needs and wants, political orientation and brand affinities.

**Quotes:** How to find representative, contextually relevant, and non-distracting comments describing the persona.

The screenshot displays a user profile for 'Phil' (Male, 20, United States) with the following details:

- Topics Of Interest:** More Interested (International Affairs, US Politics, Health and Wellness); Less Interested (Elections, News, US Social Issues).
- Quotes:** "I was in Charleston when the tragedy occurred, puts things into perspective..."; "Getting recruits at a young age is the key to taking away their future."; "Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place".
- Attributes:** Name: Phil, Gender: Male, Age: 20, Country: United States, Education Level: College grad 28%, Marital Status: Not\_married 93%.
- Most Viewed Contents:** A list of video thumbnails with titles like 'Is The South Racist? We Asked South Carolinians', 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump', 'Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+', 'Did These Swedish Cops School The NYPD?', and 'Confederate Flag Debate - We Ask People In South Carolina What They Think'.
- Audience Size:** 7,600,000 people.

**Temporal analysis:** How to analyze change and stability of personas over time?

**Information architecture:** How to choose the correct information elements and layouts for a given user, use case or industry?

**Evaluation:** 1) How to ensure personas are complete, clear, consistent and credible?  
2) How to measure usefulness of personas for individuals and organizations?

Finding better ways to automatically process and choose useful user information from vast amounts of online data.  
"Giving faces to data"

# Opportunities are endless...



## APG

- Dist(Age)
- Dist(Gender)
- Dist(Location)
- Dist(Topical interests)

≈



*Automatically inferring  
user attributes from  
tweets + followers*

- Psychographics
- Political orientation
- Brand affinities
- Needs and wants

# Research Roadmap

**Image:** How to retrieve / generate and choose correct persona profile pictures?

**Attributes:** How to infer attributes, such as psychographics, needs and wants, political orientation and brand affinities.

**Quotes:** How to find representative, contextually relevant, and non-distracting comments describing the persona.

The screenshot displays a user profile for 'Phil', a 20-year-old male from the United States. The profile includes a bio, a list of topics of interest (International Affairs, US Politics, Health and Wellness, Elections, News, US Social Issues), a list of quotes, and a list of most viewed contents. The audience size is shown as 7,600,000 people. The profile is dated 1/10, 28. February 2018, YouTube.

Attribute	Value
Name	Phil
Gender	Male
Age	20
Country	United States
Education Level	College grad 28%
Marital Status	Not_married 93%

**Topics Of Interest**

More Interested

- International Affairs
- US Politics
- Health and Wellness

Less Interested

- Elections
- News
- US Social Issues

**Quotes**

"I was in Charleston when the tragedy occurred, puts things into perspective..."

"Getting recruits at a young age is the key to taking away their future."

"Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"

**Most Viewed Contents**

- Health and Wellness: Is The South Racist? We Asked South Carolinians (30. June 2015)
- US Politics: 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump (28. October 2015)
- Health and Wellness: Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+ (12. September 2017)
- Police: Did These Swedish Cops School The NYPD? (24. April 2015)
- Turkey: Confederate Flag Debate - We Ask People In South Carolina What They Think (23. June 2015)

**Audience Size**

7,600,000 people

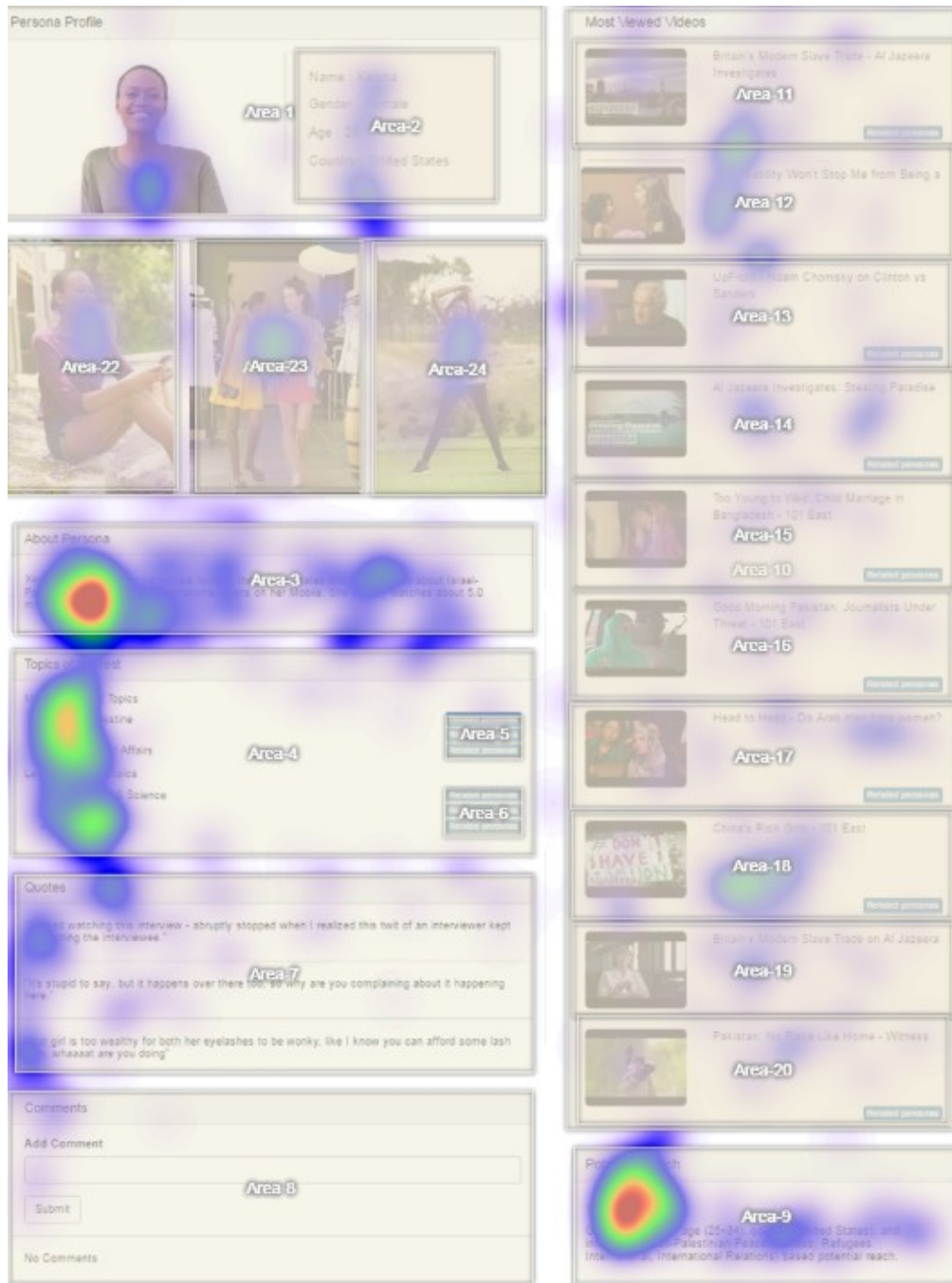
**Temporal analysis:** How to analyze change and stability of personas over time?

**Information architecture:** How to choose the correct information elements and layouts for a given user, use case or industry?

**Evaluation:** 1) How to ensure personas are complete, clear, consistent and credible?  
2) How to measure usefulness of personas for individuals and organizations?

Finding better ways to automatically process and choose useful user information from vast amounts of online data.  
"Giving faces to data"





- Did a user study to see how people interacted with personas
- Found that quotes and images cause judgment toward the persona

**The new goal: Find out if toxic comments steer attention away from other information (and develop advanced filtering)**

# Research Roadmap

**Image:** How to retrieve / generate and choose correct persona profile pictures?

**Attributes:** How to infer attributes, such as psychographics, needs and wants, political orientation and brand affinities.

**Quotes:** How to find representative, contextually relevant, and non-distracting comments describing the persona.

The screenshot displays a user profile for 'Phil' (Male, 20, United States) with 0 Favorites, 0 Comments, and 0 Shares. The profile includes a profile picture, a bio, and a list of attributes: Name (Phil), Gender (Male), Age (20), Country (United States), Education Level (College grad 28%), and Marital Status (Not\_married 93%). The 'Topics Of Interest' section lists 'More Interested' topics (International Affairs, US Politics, Health and Wellness) and 'Less Interested' topics (Elections, News, US Social Issues). The 'Quotes' section contains three quotes, including one about Charleston and another about Trump and Korea. The 'Most Viewed Contents' section lists several videos, such as 'Is The South Racist? We Asked South Carolinians' and 'Explaining The DMZ: The Militarized Border Between The Koreas'. The 'Audience Size' section shows 7,600,000 people. The bottom of the page shows a row of profile pictures and the date '1/10, 28. February 2018, YouTube'.

**Temporal analysis:** How to analyze change and stability of personas over time?

**Information architecture:** How to choose the correct information elements and layouts for a given user, use case or industry?

**Evaluation:** 1) How to ensure personas are complete, clear, consistent and credible?  
2) How to measure usefulness of personas for individuals and organizations?

Finding better ways to automatically process and choose useful user information from vast amounts of online data.  
"Giving faces to data"

# View changes in personas over time

APG

Send an email invitation

QR Test

## View Personas

Content source: YouTube

Generation: 30. November 20















Persona type: Country

Number of personas: 5

Show

Personas Reach Comparison Contents Chronology

Select generations

	10. October 2017	31. October 2017	30. November 2017	
 Shahril Male   34 Malaysia	<ul style="list-style-type: none"><li>Qatar Events</li><li>Outreach</li><li>Qatar Airways Recruiting</li></ul>	<ul style="list-style-type: none"><li>Qatar Events</li><li>Qatar Tourism</li><li>Lounges</li></ul>	<ul style="list-style-type: none"><li>Travel Offers</li><li>Employee Stories</li><li>Others</li></ul>	 Shahril Male   34 Malaysia
 Rami Male   27 Qatar	<ul style="list-style-type: none"><li>Qatar Events</li><li>Qatar Tourism</li><li>Contests</li></ul>	<ul style="list-style-type: none"><li>Qatar Events</li><li>Outreach</li><li>Qatar Tourism</li></ul>	<ul style="list-style-type: none"><li>Travel Offers</li><li>Employee Stories</li><li>Others</li></ul>	 Rami Male   27 Qatar
 Sachin Male   20 India	<ul style="list-style-type: none"><li>Employee Stories</li><li>Contests</li><li>Frequent Flyer Programs</li></ul>	<ul style="list-style-type: none"><li>Qatar Events</li><li>Contests</li><li>Outreach</li></ul>	<ul style="list-style-type: none"><li>Travel Offers</li><li>Employee Stories</li><li>Others</li></ul>	 Sachin Male   20 India
 Khaled Male   25 Egypt	<ul style="list-style-type: none"><li>Qatar Events</li><li>Contests</li><li>Frequent Flyer Programs</li></ul>	<ul style="list-style-type: none"><li>Qatar Events</li><li>Contests</li><li>Outreach</li></ul>	<ul style="list-style-type: none"><li>Travel Offers</li><li>Employee Stories</li><li>Others</li></ul>	 Khaled Male   25 Egypt
 Mert Male   40 United States			<ul style="list-style-type: none"><li>Travel Offers</li><li>Employee Stories</li><li>Others</li></ul>	 Mert Male   40 United States
 Sami Male   38 Qatar	<ul style="list-style-type: none"><li>Cabin Crews</li><li>Employee Stories</li><li>Contests</li></ul>			 Sami Male   38 Qatar
 Emre Male   25 Germany		<ul style="list-style-type: none"><li>Qatar Events</li><li>Outreach</li><li>Qatar Airways Recruiting</li></ul>		 Emre Male   25 Germany

Shows the similarities and changes in the persona set over time. **How to use? See how the content preferences of the audience are changing over time.**

# Research Roadmap

**Image:** How to retrieve / generate and choose correct persona profile pictures?

**Attributes:** How to infer attributes, such as psychographics, needs and wants, political orientation and brand affinities.

**Quotes:** How to find representative, contextually relevant, and non-distracting comments describing the persona.

The screenshot shows a user profile for 'Phil', a 20-year-old male from the United States. The profile includes a photo, a bio, and a list of attributes: Name (Phil), Gender (Male), Age (20), Country (United States), Education Level (College grad 28%), and Marital Status (Not\_married 93%). The profile also features sections for 'Topics Of Interest' (More Interested: International Affairs, US Politics, Health and Wellness; Less Interested: Elections, News, US Social Issues), 'Quotes' (e.g., 'I was in Charleston when the tragedy occurred, puts things into perspective...'), 'Most Viewed Contents' (e.g., 'Is The South Racist? We Asked South Carolinians'), and 'Audience Size' (7,600,000 people). The profile is dated 1/10, 28. February 2018, YouTube.

**Temporal analysis:** How to analyze change and stability of personas over time?

**Information architecture:** How to choose the correct information elements and layouts for a given user, use case or industry?

**Evaluation:** 1) How to ensure personas are complete, clear, consistent and credible?  
2) How to measure usefulness of personas for individuals and organizations?

Finding better ways to automatically process and choose useful user information from vast amounts of online data.  
"Giving faces to data"

# Research Roadmap

**Image:** How to retrieve / generate and choose correct persona profile pictures?

**Attributes:** How to infer attributes, such as psychographics, needs and wants, political orientation and brand affinities.

**Quotes:** How to find representative, contextually relevant, and non-distracting comments describing the persona.

The screenshot displays a user profile for 'Phil' (Male, 20, United States) with the following details:

- Topics Of Interest:**
  - More Interested: International Affairs, US Politics, Health and Wellness
  - Less Interested: Elections, News, US Social Issues
- Quotes:**
  - "I was in Charleston when the tragedy occurred, puts things into perspective..."
  - "Getting recruits at a young age is the key to taking away their future."
  - "Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"
- Most Viewed Contents:**
  - Health and Wellness: Is The South Racist? We Asked South Carolinians (30. June 2015)
  - US Politics: 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump (28. October 2015)
  - Health and Wellness: Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+ (12. September 2017)
  - Police: Did These Swedish Cops School The NYPD? (24. April 2015)
  - Turkey: Confederate Flag Debate - We Ask People In South Carolina What They Think (23. June 2015)
- Audience Size:** 7,600,000 people

**Temporal analysis:** How to analyze change and stability of personas over time?

**Information architecture:** How to choose the correct information elements and layouts for a given user, use case or industry?

**Evaluation:** 1) How to ensure personas are complete, clear, consistent and credible?  
2) How to measure usefulness of personas for individuals and organizations?

Finding better ways to automatically process and choose useful user information from vast amounts of online data.  
"Giving faces to data"

# Research Roadmap

**Image:** How to retrieve / generate and choose correct persona profile pictures?

**Attributes:** How to infer attributes, such as psychographics, needs and wants, political orientation and brand affinities.

**Quotes:** How to find representative, contextually relevant, and non-distracting comments describing the persona.

The screenshot displays a user profile for 'Phil' (Male, 20, United States) with the following details:

- Topics Of Interest:**
  - More Interested: International Affairs, US Politics, Health and Wellness
  - Less Interested: Elections, News, US Social Issues
- Quotes:**
  - "I was in Charleston when the tragedy occurred, puts things into perspective..."
  - "Getting recruits at a young age is the key to taking away their future."
  - "Only if trump hasn't talked about Korea like that we would of stayed safe and war free and now I'm worrying when my final day is here cause of the war and Korea is a nice place"
- Most Viewed Contents:**
  - Health and Wellness: Is The South Racist? We Asked South Carolinians (30. June 2015)
  - US Politics: 'Donald Trump' In Mexico: We Asked Mexicans What They Thought About Trump (28. October 2015)
  - Health and Wellness: Explaining The DMZ: The Militarized Border Between The Koreas | Direct From With Dena Takruri - AJ+ (12. September 2017)
  - Police: Did These Swedish Cops School The NYPD? (24. April 2015)
  - Turkey: Confederate Flag Debate - We Ask People In South Carolina What They Think (23. June 2015)
- Audience Size:** 7,600,000 people

**Temporal analysis:** How to analyze change and stability of personas over time?

**Information architecture:** How to choose the correct information elements and layouts for a given user, use case or industry?

**Evaluation:** 1) How to ensure personas are complete, clear, consistent and credible?  
2) How to measure usefulness of personas for individuals and organizations?

Finding better ways to automatically process and choose useful user information from vast amounts of online data.  
"Giving faces to data"

# APG linkages to Computer Science

Challenge	Potential solutions
Image	Generative Adversarial Networks
Attributes	Entity Resolution, Data Mapping, Multistage Sampling
Quotes	Detecting Hate Speech, Topic Modeling
Temporal Analysis	Tensor Factorization, Time-series, Data Stream Analysis, Anomaly Detection, Concept Drift
Information Architecture	Ethnography, Crowd Experiments, Human-computer Interaction, Adaptive Systems, Information Science
Evaluation	Factor Analysis, Structural Equation Modeling, Case Study

# References

- An, J., Kwak, H., & Jansen, B. J. (2016). Validating Social Media Data for Automatic Persona Generation. Presented at the The Second International Workshop on Online Social Networks Technologies (OSNT-2016), 13th ACS/IEEE International Conference on Computer Systems and Applications AICCSA 2016. 29 November - 2 December.
- An, J., Haewoon, K., & Jansen, B. J. (2017). Personas for Content Creators via Decomposed Aggregate Audience Statistics. In Proceedings of Advances in Social Network Analysis and Mining (ASONAM 2017).
- Chapman, C. N., & Milham, R. P. (2006). The Personas' New Clothes: Methodological and Practical Arguments against a Popular Method. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50(5), 634-636.
- Chapman, C., Krontiris, K., & Webb, J. (2015). Profile CBC: Using Conjoint Analysis for Consumer Profiles. Google Research.
- Cooper, A. (1999). The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity (1st edition). Indianapolis, IN: Sams - Pearson Education.
- Jung, S.-G., An, J., Kwak, H., Ahmad, M., Nielsen, L., & Jansen, B. J. (2017). Persona Generation from Aggregated Social Media Data. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (pp. 1748-1755). New York, NY, USA: ACM.
- Kwak, H., An, J., & Jansen, B. J. (2017). Automatic Generation of Personas Using YouTube Social Media Data (pp. 833-842). Presented at the Hawaii International Conference on System Sciences (HICSS-50). 4-7 January, Waikoloa, Hawaii.
- Nielsen, L. (2013). Personas - User Focused Design. Springer-Verlag London: Springer Science & Business Media.
- Pruitt, J., & Grudin, J. (2003). Personas: Practice and Theory. In Proceedings of the 2003 Conference on Designing for User Experiences (pp. 1-15). New York, NY, USA: ACM.
- Salminen, J., Şengün, S., Haewoon, K., Jansen, B. J., An, J., Jung, S., Vieweg, S., & Harrell, F. (2017). Generating Cultural Personas from Social Data: A Perspective of Middle Eastern Users. In Proceedings of The Fourth International Symposium on Social Networks Analysis, Management and Security (SNAMS-2017). Prague, Czech Republic.



# Thank you!

**See demo: <https://persona.qcri.org>**

To become a client, contact Dr. Jim Jansen:  
[bjansen@hbku.edu.qa](mailto:bjansen@hbku.edu.qa) (*free for a limited time*)

To collaborate on research, contact Dr. Joni Salminen:  
[jsalminen@hbku.edu.qa](mailto:jsalminen@hbku.edu.qa) (*open for collaboration*)