Shaping AI Systems with Cultural Data

Report on IVOW’s Women in History Data Ideation Challenge

By Davar Ardalan, Kee Malesky, Nikki McLay, with the IVOW Team & Advisors
Contents

1. Introduction pg. 3
2. Data Ideation Challenge pg. 6
3. First Prize Winner pg. 10
4. Advisors’ Corner pg. 15
5. The Role Of Testers And Qa Leaders pg. 26
6. Imbuing Current Ai Technology With Cultural Intelligence pg. 27
7. Sina Storyteller pg. 27
8. Women In AI Global Summit pg. 30
9. Conclusion pg. 31

With thanks to the IVOW Team: Davar Ardalan; Kee Malesky; Nikki McLay; Nisa McCoy; Karim Ardalan; Robert Malesky; and our AI, Sina Storyteller.

Cover images courtesy Wikimedia Commons. Left to right:

Dame Whina Cooper (1895–1994) was a heralded New Zealand Māori leader who tirelessly advocated for Māori rights and welfare – particularly for women.

Indira Gandhi (1917 – 1984) was an Indian politician and stateswomen, she became the first, and to this date, the only female Prime Minister of India.

Shirley Chisholm (1924 - 2005) was an American politician, educator, and author. In 1968, she became the first black woman elected to the United States Congress.

Marie Curie (1867 – 1937) was a Polish physicist and chemist who became the first woman to win a Nobel Prize, and the only woman to win the Nobel prize twice in separate scientific fields.

Frida Kahlo (1907 – 1954) was Mexican painter best known for her self-portraits that explored themes around the body, gender equality, culture and politics.
1. Introduction

This report details our learnings at IVOW, followed by deeper insights from our advisors and collaborators on our recent Women in History Data Ideation Challenge with Topcoder.

IVOW stands for Intelligent Voices of Wisdom. At IVOW AI, we are designing the storytelling chatbot Sina (currently prototyped on Google Home). We understand that current chatbots, and the decision trees and knowledge graphs feeding them, are desperately lacking in cultural and gender richness and diversity. We understand the critical need to nurture a user experience for customers and brands that is simultaneously culturally-specific, narrative-rich, globally-appealing, and interactive.

The data output by new technologies can only be as good as the data input. Unlike their male counterparts, the stories of many of the women who have paved the way in science, technology, arts and culture have been watered down, overshadowed, trivialized, or simply ignored and undocumented. We know that thousands of women's stories are archived in one way or another in public records stored online, but how have they been tagged and labelled? Can this data be improved? Can the data pulled from these many internet archives be cleaned and refined so it is accurate and reflective of the contributions women have made throughout history? Can this data be made machine-ready to be utilized by today's technology? It is clear to us that we need a central repository of data that is easily accessible and explainable in terms of its tags and keywords.

For this challenge, our goal was to find public data sources about women throughout history and suggest how that data could be used to gain new insights for AI products and solutions with a concentration on women. Our primary focus was to get a well thought-out overview of what is possible, what relevant data on women throughout history can be found and where, and how the data should be collected.

Orchestrated efforts for shaping the societal and cultural impacts of our AI systems have been long overdue. IVOW's mission to bring cultural awareness to AI through the power of storytelling will be leading the way in such efforts. Imagine a day when young girls interacting with various smart devices can interactively learn engaging stories about pioneering female scientists and leaders, instead of fairytale princesses... That's exactly the near-future reality that the Women in History Data challenge will be enabling.

– Nasrin Mostafazadeh, AI expert

Today's AI systems are powerful, but they lack knowledge of the role of women and non-western cultures across human history. For perspective, we searched a major database with more than 70,000 biographies of “the most memorable people in our planet's history,” and found that almost 90% of female public figures going back to 3501 BCE were tagged as companion, model, pornographic actor, or social activist.
To widen the diversity in our cognitive pool, AI ethics advocate/researcher and IVOW data ideation advisor Susanna Raj says we need to include the diversity factor in all stages -- the conceptualization of ideas, the collection of data, the modeling, the development of socially beneficial use cases, and the humane deployment of AI.

*AI trained on datasets that have erased the contributions of women and minorities cannot magically see and value what is not there... To train our AI systems to be effective gatekeepers of cultural context, we must train the next generation of humans to see with an ethical lens, to spot and call out the biases.*

– Susanna Raj, Cognitive science, HCI & AI ethics researcher, IVOW Advisor

The global pandemic has made our mission even more critical as humans yearn to understand one another and enterprises seek more authentic ways to engage with their customers online.

The foundational ontology of our work at IVOW is: How do we introduce storytelling and culture into artificial intelligence?

- **Conversational AI with Cultural Richness**
  The current conversational AI ecosystem is desperately lacking in cultural wisdom and diversity. With IVOW, our collaborators are able to develop conversational AIs with rich cultural knowledge, compelling narratives, and customization for worldwide appeal.

- **AI Storytelling Solutions**
  Human beings transmit knowledge to future generations through stories, and we are constantly guided by the narratives around us. By combining the power of new technologies with ancient storytelling wisdom, IVOW empowers brands to tell their story in a way that resonates with a diverse customer base.

- **Globally Inclusive Datasets**
With more representative data, we know that brands, advertisers, and researchers in all industries can have a deeper understanding of global audiences and make better-informed decisions. IVOW helps cultivate a cultural lens on data, and a new data ecology within organizations.

As a big data and AI fanatic, I see how Cultural Intelligence is what makes us human, and insights gained in the future from AI and Big Data by using a cultural lens will aid Governments and Companies to better converse and communicate. This will in turn impact every aspect of people’s lives by showing more authentic empathy and understanding.

– Mark Germishuys, AI and Big Data Pioneer and CEO at NGA

Marie Smith is chief information officer at Data 360, and also an IVOW advisor. She feels that, as more and more machines governed by internet connections and various forms of artificial intelligence are deployed globally by technology companies large and small, proper machine-ready data sets are essential for our future to ensure an equitable and just society.

For the last few years, clean and plentiful sources of data along with performance-driven data models are in high demand! McKinsey Global Institute predicts that the shortage of data scientists in the US could increase to 250,000 data scientists by 2024. This means that we have a chance now to create new and insightful data models that can heal the past, inform the present, and guide the future with companies like IVOW.ai.

– Marie Smith, CIO, Data 360

Leading bioinformatics software test engineer Aprajita Mathur is also a senior advisor to IVOW’s data ideation challenge. She knows that in the medical sciences, each human is not just figuratively but also literally unique. The concept of personalization is as important in AI as it is in medicine.

Someday my great-great-granddaughter will ask, ‘Google, why do Indians wear a red dot on their foreheads?’ I want the answer to be truly reflective of her ancestry and include the emotions that I would feel in answering that question, rather than the one-size-fits-all answer that it’s common practice to do so.

– Aprajita Mathur, leading bioinformatics software test engineer
Since 2018, IVOW has delivered on the promise of cultural intelligence in AI through storytelling, with over 2,000 investors, corporate and civic tech leaders, policymakers, AI experts, testers, QA leaders, anthropologists, and storytellers. This includes engagements with Topcoder, Soul Machines, XPrize, ITU, NASA's International Space Apps Challenge, WeWork Labs, WaiAccelerate, Amazon, IBM, and KiwiTech.

What you are doing to lower barriers to connecting across cultures and to amplify the needs of different cultures clearly is both important and inspiring. Brands are learning that to be successful in the mass-access market that we have now, they need to be able to make these connections. What I call "tyranny of popularity" doesn't work any more. Having the right connections to draw on does.

- Jana Eggers, CEO Nara Logics Inc

2. Data Ideation Challenge

IVOW launched the Women in History Data Ideation Challenge, together with Topcoder, in early July 2020. We had 15 submissions and 7 winners. The results confirmed some of our assumptions and fears:

1. Historic gender and cultural biases perpetuate in the ecology of AI;
2. Current classification of gender, ethnicity, and race in Wikipedia is flawed and lacking;
3. Improved data ecologies that account for gender, culture, and history to create better algorithms are vital for popularizing future AI products.

Not surprisingly, most of the Topcoder challenge participants used Wikipedia as a source for information on women in history and fiction. This in and of itself proves the limits of content we can draw on. Wikipedia has some 6 million articles and is a viable base for data collection and training data for future NLP-Machine Learning based data exploration. Having said this, most participants
acknowledged the difficulties with Wikipedia — in particular, biases, poor writing, and boring raw facts, as well as the lack of global cultural references.

**On Gender Bias:**
Word embeddings trained on Wikipedia disproportionately associate male terms with science terms, and female terms with art terms, which could amplify stereotypes and lead to replicating biases. This was also shown in the 2017 report, published in *Science*, “*Semantics derived automatically from language corpora contain human-like biases*.”

**On Classification of Culture & Ethnicity:**
The results show that even with many words joined under ontology and property in the query, the fields CulturalSource, TimePeriod, and Relations were still hard to fully populate. This suggests the necessity to use machine learning to extract values for these fields from TextualAbstract. Some applicants suggested creating categories or keywords, based on location, language, occupation, etc. The significance of proper tagging was noted in one submission, an important consideration given recent news about a racist, misogynistic database. Another suggestion was to use machine learning to enable AI to distinguish between “important” and “unimportant” sentences in a text, as a way to write a better profile.

**On Current Datasets:**
A few of the applicants used other free sources (Pantheon, Women in Hollywood, Fortune, etc.), which was encouraging since we know that Wikipedia is not sufficient. Pantheon datasets contain many historical female figures from various domains (politics, science, culture, sports, etc), with their cultural background (when and where they were born). For each person in the dataset, a short summary is collected from the corresponding Wikipedia article.

Overall, Pantheon and Wikipedia mainly include well-known women, and famous people’s entries are subject to a more intensive peer-review process, which can result in increased accuracy of information over time. But this also means that there are many cultures which are not properly documented because information about them on the web is missing or insufficient. We want and need to collect data not just on famous women, but also on obscure figures from history, fiction, and mythology.

In addition, we noticed that several of the sample datasets gathered via the Pantheon database included the occupation of “pornographic actor,” with 174 names included from a sample of 13,000+ famous women. This sample data also included 3,950 actors, 10 anthropologists, 7 archaeologists, 11 architects, and 211 fashion models. This indicates how crucial it is to be clear about what emphasis we want to bring to machine-readable AI datasets about women in history and their contributions.

It’s important to note that in their FAQ, Pantheon describes itself as “a project that uses biographical data to expose patterns of human collective memory. Pantheon contains data on more than 70k biographies, which Pantheon distributes through a powerful data visualization engine centered on locations, occupations, and biographies.”
We went to the [Pantheon database](https://www.pantheon.org) and filtered for females in public figure occupations from 3501 BCE to 2020 CE. The visualization created illustrates that the major occupations for females in the past 5,000+ years have been modelling and pornographic acting.

(Via Pantheon dataset visualizing occupations females in public figure occupations 3501 BC - 2020)

To focus on contemporary women, we filtered for females born in 1964 in the US in all occupations. The list brings up 59 names. Melinda Gates is the only businesswoman; the majority are labeled as actors, musicians, film directors, pornographic actors, models, politicians, or companions (we have an in-depth breakdown on ‘companions’ later in this report). The list includes no female scientists, engineers, or technologists. Let's see about adding biochemist [Jennifer Doudna](https://en.wikipedia.org/wiki/Jennifer_Doudna) to this list. Born in 1964, according to Wikipedia, “Doudna has been a leading figure in what is referred to as the ‘CRISPR revolution’ for her fundamental work and leadership in developing CRISPR-mediated genome editing.”
We also looked at women who were labeled ‘Celebrity.’ One example was Peggy Guggenheim, an important 20th century art collector, museum founder, and philanthropist. Wikipedia categorizes Guggenheim under those terms, but Pantheon merely lists her as a celebrity, because they don’t use those categories.

Pantheon’s website notes that their database is a popular resource with some history teachers and scholars in the digital humanities. “Pantheon 1.0 was visited by over half-a-million people, including students, scholars, entertainment industry executives, and history aficionados. The data from Pantheon has been used in several academic publications. For instance, we have used it in work
connecting the centrality of languages in the network of translations with fame, in work exploring the decay of human collective memory, and in work showing how the size and composition of human collective memory changes with the introduction of communication technologies.”

3. First Prize Winner

We are sharing some of the research and methodology behind the IVOW/Topcoder first prize winner. This submission focused on Wikipedia as a source because of its huge number of entries, and suggested a five-step approach:

**Step 1** – Search Keyword Generation - Idea
**Step 2** – Web Scrape 'raw html'
**Step 3** – Convert 'Raw HTML' to 'Pretty HTML'
**Step 4** – Extract Women Wiki Page Link from 'Pretty HTML'
**Step 5** – Web Scrape to Woman_Database

### Approach

**Step 1 – Search Keyword Generation - Idea**

- **City**: worldcities.csv (15000 Cities)
- **Countries**: worldcities.csv (196 Countries)
- **Languages**: language.csv (100 languages)
- **Occupation/Profession**: occupation.csv (250+ entries)

Google Search Keyword: list of women + city/countries/languages + occupation + wikipedia

Example search keyword: list of women indian activist wikipedia

Possible combination of search keyword: 150 countries X 250 Occupations = approx. 37500 keywords

Average results for each keyword - 50 women dataset * 37500 keywords = approx. **1.8 million entries**
Approach
Step 1 – Search Keyword Generation - Example

Steps:
1: Search in google and select the wikipedia page
2: Open wikipedia page
3: Append the wikipedia link in search_list.csv

Note: This step is easily automatable. But, the wikipage usually will have more than 100s of women dataset. It is worthwhile to spend 5-10 seconds to see whether the wikipage has list of women dataset or not.

Approach
Step 2 – Web Scrap 'raw html' of all links in search_list.csv

Note: Using Python Selenium Web Scraper, all the links in "search_list.csv" is scraped and raw html is stored in step_2_raw_list.txt
**Approach**

Step 3 – Convert 'Raw HTML' to 'Pretty HTML' suitable for Women wiki page link extraction

Note: By prettifying raw html, Women wiki page like '/wiki/Chris_Evert' can be extracted in future steps.

**Approach**

Step 4 – Extract Women Wiki Page Link from 'Pretty HTML'

Note: Wikipedia link has been extracted from 'step3_pretty_html.txt'
Approach
Step 5 – Web Scrap 'women wiki links' to Woman_Database.csv

Note:
Finally, the Selenium Web Scraper will extract data and keep the data in the Women_Database.csv in the required format.

Results

• More than 6200 Women dataset has been scraped automatically to show the capability of the script. The results are kept in Women_Database.csv. The script has the potential to scrape millions of famous women's data from internet in a week's time. As the current challenge is not to extract complete dataset, I have stopped at the level of 6200 women dataset.

• Even though the example shown in this presentation is based on Wikipedia pages, the same logic is extensible to other websites by suitably editing the python script.

• In order to remove the inaccurate data from the “Women_Database.csv”, “Text Classification” using Natural Language Processing (NLP) can be trained with the dataset scraped as most of the entries will be accurate due to carefully reviewed search_list.csv in step 1.

• As this algorithm can be used to scrape at least 1,000,000 entries within a couple of days, we can design a completely automated web scraper using “NLP-Text Classifier” from sources other than Wikipedia.

• As the worldcities.csv has latitude and longitude information, it can be added in the Women_Database.csv as a separate column. So, GPS location of the consumer can be compared with Women_Database_lat_long column and recommend the relevant stories to the consumers, making it interesting for the consumers to know interesting stories around them.

• As more of the data is scraped from Wikipedia, the license information is not scraped from the page because most of the pages have “Creative Commons Attribution-Share Alike License.”
• In the profile_info column in “Women_Database.csv” there is a lot of useful information, such as birth date, education, birth city, occupation, and other tags. It helps to build the knowledge graph with appropriate tagging.

• I highly recommend storing the Women_Dataset data in a scalable relational database such as Neo4j as it is faster and graph database management system.

• How do we extract text abstracts 100 words long from larger articles? We need to train the NLP-ML model to classify the sentence as “important” or “not important”.
4. Advisors Corner

We shared early results of the dataset challenge with two of IVOW’s advisors, Rogério Lourenço and Susanna Raj. Here are their thoughts, reactions, and suggestions:

Noble Companions

By Susanna Raj

Susanna Raj is a cognitive science, HCI & AI ethics researcher interested in the ethical applications of AI technology for the greater good of humanity. She has worked on human emotion model training for AI at Intel labs and co-authored a research publication on emotion UI models at MobileHCI 2018. Susanna is a member of the DataEthics4All Think Tank and the #Ethics4NGAI Hackathon Organizing Team.

Introduction

Bias in AI has many entry points; from ideation to data collection, from labeling to modeling, from analysis to deployment; it has multiple ways to get in, but not that many exit options. Once in, it is impossible to exile it out. That is why we must have gatekeepers at every entry point. Gatekeepers that are better than the all-seeing and all-hearing Heimdall guarding the Bifröst in Asgard!
IVOW.ai has embarked on this ambitious and remarkable journey of becoming the gatekeepers of Cultural Context in AI and they have given this gatekeeper the persona of Sina, a cultural chatbot.

As part of the building blocks of Sina, they launched an ideation challenge for women in history. I am excited to share one of the many key findings of Phase-1 of their ideation challenge by Topcoder that piqued my interest.

**Brief Intro to the Methodology**

The ideation challenge asked participants to find existing publicly available datasets on culture within the area of women in history and provide ideas in curating them into machine readable formats. One of the 7 finalists had chosen Pantheon 2.0 and Wikipedia as their sources. Pantheon 2.0 claims to provide a cultural export, to show the ingenuity of human history using collective human memory. The labels for this dataset were taken from Pantheon and the summaries were culled from Wikipedia.

![Pantheon Database](image)

(As featured on the Pantheon database. Michelle Obama is tagged as companion.)

**Key Findings # (1):**

**The Noble Companions**

Out of the 7, this particular dataset caught my eye as I was not familiar with the categorization of Pantheon database and one of the tags was oddly named as ‘Companion.’ The category of
Companions is 595 members strong. At first the label did not make much sense to me, but when I went to Pantheon's database, I realized this tag stood for someone's occupation and professional identity. And that's what led me down a rabbit hole of misrepresentation, mislabeling, and biased tags about women and minority ethnicities.

I prefer to call this finding the Noble Companions. Satirically, of course. For there is nothing noble or friendly about what I found with regards to women's representation in history.

Let me start with the contemporary and go back in history.

Jacqueline Kennedy and Michelle Obama were tagged as ‘companions,’ denoting their occupation as well as identity. Michelle Obama could have easily fallen under the category of writer or lawyer, yet she was placed squarely in the company of spousal companions. She ranks at number 167 out of 595 companions. Jacqueline Kennedy ranks 12 among Companions. This category also hosts two Egyptian queens, Nefertiti and Nefertari. The former, Nefertiti, ranks number 1 according to the metrics of Pantheon and Nefertari ranks 11.

Both queens, like the two former first ladies, made vastly different cultural and historical contributions to their countries, both in their personal and public lives.

These rankings have no connection or relevance to the contributions made by any of the women in the historical or cultural context. It is based on page views, and popularity index derived from the page views.

That was the least of the problems. Out of the 595 Companions, only 36 were male. That the label of Companion applied mostly to women is not only blatantly sexist but incredibly unfair.

What was even more surprising was that the majority of the Egyptian queens, princesses, consorts, and women of noble birth were also categorized under the Companion tag. Similar fate was bestowed upon royal women from Arabian, Mediterranean, and Asian countries as well.

To add salt to the wound, there was another very mysterious occupation/identity tag called ‘Nobleman’ applied to both men and women: 611 Nobleman, of which 231 were women. Most of them were of European descent: princesses, queens, consorts, and women of noble birth or noble by marriage. The graph of Nobleman paints a dramatic picture like this below.

This is the picture that an AI agent will also likely reflect – that noble birth or royalty was predominantly limited to European descent. The rest of the Noblemen were unfortunately locked into ‘Companions’.
Hidden Figures

This type of bias continues into the different occupations as well, very much skewed against women. Women who should be in the ‘Mathematics’ tag were in the ‘Writer’ tag if they wrote a book in their career path. Yet the same logic was not applied to men.

Even the category of writers was not fair to women or other nationalities. There are only 65 writers for India according to Pantheon, of which only 13 are female. In fact, there are hundreds of writers in India and the most historically prominent authors in the English language alone account to 75 or more. India as many may know, is a country of multiple languages, English being one of them.

This mis-categorization in occupation did not spare Shakuntela Devi, a prominent Indian mathematician who is the subject of a 2020 Amazon Prime movie release of the same name. In Pantheon, she was listed under the tag of ‘Writer’. There were no female mathematicians for India according to Pantheon. A total of 13 mathematicians are listed for India in total, but none of them female. None! I wonder where they were categorized? Wikipedia however lists 24 prominent women Indian mathematicians.

However, according to the World Economic Forum, the contributions of India to the field of mathematics shaped its core foundations and yet the cultural map of mathematics according to Pantheon looks like this graph below:
According to this map above and the one below, the contributions of other nationalities in the field of STEM was nothing compared to the European nations.
On the other hand, the entire occupation history of India dramatically looks like this: The industry of ‘Actors’ dominates everything.

But this was not unique to India, this was reflective of the United States as well. Fortunately for the US, the curve is much wider.
South Africa had only 4 mathematicians listed and no females – the SAMS (South African Mathematical Society) lists 300 mathematicians and independent searches about black women in mathematics reveals more than 20 historically prominent women in that field.

Therefore, South Africa’s cultural and knowledge contribution to the world looks like this: ‘Soccer Player’ seems to be leading.

![Births Over Time](image)

Notably missing from Pantheon was Euphemia Lofton Haynes, the first African American woman to hold a PhD in math as well as mathematician Fern Hunt.

This begs the question how many women were mislabeled? How much of the rich cultural history of other nations is buried under prejudicial views? How much of the contributions of different races, genders, and ethnicities are not considered in the shaping of science, arts, humanities, and technology? How many groundbreaking discoveries of men and women from indigenous communities are hidden behind these ragged mountains in these graphs? How many of them are just a line when they should form the main curve?

**Conclusion**

The answer unfortunately is not very encouraging. Hedy Lamarr, an inventor, falls under the label of an actor. Not even her European heritage could come to her aid, to give her rightful place when it came to scientific contributions by women. Without her, there would be no Bluetooth.
Yet gender bias seems to trump ethnicity bias. There are 323 inventors in their database and only 8 are female. And there are no Black inventors and no Black women inventors. Yet we know there are 14 prominent Black women inventors, like Marie Ban Brown, the inventor of the first home security system, who was also a nurse.

AI trained on datasets that have erased the contributions of women and minorities cannot magically see and value what is not there. It will fill in the gaps and make flawed inferences in the same way as humans.

A broader, deeper cultural awareness breeds relatability and acceptance among humans. In the same way, AI trained on biased, incomplete, and poorly labeled datasets will breed hostility and ingrain the existing prejudices of humans against other humans – continuing the poisonous legacy of systemic inequalities against the marginalized and the minorities.

This was evident in the areas of predictive policing against African American communities, and the use of flawed algorithms in our criminal justice systems against all minorities. And it was brought to the forefront during the pandemic with the apparent healthcare disparities in the diagnosis as well as treatment of Blacks, Asians, and Hispanics.

To train our AI systems to be effective gatekeepers of cultural context, we must train the next generation of humans to see with an ethical lens, to spot and call out the biases. In order to do that, we must widen the diversity in our cognitive pool, to include everyone in developing AI. The diversity factor must be there for all stages -- the conceptualization of ideas, the collection of data, the modeling, the development of socially beneficial use cases and the humane deployment of AI.

Just as IVOW.ai has boldly taken on the challenge of bringing cultural context to AI, there are organizations like DataEthics4All that has taken on the challenge of training the next generation of ethical AI by directly tackling the current biased AI systems in predictive policing, criminal justice, and COVID contact-tracing apps in their inaugural #Ethics4NGAI hackathon launching in September. Together, the landscape of AI is indeed getting the much-needed ethical makeover.

References

About Egyptian Queens
About Black women in history
About African Women Mathematical Society
About Indian Authors
About Indian contribution to math by World Economic Forum
About 14 Black Female Inventors
About Indian Female Mathematicians
About Indian Mathematicians
A Three-step Suggestion Path To Build Anti-bias Cultural Datasets

By Rogério Lourenco

Rogério Lourenco is an information ethnographer living in Rio de Janeiro and an IVOW Advisor for this data ideation challenge. As an anthropologist who holds a PhD in linguistics, he works with digital content research, production, and documentation. He has professional and academic experiences in discourse analysis and other language needs, like controlled vocabularies and metadata. Rogério has worked with youth community media, open digital culture, public education, games, and UX design for interactive voice interfaces.

Today’s Artificial Intelligence technologies lack capacity for dealing with one of the most basic human characteristics: culture. The use of symbols in order to organize our collective psychological experience. From the start, within these symbols, language is the most fundamental.

The objective of IVOW’s ideation challenge “was to get public data sources of women throughout history and suggest how that data could be used to gain new insights for AI products and solutions with a focus on women.” The results from the challenge made clear that our collective organization of psychological symbology lacks a broader cultural landscape. It brings a narrow comprehension of needs that customers have in order to see products as part of their social reality.

There are ways to make this gap less present, or even fill it in such a manner that preserves the user experience at the cultural level. It is important to develop tools for promoting diversity, therefore the necessity to build sophisticated AI that recognizes more than traditional features of customization, like the data points, so to speak.

Traditional digital data points, such as language, geographical localization, personas or psychological typologies, among others, do not include “culture” as a standalone variable. The actual standard for user experience to analyse and understanding are the Big Five personality traits: extraversion, agreeableness, openness, conscientiousness, and neuroticism. By doing so, products are generalized to fit the “for all” mantra of marketing and the “universal affordance” of design criteria. Although it brings usability to a sufficient level, nevertheless, it also brings gaps that go from color of skin, gender roles, to dietary specificities. Along with this, such an approach brings problems like the ones we see in privacy or adaptability to diverse backgrounds.

As IVOW’s blog about the first impressions of the challenge points out: “Current classification of gender, ethnicity, and race in Wikipedia is flawed and lacking.” Here I will delineate a suggestion of prospecting possible solutions based on the concept of algorithmic path, instead of adopting the one of structure (chosen by the majority of challenge candidates). By proposing three steps of exploratory stages, I suggest we can look to human languages as a blueprint for construction of a dataset based not on content, which is biased, but by the cultural bases it is constructed on.

Given that there are already constructed Wikipedia datasets, those can be repurposed to meet language bases from the ISO 639 - codes dataset (which allows commercial use) and build from that.
STEP 1: Information on ISO 639-3

There are around 7000 languages in the world. For all those, there is a unique code for identification. Here I quote the ISO 639-3 as “a set of codes that defines three-letter identifiers for all known human languages. At the core of ISO 639-3 are the individual languages accounted for in ISO 639-2.” These languages were collected from different sources, but primarily from Ethnologue (15th edition) produced by the Summer Institute of Linguistics, as the only official institution authorized by the ISO, the International Organization for Standardization in this matter. By starting with only the languages without direct relation to Wikipedia, we then look at the other possible datasets in order to build the next layer of information.

STEP 2: Language and datasets structures

The ideation challenge had Wikipedia content as the main source. The Wikidata dataset is organized by properties, as a keyword allowing the categories to be grouped. Since Wikidata datasets also have a mapping of language to territories page, based on the Ethnologue, it faces the same kind of barriers IVOW has, on a larger scale. For the first, the necessity of filling the gap with articles in general, for the later, the localization of information on women inside of what is already organized by Wikipedia. As noted with the challenge results, information on women, not only on Wikipedia, is poorly collected, organized, and published. Although it can be observed that in the case of Wikidata, there is a property list related to women and feminism (Q28649030). Nevertheless, entries still barely catalogued.

As Wikipedia categories are organized to cover the whole spectrum of human knowledge, IVOW needs to characterize itself inside its own effort as a gender specific task. So the need to look at categories has the double task of finding women inside the datasets, but also recognizing the relevance of what is found.

The Wikipedia Cultural Diversity Observatory

Although there is more on the internets than Wikipedia data to search for, information on women regarding expressive cultural figures, such effort must start from Wikipedia dataset. This is comprehensible once we look at the way its dataset is organized. As noted above, there is already a link between languages and geographical pages with content. What is needed is to link the gender specific content to these pages, since they are designed to have general knowledge. One way to do so is to look at culture-specific categories that Wikipedia has:

2.1 Wikipedia Cultural Diversity Observatory/Languages

This base from ISO 639-3 would be useful for other data beyond Wikipedia, but also would fit into it. In this manner, by scraping content that adjusts to the world languages, but also to the cultural categories, it is possible to gather a variety of other sources with a standard methodology.

2.2 Wikipedia Cultural Diversity Observatory/Cultural Context Content (CCC)
Once the number of possible recipients (ISO 639-3) is given, the CCC is a second layer available to use. Since the datasets were constructed by using Wikipedia content, the CCC criteria could be applied to guide what should be considered relevant to the cultural context.

In terms of path, it should be also possible to do the other way around, and not start from the three main ones: i) gender, ii) ethnicity, and iii) race classification. Instead, to begin with language (ISO 639-3, and then look for the distributed list of vital topics and articles. The first deals with broader concept areas and later with specific thematic fields. For IVOW's purpose, to find relevant data on gender, ethnicity, and race, there are several categories that could be adapted but are not indexed as such.

Step 3: Disciplinary Relations

As mentioned, both Wikipedia and the output from the ideation have gaps that need to be filled, even though each one has its specific reason. For IVOW, in order to be feasible to function in the future digital agents with Cultural IQ, the incorporation of some categories from the Wikipedia dataset could then be expanded based not only on its content, but on the previous layers of language. There are various well known linguistic and ethnographic techniques (1) (2) for eliciting data that could be applied to construct specific datasets. From those, the subjacent work would be to connect with other data points, such as the ones above mentioned, in order to assure the quality of what is gathered.

Besides this dataset, many others could eventually be adapted to organize scraping paths, not from its content, but from its categories / language parameters. Therefore, by looking at scientific procedures and resources of disciplines like Anthropology, Archaeology, Linguistics, or History, instead of searching mainly in the Wikipedia data available, a more robust corpus could be constructed by autonomous algorithms created for future investimes from IVOW.

References


Max Planck Institute for Evolutionary Anthropology: Typological tools for field linguistics, 2020.

5. The Role of Testers and QA Leaders

In late October 2019, IVOW travelled to Dublin, Ireland, to present at the Quest for Quality Summit sponsored by Comtrade Digital Services. Software engineers, testers, data scientists, and AI specialists from as far as New Zealand came together to engage around the Quest for Quality in AI, data, and testing.

It was at this summit that we soft-launched the Women in History Data Ideation Challenge and talked about why it’s necessary to understand what data is out there for historical context and for the preservation of culture with an emphasis on the role of women.

One of the keynote speakers was Tariq King, a thought leader in quality assurance and testing. Tariq said that testing has typically lagged behind development, but now it’s almost as if that problem is being flipped when it comes to innovation with AI. He noted that if AI is to be successful in any industry, “it will require solid validation and verification practices to make sure these systems are functional, safe, secure, performant, culturally-aware, law-abiding, and ethical. This is where a community of testing professionals becomes invaluable.”

Today, Tariq is the chief scientist at test.ai where he leads research and development of their core platform for AI-driven testing. In July, 2020, IVOW reached out to Tariq and Jason Arbon, CEO of the company, to look over the early results from our data ideation challenge. Tariq and Jason shared their work on Quantifying the Risk of AI Bias and Testing and AI and Bias, respectively. Tariq gave us permission to include this excerpt from his insights:

The World is Filled with Bias. A quick search on bias reveals a list of nearly 200 cognitive biases that psychologists have classified based on human beliefs, decisions, behaviors, social interactions, and memory patterns. Certainly, recent events stemming from racial inequality and social injustice are raising greater awareness of the biases that exist in the world today. Many would argue that our social and economic system is not designed to be fair and that the system is even engineered in a way to marginalize specific groups and benefit others. However, to be able to improve a system that is unfair, we first have to be able to identify and measure where, and to what degree, it is biased.

Since the world is filled with bias, it follows that any data we collect from it contains biases. If we then take that data and use it to train AI, the machines will reflect those biases. So how then do we start to engineer AI-based systems that are fair and inclusive? Is it even practical to remove bias from AI-based systems, or is it too daunting of a task?

It is evident that the testing community has a role to play in the engineering of trustworthy AI-based systems. More specifically, on the issue of AI bias, testers have the necessary skills to directly contribute to preventing, detecting, and reducing unwanted bias in AI systems. As a starting point, we developed three heuristic-based artifacts to support testing AI systems for unwanted bias: (1) Testing Mnemonics Based on a Survey of Bias and Fairness in Machine Learning, (2) Testing AI and Bias Questionnaire Checklist, and (3) AI Bias Risk Assessment Tool. The testing mnemonics are displayed graphically on the AI bias cycle as follows:
Technology giants Google, Microsoft, and IBM have all developed open-source toolkits for testing and mitigating unwanted bias in AI systems. Wondering how to get involved? Here are some ways we believe folks can have an impact on testing AI and bias:

1. Devising practical testing methods and processes for preventing and detecting unwanted bias in datasets.
2. Developing new coverage models and static/dynamic analysis tools for validating AI and ML systems.
3. Mastering and contributing to the existing toolkits for measuring fairness and detecting/mitigating unwanted AI bias.
6. Imbuing Current AI Technology With Cultural Intelligence

Chatbots are taking the internet by storm. According to Forbes, 50% of Americans used a chatbot to interact with a brand in the past year, and 64% of organizations plan to increase their AI investment over the coming year.

With Artificial Intelligence quickly becoming a part of our daily lives, it is vital for AI to represent all the people who use it. Unfortunately, as our report shows, current AI technology and the datasets underlying them are lacking in diversity and cultural knowledge, which can play a role in perpetuating harmful historical biases. By adding diversity and culture, we can make AI not only intelligent, but also move it toward being wise and just.

Adding cultural richness to AI is also important for preserving our traditions for future generations. Just as the printing press and the world wide web revolutionized the way humans share information, AI is now reshaping how people learn and interact with one another. While nothing beats having a human teacher, AI chatbots can make cultural knowledge available 24/7 to millions of people simultaneously around the world.

7. Sina Storyteller

IVOW's vision for our culturally intelligent voice assistant Sina (SEE-na) is to create an AI character with deep interest in human culture. Sina Storyteller by IVOW is a conversational AI with a unique appreciation for global cultures. Sina (available in prototype via Google Home) can tell you stories about great women in history, ancient traditions, and vibrant food cultures from around the world.

In October 2019, at the same Quest for Quality summit in Dublin where we met Tariq King, we also met Pramod Bommisetty, lead QA engineer for Soul Machines based in New Zealand. Our focus on cultural intelligence in AI was intriguing to him, and we followed up. Together with several other Soul Machines team members we came together in July 2020 for a limited engagement.

Sina is inspired by multiple sources of wisdom and represented by a dove, an ancient symbol synonymous in many cultures with storytelling and unity. It is intended that Sina can take on many forms and as a result of our new collaboration with Soul Machines, Sina takes the form of three digital humans. Soul Machine’s Autonomous Animation Platform is powered by the world’s only “digital brain” and their digital people are astonishing. Our work to bring a cultural lens to this world of exponential AI is exciting. As a messenger of human stories, having an expressive human face helps Sina advance her mission of connecting with people through culture.
You can also access a prototype of Sina on Google Home by asking for “Sina Storyteller.” Next time you are in your kitchen, ask Sina to tell you a story about the traditional foods of the ancient festival of Diwali; she also knows recipes from the Navajo and Cherokee nations. Or ask her how machine learning can help us keep ancient cultural knowledge alive.

Automation does not need to be synonymous with the loss of ancient wisdom and cultural knowledge. With Cultural Intelligence, AI can help, rather than hinder, humans in our quest to tell our unique stories.
8. Women in AI Global Summit

On September 8, IVOW will be part of two panels from 5-7pm ET as part of the Women in AI Global Summit. These are interactive sessions where we will share results from our data ideation challenge with Topcoder, and demonstrate how we integrated cultural intelligence into our current AI product. Sponsors and community partners for this summit include Jessica Davis at Microsoft; Coach Kathy Kemper of Institute for Education; Carrie Jaquith of AXA_XL; and Erica Lloyd and Holly Peck of Soul Machines.

**September 8, 2020: 5-6pm ET**

Title: Cultural Data: Richer AI

Short Description of our panel: How can the creation of machine-ready cultural datasets inspire richer AI solutions, bring visibility to global communities, and diversify the tech workforce? Panelists will include: Carrie Jaquith, global head of digital enablement at AXA XL; Coach Kathy Kemper CEO of the Institute for Education; Jessica Davis of Microsoft; and Shaun Chavis, founder and CEO of LVNGbook. Moderator is Davar Ardalan of IVOW.AI.

**September 8, 2020: 6-7pm ET**

Title: Cultural AI and Brands

Short Description of our panel: Join our interactive panel of storytellers, scientists, and conversational AIs to consider how chatbots can become culturally diverse and deeply inclusive. Panelists will include: renowned AI and storytelling scientist Nasrin Mostafazadeh; Erica Lloyd and Holly Peck of Soul Machines; and Nikki McLay of IVOW.AI. We will also feature our digital storyteller Sina and digital people from Soul Machines. Moderator is Davar Ardalan of IVOW.AI.
9. Conclusion

Our Women in History Data Ideation Challenge proves that:

A. we need to create more free and open source datasets featuring women in science, engineering, technology, arts and culture;
B. we need to clean-up current datasets and tag women based on their contributions in history and not a generic label (companion);
C. we must create machine-ready datasets focused on gender and culture so that AI products and services can be more inclusive.

Exploring this new frontier of “Cultural Intelligence” will require a new data ecology, inclusive datasets, and traditional storytelling solutions.

For those of us marching into automation, we must understand that much of the current data ecology is incomplete, flawed, racist, misogynistic, and derogatory. This lack of diverse data and prevalence of biased data hampers innovation and causes significant setbacks for future AI products and solutions.

There is nothing more important to the story of our future than protecting the story of our past. Particularly for me as an African American, for a group of people whose history has been erased, preserving the truth of our stories will ignite who we are to become. Thank you IVOW for holding the torch.

— Lisha Bell, Silicon Valley business leader and member of Pipeline Angels

It’s easy to document the prevalence of bias in today’s digital environment. Algorithms that are not trained by broadly diverse data will discriminate against women and people of color or other groups in ways that impact their employment, health care, housing, and financial well-being. There are facial recognition programs that can’t “see” Black faces; databases of images that have been tagged with misogynistic and derogatory terms; and women often find themselves steered to lower-paying jobs or inappropriate advertising. Here is one more example from Pantheon: in a search for American women in all science & technology fields, we found 71 females (one was actually male); only four were African American.

Recently, MIT announced that its 80 million “tiny images” dataset would be taken down immediately. According to the authors of the dataset, “Biases, offensive and prejudicial images, and derogatory terminology alienates an important part of our community -- precisely those that we are making efforts to include. It also contributes to harmful biases in AI systems trained on such data.” The statement further acknowledged that the presence of such prejudicial data hurts efforts to foster “a culture of inclusivity in the computer vision community.”
The technology is progressing in leaps and bounds, and every day of not making things right translates into an impact of multiple years in AI development terms. We need to identify and address these biases now. Cultural intelligence in AI is no longer just a feel-good, politically correct thing to do. It is a critical aspect of AI development in a world that is becoming more and more technology-reliant, yet increasingly polarized.

– Jeanne Lim, founder and CEO of Nama Institute and ex-CEO at Hanson Robotics

Future automated decisions cannot be made based on current data ecosystems. The next generation of developers, coders, engineers, and other computer scientists need to be cross-trained on the promise of diverse and deeply inclusive AI/ML.

As we look into that future, this is the environment that is needed for success:

**Educational Institutions** that create cross-campus learning hubs where teams of data scientists, cognitive scientists, computer scientists, cultural anthropologists, social scientists, historians, artists, and musicians come together to educate each other on how to apply AI/ML tools.

**Quality Assurance & Testing Units** where brands work closely with their Testing and QA teams to guarantee that what they are serving their customers is not only user friendly from a tech standpoint, but also culturally relevant.

**Data Architect Hubs** where diverse data scientists will come together in daily or weekly forums to accelerate AI innovation by making sure their community’s data ecosystem is growing, their taxonomy is relevant and being trained ethically. These hubs will also solve challenging NLP extraction problems and share knowledge on how these findings can be applied to other use cases. This can be done on an honor system to ensure that people are contributing relevant and accurate information.

The largest gap in designing new frontiers for media and communications is the lack of social sciences and deep learning from our heritage and traditions. So far, we have gathered enormous amounts of data while leveraging AI algorithms to foster innovation in self-driving cars, cybersecurity, and healthcare applications. We now need to focus on creating better data ecologies that account for human culture and history.

---

**We are grateful to our sponsors, collaborators and advisors:**

Topcoder; Comtrade Digital Services; KiwiTech; Women in AI; #WaiAccelerate; WeWork Labs; Aprajita Mathur of Guardant Health; Jessica Davis of Microsoft; Erica Lloyd and Holly Peck of Soul Machines; Jason Arbon and Tariq King of Test.ai; Coach Kathy Kemper of Institute for Education; Jenn Bonine, CEO of AI Appstore, Inc; AI prodigy Kashyap Murali; and renowned scholar Laleh Bakhtiar of Kazi Publications who have been on this journey with us.