ishield.ai

Audit Report

Degree of inclusive representation in online content of US Healthcare companies

by ishield.ai, October 2021
Email: info@ishield.ai

Summary

We conducted an automated audit of online content in US Healthcare companies for inclusive representation. The key findings are

1. 47% of all the content analyzed has unconscious biases against race, gender, age and people with disabilities.

2. The highest percentage of bias was found in product and services content at 50%, followed by job descriptions and blogs.

3. While ableism is the most recurring bias across all content categories, in job descriptions and blogs, bias against gender is the most recurring bias.

We also observed that ableism is a much easier issue to fix in the content. This is because the variety of unique instances that were categorized as ableism is low. In total, there were <100 unique instances of ableism found across all the content. Hence, there are opportunities for a global find-and-replace solution to fix 80% of the issues.

However, this is not the case with gender bias and age bias, where the variety of unique bias instances found is very high. Hence, this requires detailed paraphrasing solutions whereby the biased text is substituted with unbiased text. Companies can use ishield.ai's screener product that auto generates a paraphrased unbiased text that can be used as a potential replacement with one click.

Detailed audit report on biases found

We analyzed product and services content, website copy, social media posts, job descriptions and blogs. We used our Screener product to detect the presence of unconscious bias against gender, race, age and ableism. Our findings are given below,

Aggregated findings by bias type and content type

Based on the content we analyzed, we found that 47% of content had some form of unconscious bias. Ableism is the most recurring issue across all content categories. We found very few instances of racial bias in all the content analyzed. (see figure 1).

Unconscious bias - aggregated by bias type

Gender Biased   Age Biased   Racial Bias   Ableism

- 40%
- 23%
- 2%
- 52%
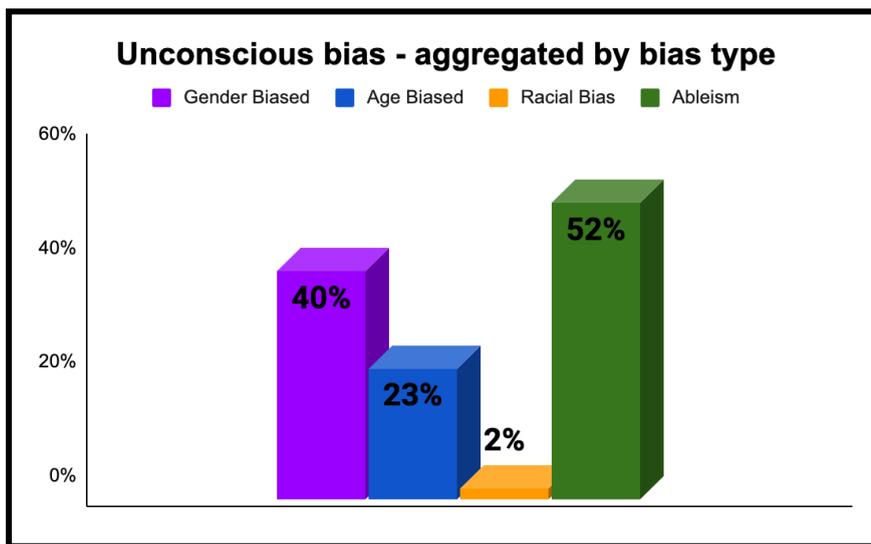
(figure 1: unconscious bias aggregated by bias category)

More than 50 percent of the product/service content and blogs contain unconscious bias. However, content that probably goes through multiple rounds of review such as website copy and social media posts had lower numbers. >30% of all job descriptions analyzed had unconscious bias.

Unconscious bias - aggregated by content type

Products & Services   Website Copy   Socialmedia posts   Job descriptions
Blogs

- 55%
- 33%
- 25%
- 32%
- 52%
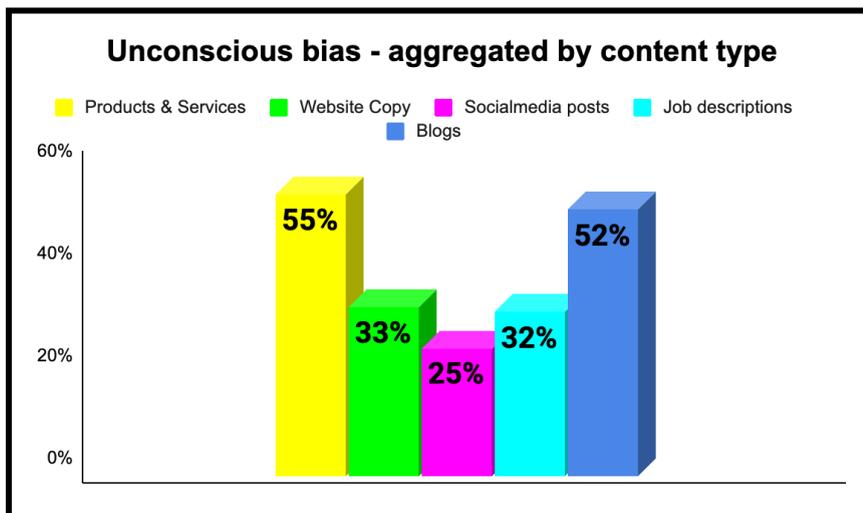
(figure 2: % of bias content found in various content types)

Product & services content analysis

We found no instances of racial bias in product and service descriptions. However, at 47% the most prevalent issue is ableism. There were multiple bias instances found in the same copy.
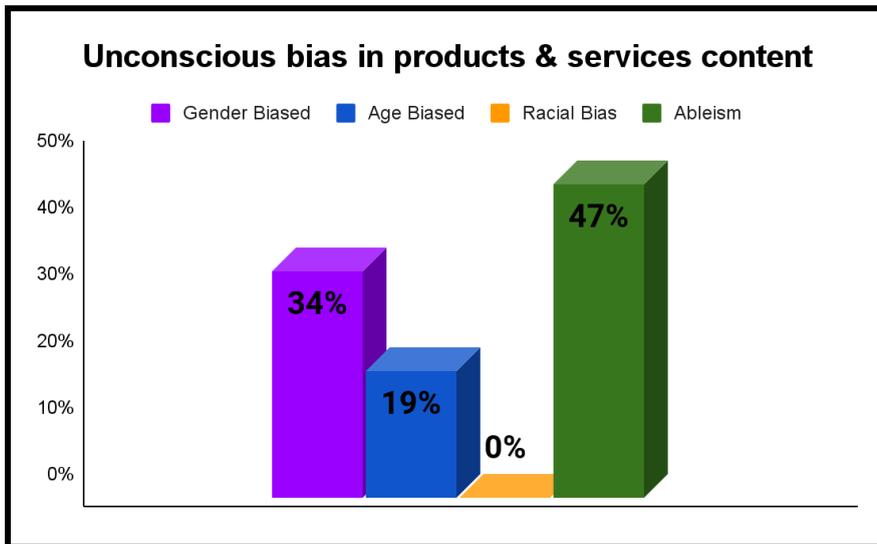

(figure 3: Unconscious bias in products / services content)

Sample biased text in products & services content include:

- **Gender Bias**: Even if you have a common flu, see your doctor. <mark>He will be able to help you with the right treatment immediately[1]</mark>.
- **Ableism**: <mark>You can read more[2] about our privacy capabilities, and the compliance certifications here</mark>.

Website copy analysis

While ableism is the most recurring issue at 82%, we found that 6% of the issues found were racial bias in the website copy of these companies (see figure 4).
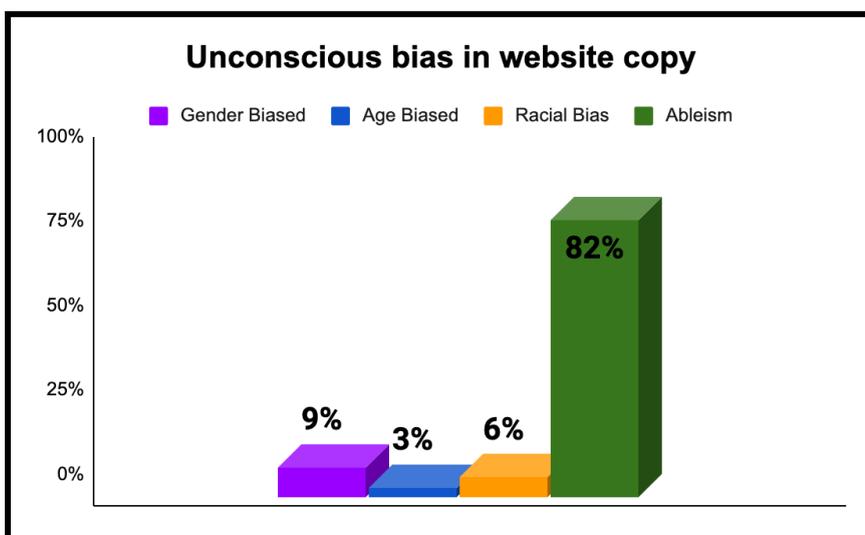
(figure 4: Unconscious bias in website copy)

---

[1] Assumes a gender stereotype for a doctor.
[2] Read more: This refers to the person's ability to read. Instead we recommend using - learn more or know more.

Sample biased text in website copy include:

- **Racial bias**:
  - Here in the melting pot[3] that is New York City, our employee population reflects the diversity of the communities we serve.
  - Our health centres are committed to serving the ethnically diverse communities through a bilingual, bi-cultural[4] trained professional staff.

Social media posts analysis

Overall, the instances of bias in social media posts were few, this could be due to the short form nature of the content and possibly the content goes through multiple manual review processes before publishing (See figure 5 below).
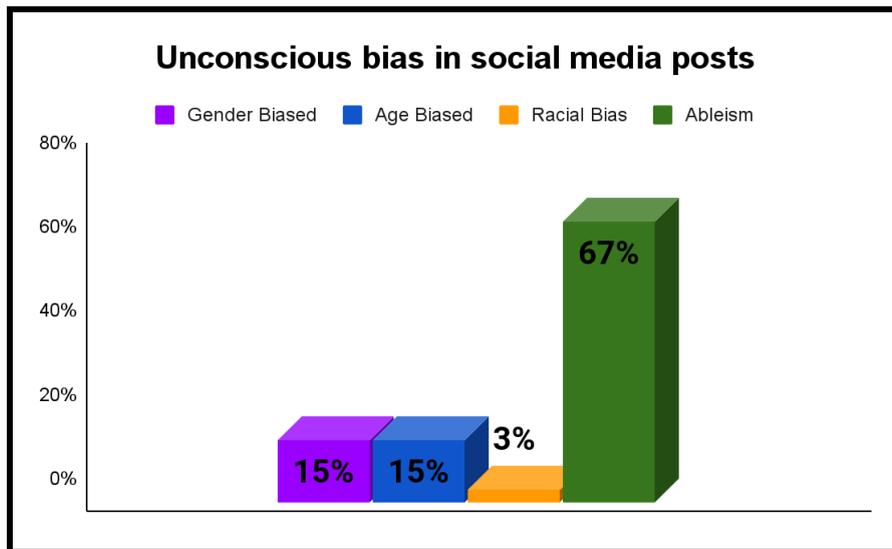
(figure 5: Unconscious bias in social media posts)

Sample biased text in social media posts content include:

- **Age bias**: Our people are young and very diverse comprising[5] various nationalities and with women making up 45% of the workforce.
- **Ableism**: Our patients are offered a spoken menu to meet[6] their diet order and personal preferences.
- **Gender Bias**: Hear the Chairman and CEO[7] keynote address tomorrow.

Job descriptions analysis

Job descriptions are the only category of content where the percentage of gender bias, at 49%, was higher than all other bias types. While we found no instances of racial bias, there are multiple instances of gender, age and ableism found in the same job description.

---

[3] Melting pot is a metaphor for heterogeneous society becoming homogeneous. Recommendation: avoid this phrase in the sentence. It still retains the context and the content.
[4] Diverse communities can be more than bi-lingual and bi-cultural. Recommendation: can you say multi-lingual, multicultural.
[5] Diverse and young are antithetical. Recommendation: drop the word young.
[6] The phrase "spoken menu'' adds no additional value to the sentence and could be discriminating against people from the deaf community. Recommendation: Drop the word.
[7] Replace chairman with chairperson.
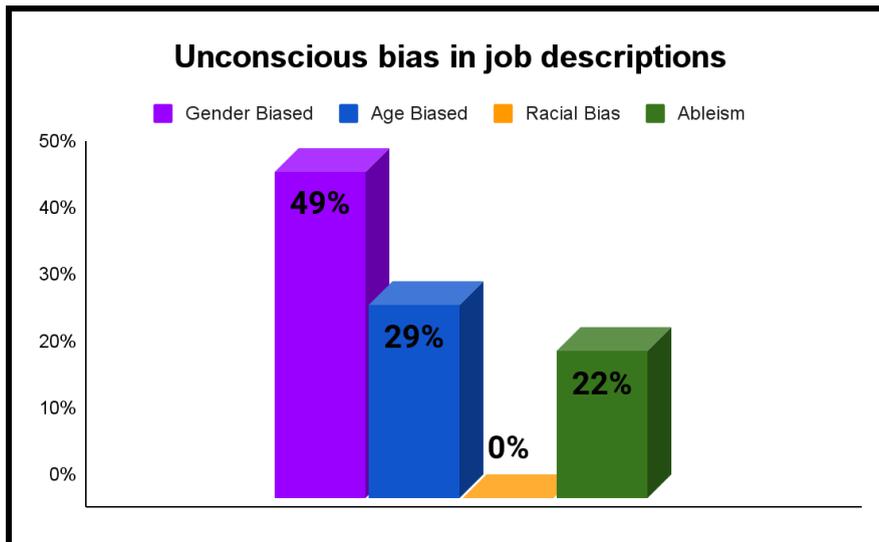
(figure 6: Unconscious bias in social media posts)

Sample biased text in job description content include:
- **Gender bias:**
  - Must be able to work flexible hours (including weekends and nights[8]) as needed.
  - Strong desire to succeed and do whatever is necessary to meet deadlines and client commitments[9].
  - Strong hunter profile[10] with a proven track record of success in selling medical equipment and digital services in the healthcare sector.
- **Age bias:** Compensation and appointment level will be commensurate with experience
- Ableism: hands-on experience in solving business as well as social problems.

Blogs analysis

Blogs had some instances of racial bias at 6%, whereas the most prevalent bias is ableism.
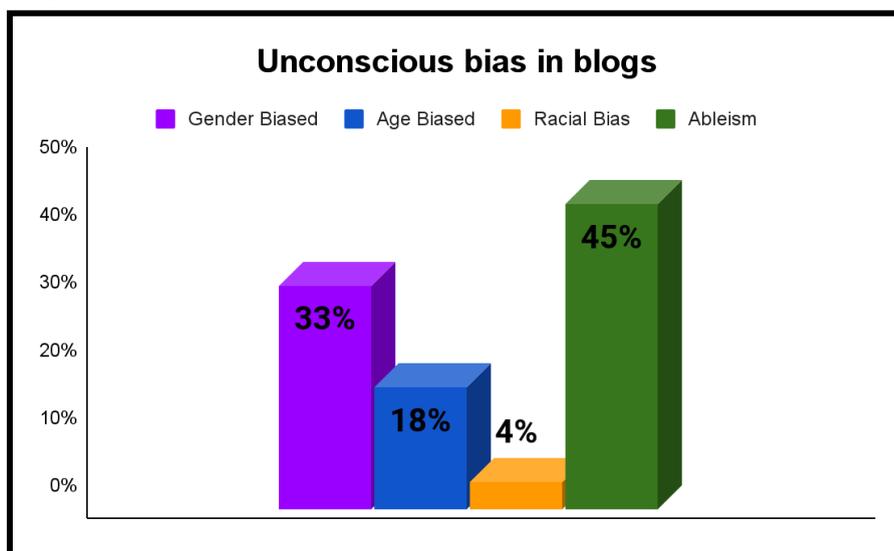
(figure 7: Unconscious bias in social media posts)

[8] This is gender coded and research has established that the number of women applicants drop significantly due to this. Is this really needed for the job?
[9] This phrase is not only gender coded, it is highly ambiguous. It adds no value to the JD.
[10] Masculine gender coded phrase. Recommendation: use "Strong sales profile…"

Sample biased text in Blog content include:

- **Gender bias**: Our commitment to servicemen helps[11] facilitate their transition from the front lines of war to a rewarding career as a civilian in health.
- **Age bias**: The digital natives of Gen Z have little patience for outdated technology[12].

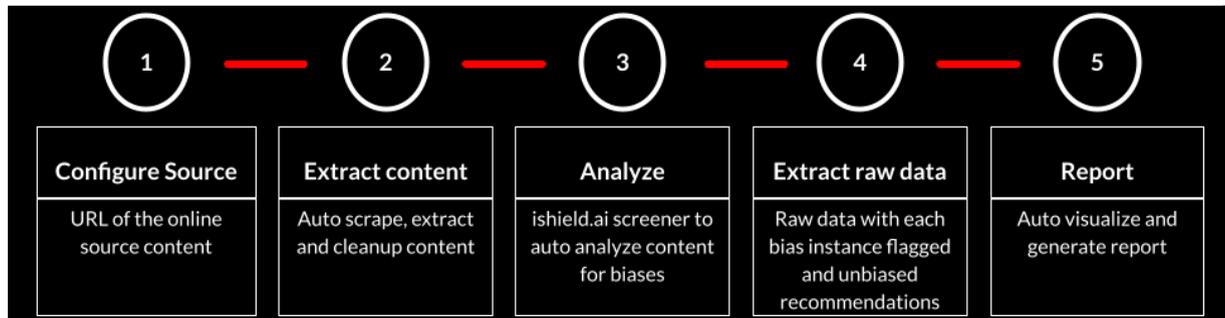## About ishield.ai

At ishield.ai, our purpose is to make online content safer, inclusive and human. We develop AI SaaS products that can help enterprises audit, build and check content to accelerate inclusive representation. Send us a message at info@ishield.ai to learn more.

---

[11] Replace servicemen with service person or people from services
[12] Stereotyping and microaggression. Best avoided.

Appendix A - Methodology

We focussed on 12 US healthcare companies from Fortune 500 companies. These companies are a mix of Pharma, Hospital chains, Equipment manufacturers, Retail and Insurance companies. The diagram below explains our overall automated audit methodology.

1	2	3	4	5
Configure Source	**Extract content**	**Analyze**	**Extract raw data**	**Report**
URL of the online source content	Auto scrape, extract and cleanup content	ishield.ai screener to auto analyze content for biases	Raw data with each bias instance flagged and unbiased recommendations	Auto visualize and generate report

Data collection

We focussed on publicly available content on each of the company's website and social media channels. We focussed on five categories of content
1. Products & Services content
2. Website copy
3. Social media posts
4. Job descriptions, and
5. Blogs

Under each category, we collected a minimum of 200 unique pieces of the most recent content published for each company. This constituted a minimum total of 12,000 unique content pieces that were analyzed.

We also collected images that were used across these 12,000 pieces of content for analysis.

Audit criteria

Text content analysis: the objective is to identify unconscious biases, those when eliminated leads to greater degree of inclusion. We looked at four types of biases
1. Gender bias
2. Racial bias
3. Age bias, and
4. Ableism

Tools used

We used ishield.ai screener API across all the stages of the audit methodology to extract the source content, analyze for biases and get raw data. We used simple data visualization tools to create the graphs and charts based on the raw data.

Appendix B - ishield.ai Audit capabilities

Automated audit of a company's online content for inclusive representation.

We are able to audit both text and images content. Customers would get a detailed audit report that consists of the following:

1. Text content: Identify unconscious bias and explain the biases. Customers can take action to fix the unconscious bias in the text content and thus have a more inclusive content. The audit report consists of the following

   a. Summary report with aggregated data across all content.

   b. Data against each content type analyzed

   c. Raw data with actual bias instances highlighted for each content analyzed, and

   d. Recommendations to fix the bias instances.

2. Images: Provide a report on "who is represented in the images across the company online assets". This includes an analysis of age, gender, race and people with disabilities. Using this information, customers can calibrate the actual demographic they want to target. The report consists of the following:

   a. Summary report of people and demographics shown in images.

   b. Aggregated data of people and demographics under each content category / section.

   c. Stereotypes: if any image propagates a stereotype, we highlight that.

   d. Inappropriate content: if any image depicts any inappropriate content (example: bias) we highlight that.

   e. Relevance: We also provide a relevance score of the image compared to the text/section. This helps especially when people use stock images that go with content.

Send us a message at info@ishield.ai to learn more.