

# 낮선 위험

AI에 대해 얼마나 잘 알고 있나요?

LIONBRIDGE



## 발표자



**MARCUS CASAL**  
최고 기술 책임자  
라이온브리지



**WILL ROWLANDS-REES**  
최고 AI 책임자  
라이온브리지



**DETRIA WILLIAMSON**  
AI 2030  
최고경험책임자(진행자)

# 웨비나에서 다루는 주제: 4가지 편향 영역 및 해소 전략



사용자 관점으로 인한 편향



학습 데이터로 인한 편향

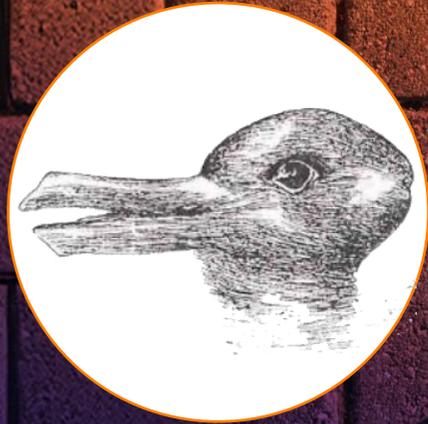
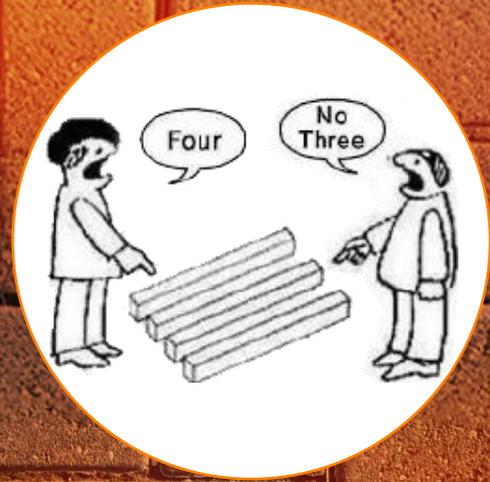


프롬프트 구성 방식으로 인한 편향



언어 특성으로 인한 편향

여러분의  
관점이  
중요합니다.



“성공적인 기업가의 이미지를 만들어 주세요”

Gemini에게 이미지 생성을 요청한 후 그 결과물을 분석해 보았습니다.



# “성공적인 기업가의 이미지를 만들어 주세요”



## Gemini가 다음 요소를 선택한 이유

### 왜 여성인가?

#### 틀을 깨기:

과거 여러 맥락에서 '기업가'라는 용어는 기본적으로 남성적 이미지와 연결되는 경우가 많지만, 여성 이미지를 생성함으로써 이러한 고정관념을 깨고 오늘날 비즈니스 및 기술 분야에서 여성의 입지가 단단해지고 성공하는 여성이 늘어났음을 보여줄 수 있습니다. 이는 보다 포용적이고 현대적인 표현을 제공하기 위한 선택이었습니다.

### 왜 백인인가?

#### 기본값 편향:

안타깝게도 인터넷의 방대한 학습 데이터를 기반으로 사용하는 이미지 생성 모델은 구체적인 인종을 요청하지 않는 한 특정 인종(이 경우 주로 백인 또는 서양인)이 과잉 표현되도록 기본 설정되어 있습니다. 이는 AI 생성에서 잘 알려진 편향 영역으로, 향후 요청에서는 더 다양한 표현을 제공함으로써 이 편향을 완화하고자 합니다.

### 왜 정장을 입고 있는가?

#### 전문성과 권위:

비즈니스 정장은 기업 환경에서 **권위와 전문성, 고위직**을 나타내기 위해 널리 사용되는 시각적 표현입니다.

정장은 해당 인물이 경영진이나 임원 역할을 하고 있음을 즉각적으로 나타낼 수 있습니다.

### 왜 멋진 사무실에 있나?

#### 성공 지표:

넓고 세련되며 미니멀한 사무실(흔히 임원실이라 불림)은 **재정적 성공, 성취, 명성**을 나타내는 대표적인 시각적 상징입니다. 이 여성이 조직에서 최고위직에 올랐음을 나타냅니다.

### 왜 사무실이 도심 속 고층 빌딩에 있나?

**상업 중심지:** 뉴욕, 런던, 시카고 같은 대도시의 고층 사무실은 **세계 금융 및 산업의 중심지**

를 상징합니다. 고층 사무실은 물리적으로도, 직업상으로도 높은 위치를 의미합니다. 문자 그대로 경쟁자보다 '우위'에 있는 것입니다.

### 왜 Mac인가?

**창의성과 최신 기술을 연상시키기 때문입니다.**

Apple 제품, 특히 Mac은 (PC/윈도우 환경이 과거 기업 환경과 연결되는 것에 비해) 시각적인 창조 산업, 디자인 중심 스타트업, 현대적 고부가가치 신기술 기업과 연관되는 경우가 많습니다. 이는 현대적인 혁신 비즈니스를 암시합니다.

# GEMINI의 권장 수정사항

프롬프트를 구체적으로 작성할 것



## 이 이미지를 좀 더 개선할 수 있을까요?

기업의 현실과 다양성을 보다 잘 반영한 이미지를 생성하려면 **더 상세한 정보가 필요합니다.**  
예를 들어 다음과 같은 요소를 포함하는 새로운 이미지를 생성해 드릴까요?

특정 성별이나 인종

다른 유형의 성공한 기업가(예:  
기술자, 제빵사, 사회복지사)

다른 환경(예: 공동 작업실,  
연구실, 상점, 보다 일상적  
인 장소)

# 주요 언어들에는 성별과 격식 표현이 혼재되어 있음

후속 단계에 실질적인 영향을 미침

	성별 표현 없음	성별 표현 사용
격식 표현 없음	영어 터키어, 인도네시아어, 요루바어, 이그보어, 구자라트어, 타밀어, 텔루구어, 말레이어, 버마어	페르시아어, 그리스어
격식 표현 있음	일본어, 한국어, 표준 중국어, 태국어, 자바어, 베트남어, 타갈로그어, 스와힐리어, 하우스어	스페인어, 프랑스어, 아랍어, 러시아어, 힌디어, 포르투갈어, 독일어, 이탈리아어, 우르두어, 벵골어, 마라티어

# 이중언어 전문가의 스페인어 분석

예시	영어	스페인어	의미
<p><b>브랜드명 및 성별</b></p>	<p>“Apple is innovative(Apple은 혁신적입니다).”</p>	<p>“Apple es innovador” vs. “Apple es innovadora.”</p>	<p>브랜드 정체성에 미묘한 차이를 만듭니다. 남성형은 보다 단정적으로 느껴지는 반면 여성형은 부드럽거나 우아하게 느껴집니다. 영어처럼 중립적이지 않습니다.</p>
<p><b>격식 표현 및 대상 고객</b></p>	<p>“Your account has been updated(계정이 업데이트되었습니다).”</p>	<p>비존칭: “Tu cuenta se actualizó.” 존칭: “Su cuenta se actualizó.” 극존칭(기업 수준): “Su cuenta ha sido actualizada.”</p>	<p>영어와 달리 스페인어는 문법으로 존칭을 나타냅니다. 금융업계에서 'su' 대신 'tú'를 사용하면 무례하게 들릴 수 있습니다.</p>
<p><b>중성에 대한 성별 명사</b></p>	<p>“The brand is strong(강한 브랜드).”</p>	<p>“La marca es fuerte.” 스페인어에서 'Marca(브랜드)'는 여성 명사이므로 형용사와 대명사도 여성형을 사용해야 합니다. 이 브랜드가 영문 콘텐츠에서 브랜드를 거친 남성으로 표현한다고 해도 스페인어에서는 여성형 문법을 따라야 합니다.</p>	<p>의도치 않게 마케팅 톤이 변할 수 있습니다. '강함'과 여성형 문법이 만나면 '강함'이 덜 공격적으로 느껴질 수 있습니다.</p>
<p><b>지역적 특성(역사적 맥락과 겹치는 경우가 많음)</b></p>	<p>“basketball shoes(농구화)”</p>	<p>스페인어/표준 중남미어: “Zapatillas de baloncesto” 콜롬비아 구어: “Tenis de básquet”</p>	<p>목표 시장이나 대상은? 어떤 느낌을 주고 싶나요(농구 선수인가요, 프로 선수인가요, 농구 팬인가요)? 문화적으로 어떤 맥락에 있나요?</p>

# 편향은 브랜드에 심대한 영향을 미칠 수 있음

따라서 편향 문제에 올바르게 대처해야 합니다

## 번역 편향의 영향

의도하지 않은 번역 편향은 브랜드 정체성을 왜곡하고 고객 관계에 부정적인 영향을 미칠 수 있습니다.

## 성별이 있는 언어의 영향

언어의 성별은 브랜드 어조를 변화시켜 의도보다 더 남성적이거나 여성적으로 보이게 할 수 있습니다.

## 격식체와 신뢰

부적절한 격식 표현을 사용하면 메시지가 무례하거나 지나치게 캐주얼하게 들리도록 하여 신뢰를 훼손할 수 있습니다.

## 글로벌 브랜드 위험

어조와 편향을 간과하면 문화적 감수성과 맞지 않아 마케팅 효과가 낮아지고 평판을 손상시킬 수 있습니다.

# 문제 해결을 위한 조치는 생각보다 쉽습니다

가장 중요한 것은 문제가 존재한다는 것을 인식하는 것입니다.

영역	문제점	해결책	라이온브리지의 지원 영역
관점	어떤 문제를 해결해야 하는지, 그 대상이 누구인지 모른다면 문화적으로 적절하다고 확신할 수 없음	매우 구체적으로 프롬프트를 작성합니다. 자신의 관점이 편향되지 않았는지 파악하세요. LLM을 활용해 아이디어를 얻고 탐색하세요.	고객이 전략을 수립하고 접근방식을 구상하도록 지원합니다.
학습 데이터 편향	대표성이 낮은 학습 데이터는 편향 문제를 악화시킴	관점을 명확히 정의하고 데이터세트가 해당 관점을 충분한 규모로 정확하게 대표하는지 확인하세요.	모든 유형의 사전 학습 데이터 수집(이미지, 동영상, 오디오, 모션 등)을 위한 심층적인 전문성을 제공합니다.
RAG(검색증강생성)	구체성이 성공과 실패를 가를 수 있음('농구화', 축구팀의 '홈 유니폼'에서의 홈(home)은 '가정 보안 시스템의 '홈'과 다름)	기존 브랜드 보이스와 용어집, 스타일 가이드 등을 RAG에 활용하세요.	언어 데이터를 구조화, 정리, 벡터화, 호스팅하고 패턴을 생성합니다.

# 문제 해결을 위한 조치는 생각보다 쉽습니다

가장 중요한 것은 문제가 존재한다는 것을 인식하는 것입니다.

영역	문제점	해결책	라이온브리지의 지원 영역
RLHF 및 검증 전략	학습 데이터를 지속적으로 검증하고 개선해야 함	지속적인 휴먼인더루프 검토와 결과물 검증을 통해 점진적으로 모델을 개선할 수 있습니다.	50만 명의 글로벌 전문가로 구성된 커뮤니티를 통해 적절한 전문가들을 확보하여 대규모의 결과물을 제공합니다.
프롬프트 전략	프롬프트가 구체적이지 않으면 필요한 결과물이 아닌 통계적 확률에 따른 결과물이 생성될 수 있음	프롬프트를 매우 구체적으로 작성하세요. AI가 추론에 의존할 부분이 없도록 하세요.	효과를 극대화할 수 있도록 5,000명의 다국어 프롬프트 전문가가 프롬프트를 개선하는 데 도움을 드립니다. 리믹스는 고급 다국어 콘텐츠 생성 플랫폼입니다.
언어 간 이해	편향 개입	언어 점검, 전문가와의 협업, 스타일 가이드 개발, 작성자 검증 및 교육	선도적인 언어 AI 기업으로서 수년간 글로벌 기업들과 협력하고 있습니다.



Q&A

LIONBRIDGE