AI and Diversity

Marco Buemi

IMPACT OF IA

- Over the next 10 years, AI could produce an annual increase in labour productivity of 1.44%.(Goldman Sachs, 2023)

- Worldwide electricity consumption by data centres will double between 2022 and 2026. (IEA, 2023)

- About 40% of global employment is exposed to AI. In advanced countries, about 60 per cent of jobs could be impacted by AI. (IMF, 2023)

- The transformative impact on the economy and society will be significant but we cannot yet quantify it;

- exposure to AI concerns more High-medium skilled workers in particular in the Services sector (mathematicians and actuaries, professionals in finance, legal and administrative professionals, but also university lecturers or software developers are among the professions most exposed to AI);

- Greater exposure to AI does not necessarily imply a greater risk of substitution. AI rather than substitution can complement/increase the ability of workers to perform their tasks;

- At present, territories with a higher level of AI exposure of employees have higher employment growth rates.

What we don't know about AI

- Whether complementarity effects will prevail over substitution effects;

- Whether the combination of AI with automation (robots) will severely impact employment;

- Whether the productivity gains and growth of economies triggered by AI can compensate or more than compensate substitution effects.

Percentage of companies using AI tools (2021)

- larger companies have more expertise,
  capital to invest, scope and return on applications

- importance of policies for a wider deployment of AI in enterprises
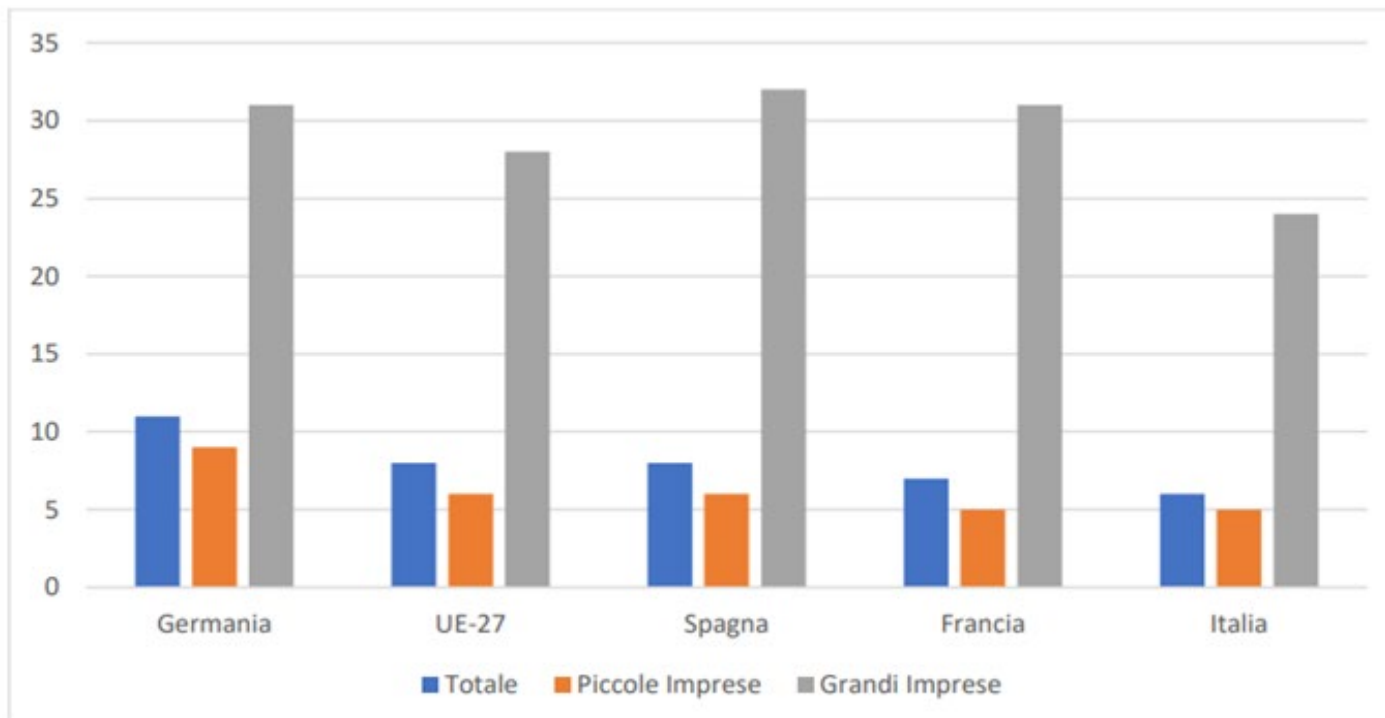
Territorial gaps

- strong polarisation between the centre (Germany and Scandinavia) with the most digital skills and the periphery
- the recent evolution of the Italian position compared to the rest of the EU is not positive, with a widening of the gap with the most advanced countries
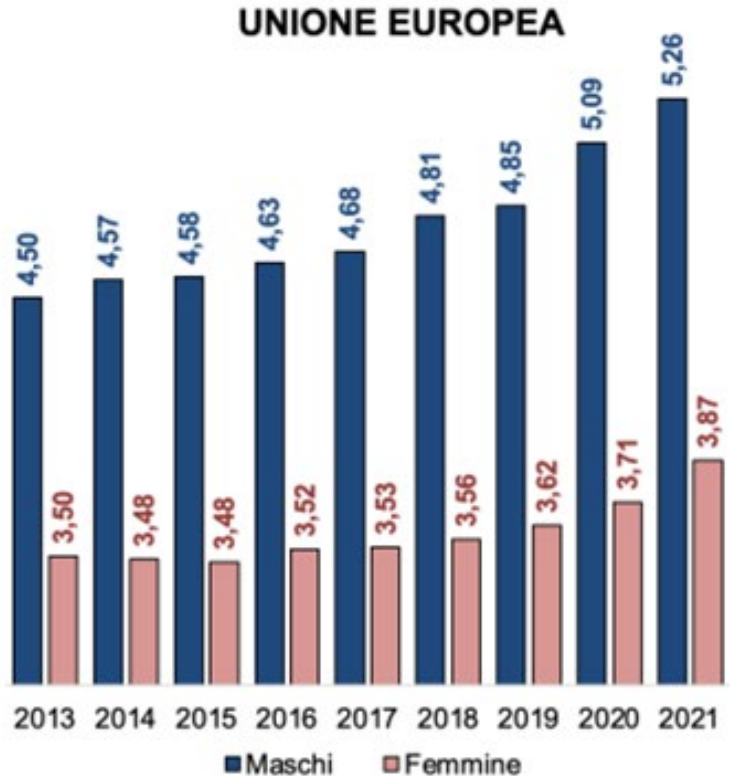
Legenda
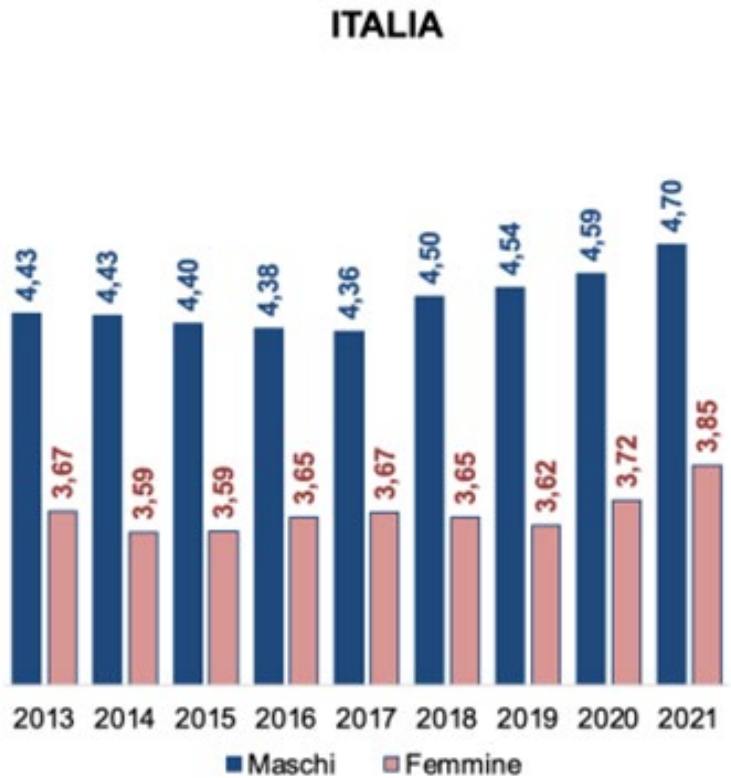- 4 quartile
- 3 quartile
- 2 quartile
- 1 quartile

2013

2021
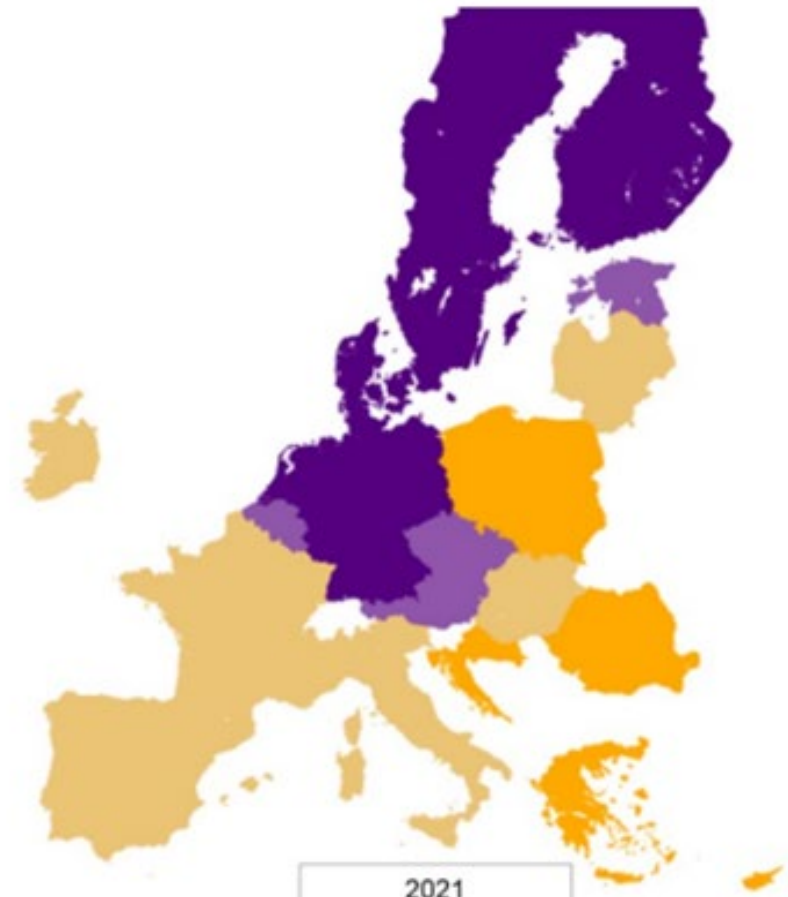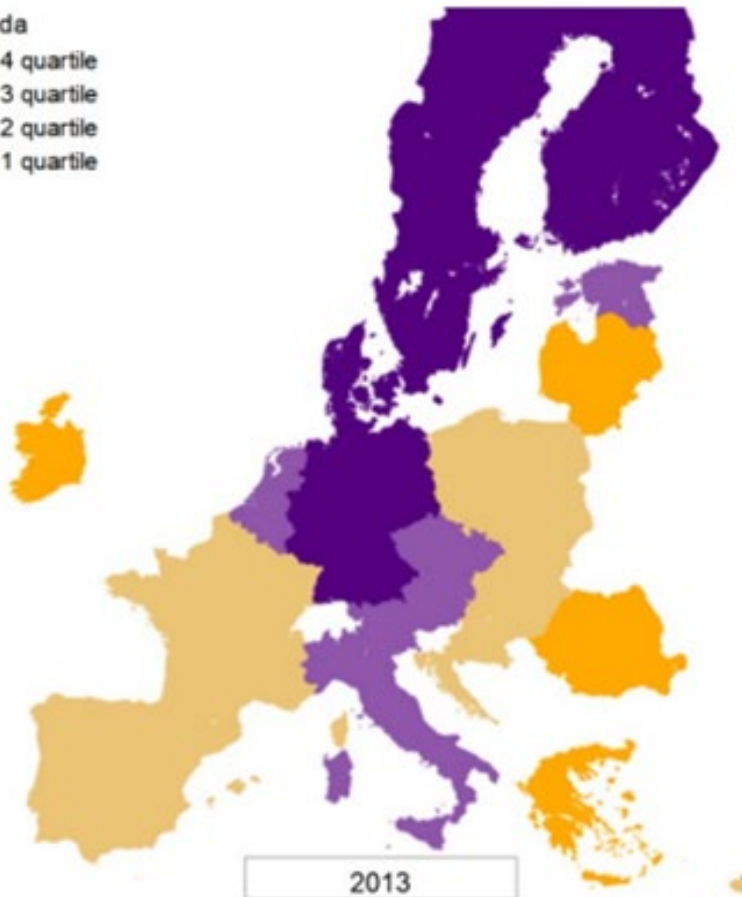
Social gaps-Gender

Differences between territories are compounded by differences between social categories.
women are systematically penalised, showing a level of digital competence that in Italy is between 20 and 30% lower than that recorded for men.
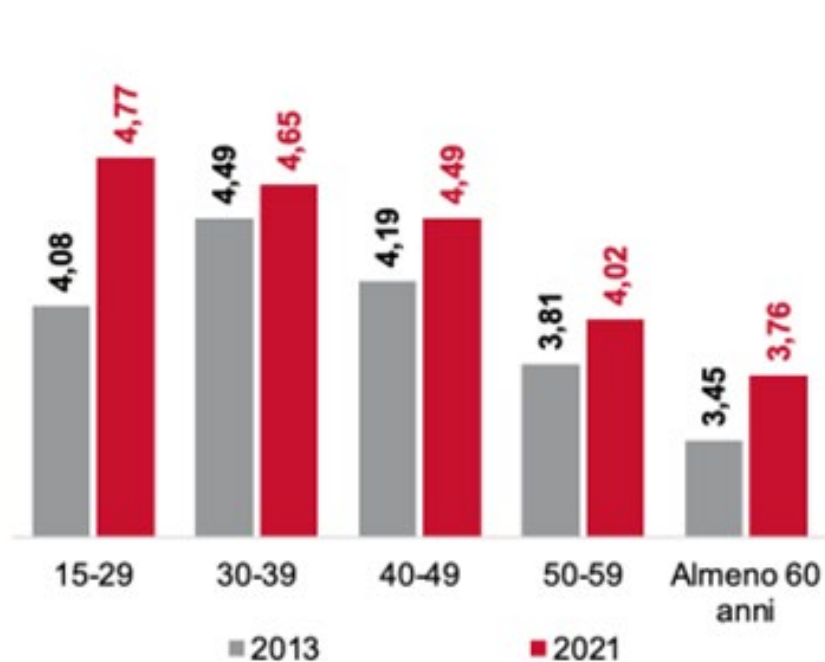
Social Gaps-Age

- digital skills are highly concentrated in younger ages and tend to decline as age increases

- a very important critical element that points to the need to develop specific policies for the reskilling and upskilling of workers, especially older workers

ITALIA

Age	2013	2021
15-29	4,08	4,77
30-39	4,49	4,65
40-49	4,19	4,49
50-59	3,81	4,02
Almeno 60 anni	3,45	3,76

UNIONE EUROPEA

Age	2013	2021
15-29	4,19	4,93
30-39	4,45	5,12
40-49	3,99	4,53
50-59	3,68	4,15
Almeno 60 anni	3,36	3,88
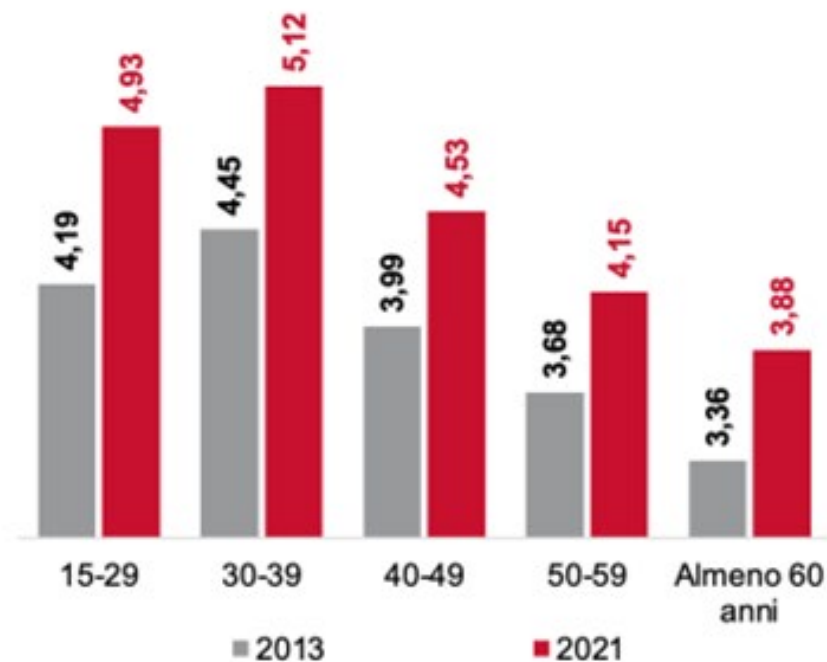
Social gaps-Income

- Digital skills seem to be a peculiarity of middle-income earners
- The digital skills increases with the income level of workers
- this evidence, which is confirmed both in the Italian case and for the EU as a whole, introduces an element of concern regarding the relationship between digitisation and polarisation, in this case linked to inequalities in income distribution
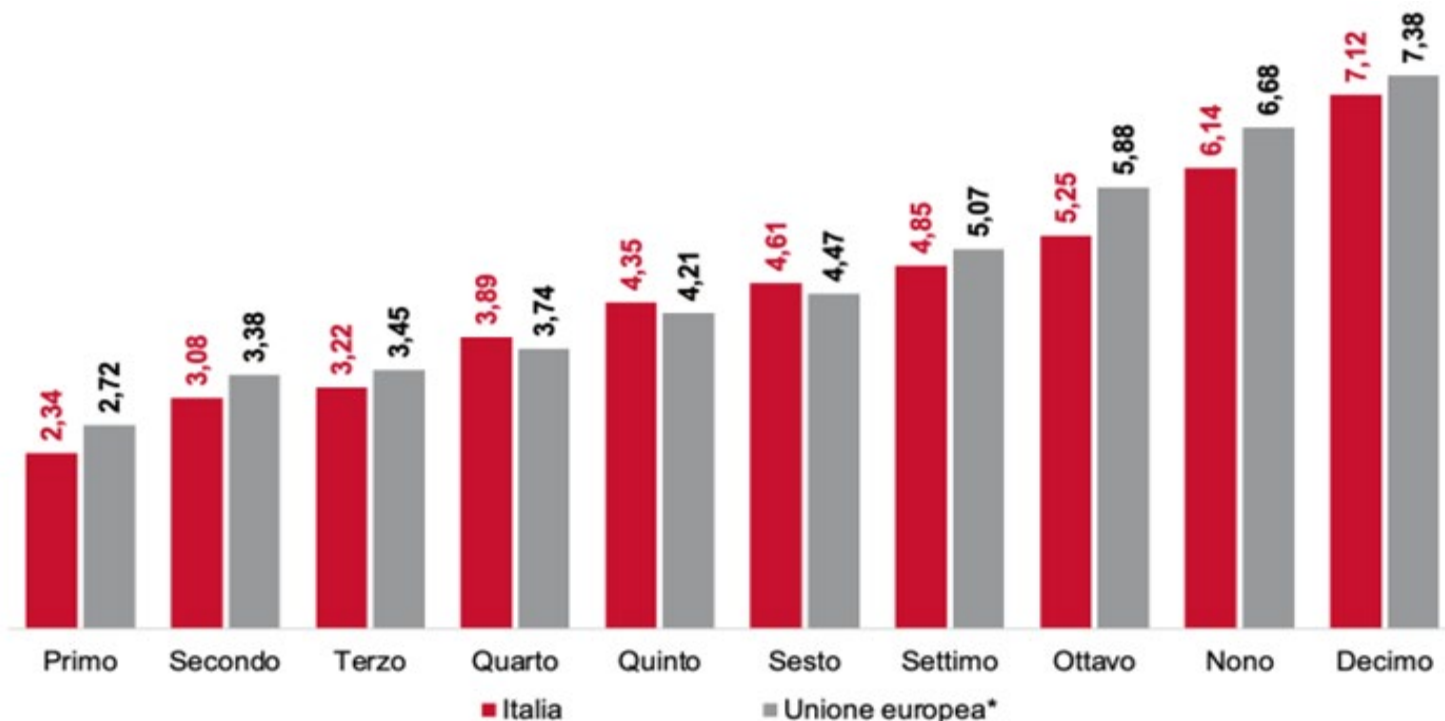
The impact of AI on diversity and inclusion

From chat bots to machine learning algorithms, AI has the potential to transform the way we work, communicate, and interact with the world around us.

However, as with any technological advancement, it is important to consider the potential impact that AI may have on diversity and inclusion.

In this video, we will explore the ways in which AI is impacting diversity and inclusion and what steps can be taken to mitigate any negative effects.

https://www.youtube.com/watch?v=J_c3F-EX3fA

UNCONSCIOUS BIAS

Unconscious Bias is a social stereotype that every individual formulates regardless of his or her own will

1998: **IMPLICIT ASSOCIATION TEST (IAT)**
Yale and Washington Universities define 'a new tool that measures the unconscious roots of prejudice'

The study recognises how prejudice affects 90/95% of the population

Prejudice: determined by the need to categorise and divide into groups in order to better manage the flow of information

Consequences: on the labour market, on the delivery of care in the medical sector, in the application of law

ALGORITHMIC BIAS

In a nutshell, "**[algorithmic] [b]ias happens when seemingly innocuous programming takes on the prejudices either of its creators or the data it is fed**."

As a consequence, women for example (especially from minority groups) may be denied loans and credit, and speech recognition programs may misidentify words spoken by black people at much greater rates than for white people

five sources and types of bias in AI:

"historical bias" describes how social hierarchies and institutionalised disadvantage shape social data. Data is therefore not neutral because it reflects the unequal society in which we live. *For example, as women have historically earned less than men, they may be given less credit or, in the context of advertising, be served adverts for lower paid job posts.*

"representation bias" arises in data collection. *For example, if an organisation's marketing team advertises in predominantly white neighbourhoods, the resulting customer base would not be representative of the wider population. That dataset would generate bias if used for example to train an algorithm later used to cater broader population groups*

"measurement bias", which "occurs when choosing, collecting, or computing features and labels to use in a prediction problem". *For example, postcode could be a proxy for race or sexual orientation, occupation could be a proxy for gender and first names are often used as proxies for age. Alternatively, if proxies overly simplify the feature to be measured or the proxy reflects variations in the quality of measurements across groups, measurement bias could arise.*

"Aggregation bias" relates to how data is combined. It occurs when data groups are inappropriately combined, resulting in a model that does not perform well for any group or only performs well for the majority group. The researchers mention the example of local meanings ascribed by specific communities to emoji, hashtags and sentences on social media, which differ from the meanings in the broader social media user population

"evaluation bias", which occurs when evaluating a model, if the benchmark data (used to compare the model to other models that perform similar tasks) does not represent the population that the model will serve. *For example, the Gender Shades paper discovered that two widely used facial analysis benchmark datasets (IJB-A and Adience) were primarily composed of lighter-skinned subjects (79.6% and 86.2%, respectively)*

The discriminatory impact of AI: some concrete examples

Recruitment: Reuters reported in 2018 that Amazon developed a program relying on machine-learning to identify top candidates in pools of CVs. The program systematically disadvantaged women's CV because it reflected the gender gap in the workforce recruited over the past ten years. Neutralising words like "women" did not redress the discriminatory outcome as the system was able to infer sex from other data.

The online targeted distribution of job adverts powered by optimisation services offered by social media platforms such as Facebook also serves to reinforce gender stereotypes as well as gender segregation within the work place.43 An experiment conducted by AlgorithmWatch in 2020 showed that when asking Facebook to distribute ads "neutrally" (without targeting a specific audience), an ad for a truck driver position was shown to a public composed of 93% men and 7% women. Conversely, an advert for a position as educator was distributed to an audience composed of 96% women and 4% men

Access to goods and services, banking and insurance: In Finland the National Non-Discrimination and Equality Tribunal found direct multiple discrimination in a case where the applicant was denied a loan online. After investigating the case, the Equality Body (the Non-Discrimination Ombudsman) found that the company had used statistical models to assess credit worthiness that relied on an applicant's age, gender, language and place of residence while not taking into account an applicant's actual credit history.

In that case, the applicant being male, Finnish speaker and from a rural area were treated as factors of disadvantage in the assessment performed by the financial institution.

Media and search engines: Research shows that representations of women in images returned by search engines online are biased and reflect sexist, racist and intersectionally discriminatory stereotypes. For instance, Noble shows in an experiment with the Google search engine how images of black girls and black women are sexualised.73 Other groups of minority women are also subjected to sexualised stereotyping in search engines results, for example in searches related to the word "lesbian"

Online gender-based violence, hate speech, harassment: Digital discrimination also takes the form of gender-based violence, for instance when deepfake videos are used to harass women in the context of so-called "revenge porn" cases. Unconsented dissemination of sexual content, often in the form of images, has also been recognised as a form of gender-based violence that especially affect women and girls who are young or public figures such as journalists, human rights defenders, or politicians.77 In addition, sexist and other forms of online hate speech have been highlighted as contingent on the rising use of social media platforms

Challenges with the role of AI in DEI

Facial Recognition:
Facial recognition systems have shown biases in correctly identifying individuals based on their race, gender, or age. Studies have revealed higher error rates for women, people of color, and the elderly. This bias is often a result of imbalanced training data that predominantly includes images of white males, leading to reduced accuracy for underrepresented groups.

Criminal Justice:
AI algorithms used in predicting recidivism or determining parole decisions have demonstrated bias against certain minority groups. Research has found that these systems are more likely to falsely flag individuals from underrepresented communities as high-risk or to assign longer sentences to them, compared to white defendants with similar characteristics.

Hiring and Recruitment:
AI-based systems used in resume screening or candidate selection have displayed bias against candidates based on their gender, race, or ethnicity. Biased algorithms can perpetuate historical disparities by favoring certain demographics or penalizing underrepresented groups, reinforcing discriminatory practices.

Natural Language Processing:
Language models trained on large datasets can inherit biases present in the training data. These biases can manifest in the generation of offensive or discriminatory content, reinforcing stereotypes or marginalizing certain groups. Biases related to race, gender, and other sensitive attributes have been observed in AI-generated

Opportunities for AI to advance DEI efforts

Inclusive Design and User Experience:
AI can be utilized to create inclusive design solutions that accommodate a wide range of users, including disabled people and people with different cognitive abilities. Natural language processing and computer vision technologies can enable more accessible interfaces and assistive technologies, making digital products and services more inclusive for all users.

Decision Support and Transparency:
AI can provide decision support tools that help reduce bias in human decision-making processes. By leveraging machine learning algorithms, AI systems can offer objective insights and recommendations, minimizing subjective biases that can arise from human judgment. Transparent AI systems also allow for better accountability and auditing of decisions made, ensuring fairness and addressing concerns about hidden biases.

Education and Awareness:
AI can be employed to raise awareness about diversity and inclusion issues. Chatbots, virtual assistants, and educational platforms powered by AI can provide information, resources, and training on topics such as unconscious bias, cultural sensitivity, and inclusive practices. These technologies can help foster a more inclusive and informed society.

Personalized Experiences and Recommendations:
AI algorithms can be leveraged to provide personalized experiences that cater to the individual needs and preferences of diverse users. By considering a person's unique characteristics and adapting recommendations or content accordingly, AI can contribute to a more inclusive and engaging user experience.

It is important to recognize that AI is a tool and its impact depends on how it is designed, developed, and deployed.

To ensure AI's positive impact on diversity and inclusion, it is essential to have diverse teams of developers and researchers working on AI projects and incorporating ethical considerations throughout the entire AI development lifecycle.

Boston Consulting Group

- Demonstrate a direct link between a higher level of diversity in leadership and a higher rate of business innovation and higher EBIT (Earnings Before Interests and Taxes) margins

- Research sample: approximately 1700 companies of different sectors and sizes in 8 countries (Austria, Brazil, China, France, Germany, India, Switzerland, USA)

- Research subject: Management perceptions of six dimensions (gender, age, nationality, career path, industry background, education)

- Index of the level of innovation taken as a reference: percentage of total turnover from new products and services launched in the last 3 years

- 75% of respondents stated that DM is gaining value in their organisations

- employees of companies in emerging countries (China, Brazil, India) report more progress in recent years than companies in developed markets

- Found a strong correlation between the level of management diversity and the level of business innovation

- Companies with higher diversity reported 19 percentage points higher turnover than companies with low levels of management diversity

EXHIBIT 1 | Companies with More Diverse Leadership Teams Report Higher Innovation Revenue

Companies with
below-average diversity scores

Companies with
above-average diversity scores

26%
average innovation
revenue reported
by companies
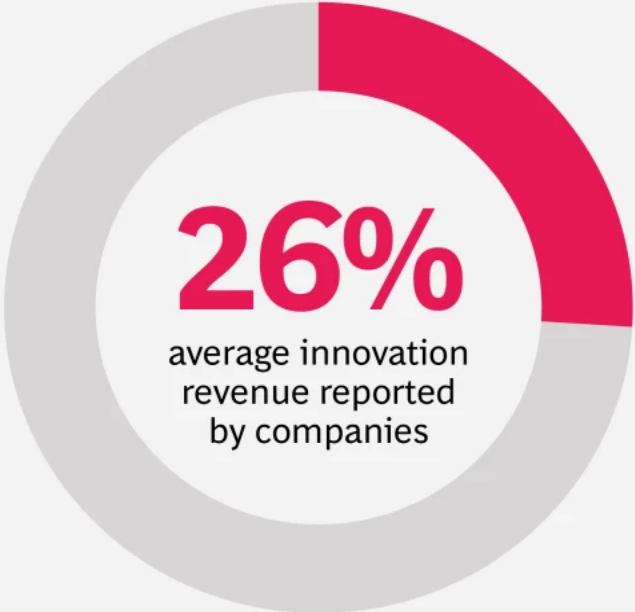
45%
average innovation
revenue reported
by companies
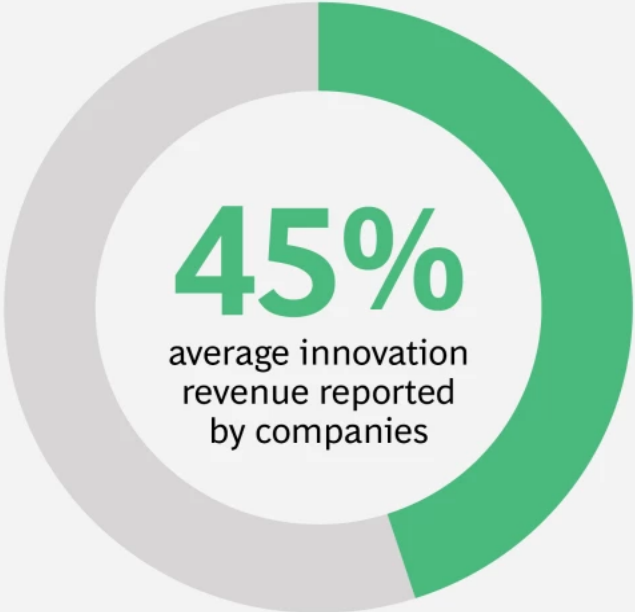
Source: BCG diversity and innovation survey, 2017 (n=1,681).
Note: Average diversity score calculated using the Blau index, a statistical means of combining individual indices into an overall aggregate index.

Companies with Above-Average Diversity Also Have Higher Innovation Revenues

SHARE OF INNOVATION REVENUES FROM PRODUCTS LESS THAN THREE YEARS OLD
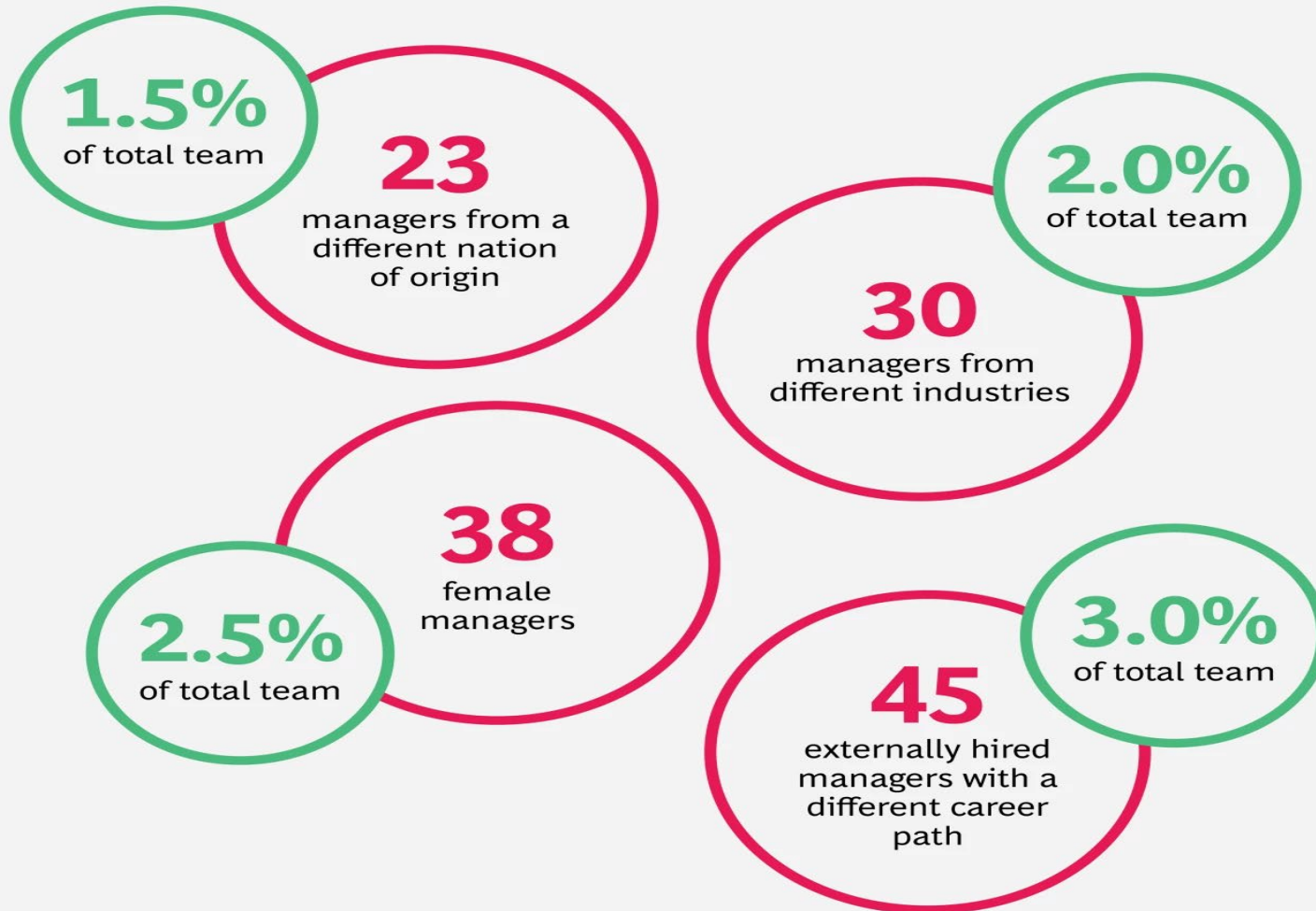

TOTAL DIVERSITY INDEX* (0–1)
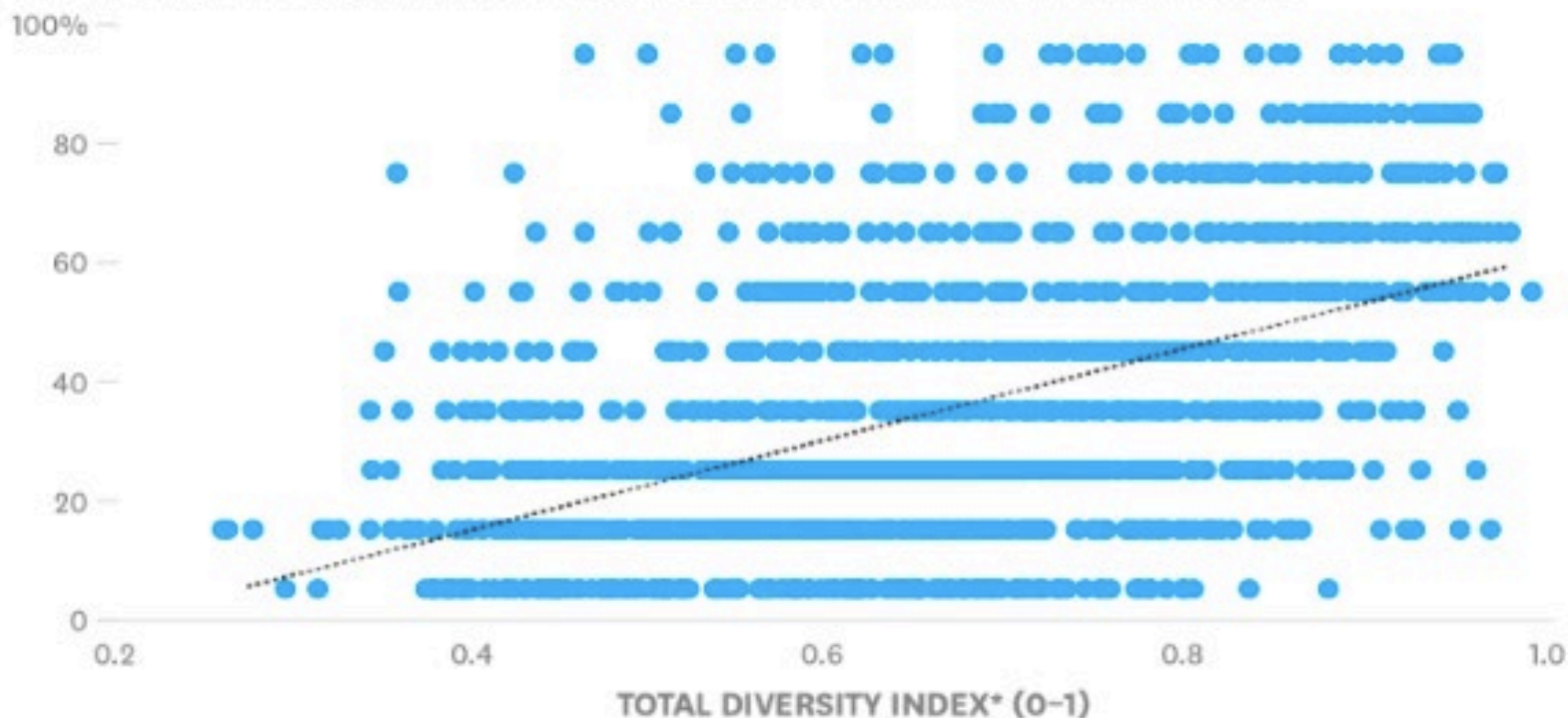
NOTE N=1,606, R^2=0.257 (SIGNIFICANT AT P=0.001 LEVEL); *TOTAL DIVERSITY INDEX IS THE AVERAGE OF THE BLAU INDICES FOR SIX DIMENSIONS OF DIVERSITY: MIGRATION, INDUSTRY, CAREER PATH, GENDER, EDUCATION, AND AGE.
SOURCE BCG ANALYSIS OF MORE THAN 1,600 COMPANIES ACROSS EIGHT COUNTRIES

© HBR.ORG

IMPACT OF GLOBALISATION AND NEW TECHNOLOGIES

- Globalisation and technology boost business performance

- Higher diversity impacts in companies with an emphasis on digital innovation

- Relationship between diversity and performance stronger in companies with interests and activities in different countries

EXHIBIT 2 | Changes in Leadership Can Lead to a Percentage Point Gain in Innovation Revenue

1.5% of total team

23 managers from a different nation of origin

2.0% of total team

30 managers from different industries

38 female managers

2.5% of total team

45 externally hired managers with a different career path

3.0% of total team

IMPACT OF DIVERSITY ON THE LEVEL OF INNOVATION

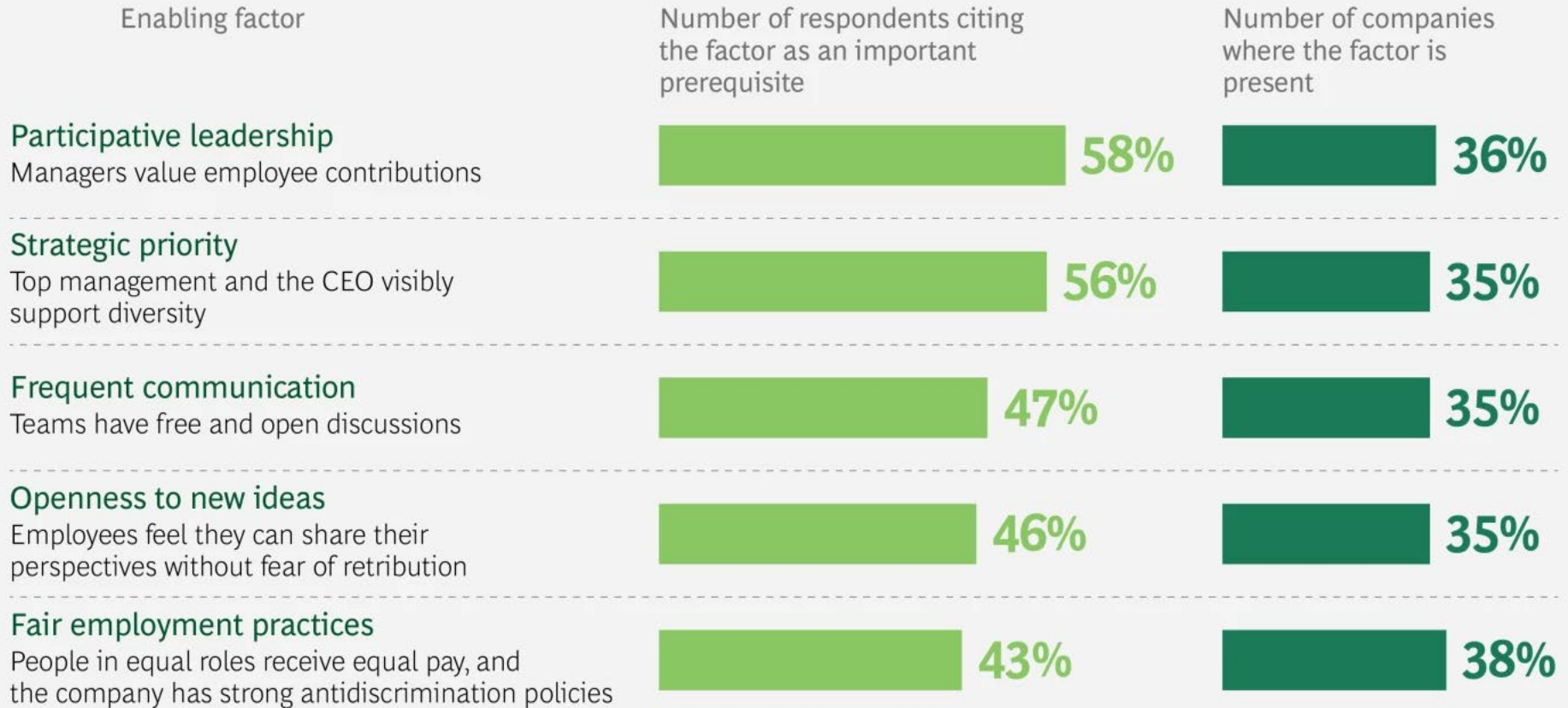- Fair employment practices, fair and equal pay, participative leadership, support from top management, open communication practices (12.9 percentage points higher on innovation)

- Greater diversity of managers could imply a 1% higher innovation rate

- Diversity in geographical origin + 1.5%

- Diversity of employees by sector of origin + 2%

- Diversity of employees by gender + 2.5%.

- Diversity in career progression + 3%

Small changes in the makeup of management can have a significant impact.

For example, if our hypothetical company were to hire 30 managers from a different industry (2% of the total management team), it would improve its innovation revenue by a full percentage point. Hiring 38 female managers (2.5% of the team) would have the same result, as would hiring 23 managers (1.5% of the team) from a country other than the one in which the company is based.

To be clear, these are not incremental new hires but rather replacements for existing managers and executives; the overall size of the management team remains the same, but it is more diverse.
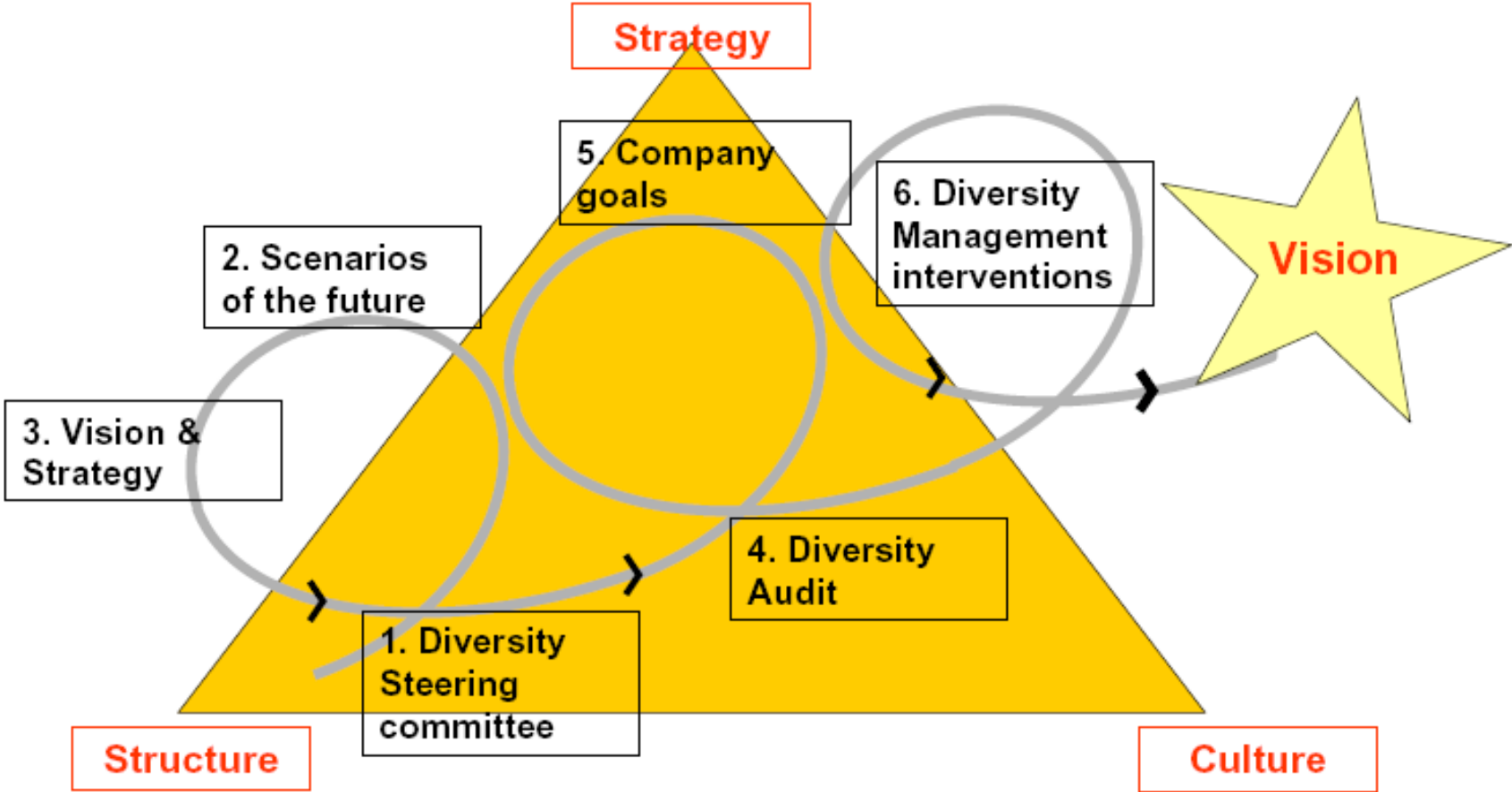
EXHIBIT 3 | Enabling Factors Exist in Fewer than Half the Companies Surveyed

Enabling factor	Number of respondents citing the factor as an important prerequisite	Number of companies where the factor is present
Participative leadership Managers value employee contributions	58%	36%
Strategic priority Top management and the CEO visibly support diversity	56%	35%
Frequent communication Teams have free and open discussions	47%	35%
Openness to new ideas Employees feel they can share their perspectives without fear of retribution	46%	35%
Fair employment practices People in equal roles receive equal pay, and the company has strong antidiscrimination policies	43%	38%

Source: BCG diversity and innovation survey, 2017 (n=1,681).

DIVERSITY MANAGEMENT IMPLEMENTATION

Source: synetz – the management consultants www.synetz.de

PROS AND CONS OF AI IN EDUCATION

PROS	VS	CONS
Improved Student Engagement and Motivation		Threat to Teacher's Job Security
Enhancing Student Performance		Dehumanized Learning Experience
Cost-Effective Learning		Costly to Implement for Teachers
Continuous Evaluation and Improvement in the long run		Dependence on Technology

THANK YOU!

Email: m.buemi72@gmail.com