INDEPENDENT THINKING ABOUT ARTIFICIAL INTELLIGENCE

BY ADRIAN DUNKLEY

AI BOSS

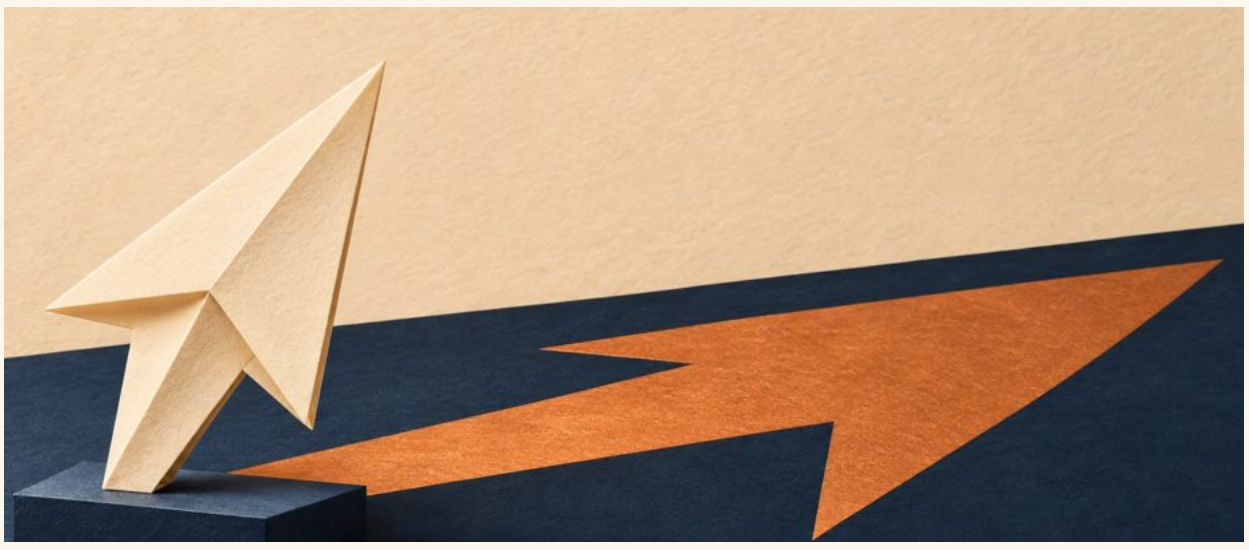
THE DAILY AI FLUENCY NEWSPAPER

SUNDAY 8 NOVEMBER 2026 | ISSUE 007 | AIBOSSHQ.COM

MODEL EVALUATION

Test AI confidence against its mistakes

A 2024 study found that a model could give the same high score to right and wrong answers. Test whether the scores help staff recognise mistakes.



The Long Shadow of a Guess
AI-generated artwork · AI Boss Ltd

Flag mistakes that could charge the wrong project or grant. See whether they get lower scores than correct answers.

If they routinely get the same scores, leave the percentage off the screen and show the relevant policy entry when you can verify it.

Continue reading at aibosshq.com

AI AND CULTURE

Check occupation illustrations as a set



Sparsh Paliwal / Unsplash · Illustrative photograph
Unsplash Licence

A 2023 study found that a model repeatedly showed certain jobs in gender-stereotyped ways. Review the full batch before choosing pictures for a careers guide.

Compare roles side by side and look at who appears to do the work, use equipment and lead.

Continue reading at aibosshq.com



Cht Gsml / Unsplash · Illustrative photograph
Unsplash Licence

AI AT WORK

Count AI-assisted work as complete after verification

Put verified and returned drafts on the same team scorecard, alongside the hours spent checking or correcting them.

Continue reading at aibosshq.com



A Fresh Page for the Test
AI-generated artwork · AI Boss Ltd

BENCHMARKS

Keep a fresh set of AI test cases private

Save the dated inputs and answer key before the test. Replace any case used to tune the workflow.

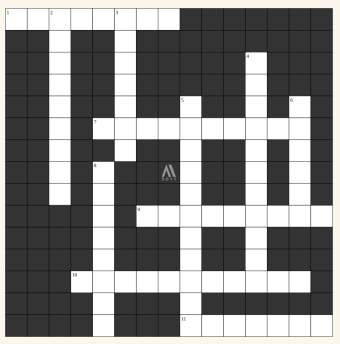
Continue reading at aibosshq.com

THE CROSSWORD DESK

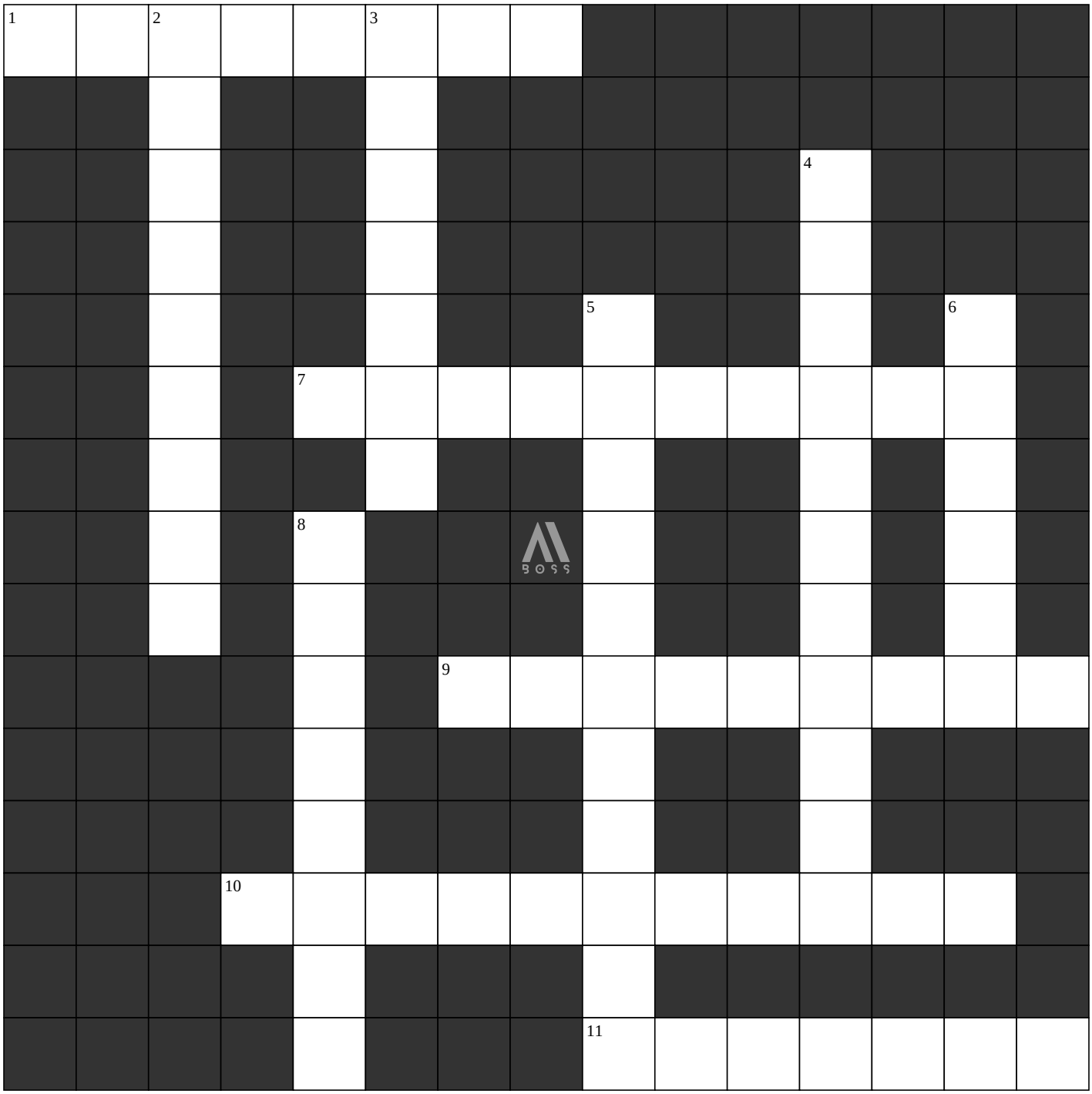
How well do you know your AI?

Practise AI vocabulary and judgement through workplace clues. Play the crossword or fill the grid with Quiz Mode at **aibosshq.com/games/daily/2026-11-08/**.

The printable puzzle follows on page 2. Its solution is on page 3.



Article sources and illustration credits appear with the full essays.



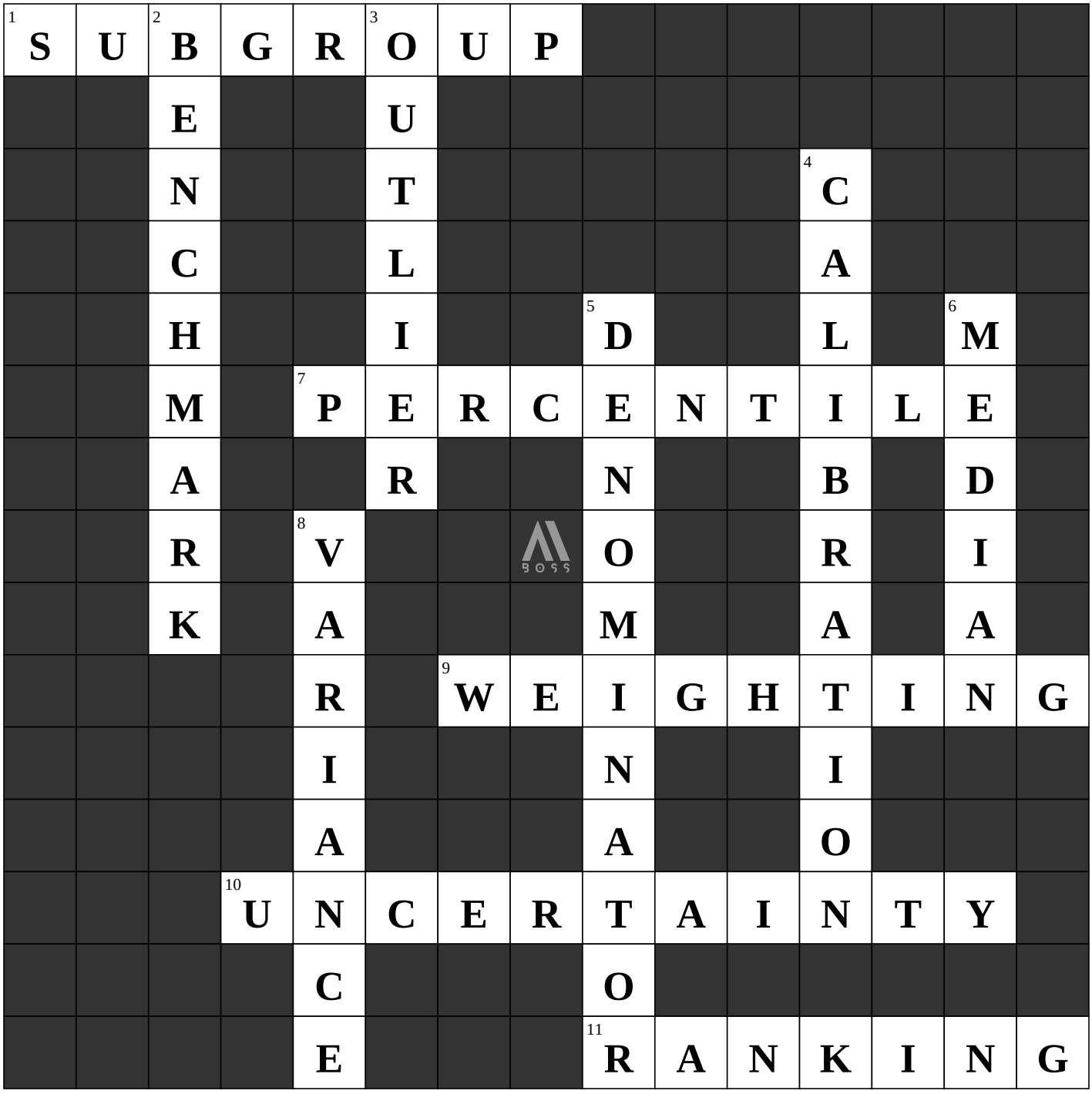
ACROSS

1. Part of a larger group that shares a characteristic. (8)
7. Cutoff that marks a percentage of ordered values. (10)
9. Giving some data points more influence. (9)
10. Doubt about the true value after measurement. (11)
11. Order of items according to their scores. (7)

DOWN

2. Standard set of tasks for comparing systems. (9)
3. Observation unusually far from the other values. (7)
4. Fit of predicted chances to observed rates. (11)
5. Number below the fraction line. (11)
6. Middle value when data are ordered. (6)
8. Spread based on squared gaps from the mean. (8)

Solution


1 S U 2 B G R 3 O U P
E U
N T 4 C
C L A
H I 5 D L 6 M
M 7 P E R C E N T I L E
A R N B D
R 8 V O R I
K A M A A
9 W E I G H T I N G
I N I
A A O
10 U N C E R T A I N T Y
C O
E 11 R A N K I N G