AI-02 · Bài 4

Cách kiểm một câu trả lời trước khi tin nó

Lớp Trợ lý AI giúp được việc gì, và không giúp được việc gì

m.school by MES LAB

Học xong bài này bạn sẽ

- chia được một câu trả lời thành bốn loại khẳng định, và nói được loại nào kiểm bằng cách nào.
- Chọn được mức kiểm tương ứng với hậu quả nếu nội dung ấy sai, thay vì kiểm đều tay hoặc không kiểm gì.
- Nói được vì sao một phép kiểm chỉ đáng gọi là phép kiểm khi nó dựa vào một chỗ đứng độc lập với thứ đang bị kiểm.

1 · Một câu trả lời gồm bốn loại khẳng định, và mỗi loại kiểm bằng một cách khác nhau

- Việc đầu tiên không phải là kiểm, mà là chia.
- Loại thứ nhất là khẳng định có địa chỉ
- Loại thứ hai là phép tính và con số rút ra từ chất liệu của bạn
- Loại thứ ba là mô tả cơ chế và kiến thức chung của nghề
- Loại thứ tư là suy luận, đánh giá và khuyến nghị
- Chia xong bốn loại là xong phần lớn việc.

2 · Phép kiểm phải đi ra ngoài văn bản, và phải tới tận bản gốc

- Một phép kiểm chỉ đáng gọi là phép kiểm khi nó dựa vào một chỗ đứng không phụ thuộc vào thứ đang bị kiểm.
- Chỗ hỏng thứ nhất là ở lại bên trong cuộc trò chuyện.
- Chỗ hỏng thứ hai tinh vi hơn, và nó đã làm hỏng việc của nhiều người cẩn thận: dừng lại ở một bản tóm tắt.
- Là nơi chịu trách nhiệm phát hành thứ đó.
- Có một hệ quả cần nói thẳng, vì nó ảnh hưởng tới cách bạn giao việc.

3 · Tính lại con số bằng tay, vì trình bày gọn không phải là bằng chứng

- Với loại khẳng định thứ hai, có một việc không thay thế được: tự làm lại phép tính.
- Bài 3 đã nói vì sao lời giải của máy có thể trình bày đúng mà kết quả sai.
- Có hai chỗ nên soi riêng vì chúng hay sai mà khó thấy.
- Còn một thứ nữa phải đòi, và nó quan trọng ngang con số: điều kiện đo.

4 · Mức kiểm đi theo hậu quả nếu sai, và ai ký tên thì người đó chịu

- Không thể kiểm mọi thứ đều tay.
- Thứ quyết định mức kiểm không phải là độ dài câu trả lời, cũng không phải cảm giác của bạn về nó.
- Một câu hỏi gọn để xếp hạng: nếu chỗ này sai và nó đi tới tay người ngoài thì chuyện gì xảy ra, và ai phải đứng ra trả lời.
- Phần cuối là phần không đổi dù máy khá lên tới đâu.
- Có một chi tiết đáng để ý trong vụ kể ở phần ví dụ dưới đây, vì nó lật ngược một lập luận nghe rất quen.
- Cùng một chất lượng đầu ra, ba việc có hậu quả khác nhau thì cần ba mức kiểm khác nhau; mức kiểm đi theo hậu quả, không đi theo độ dài câu trả lời. -->

Bảng tổng hợp

Bảng tổng hợp bốn loại khẳng định và cách kiểm từng loại

Loại khẳng định	Ví dụ trong một câu trả lời	Kiểm bằng cách nào	Chỗ phép kiểm hay hỏng
Có địa chỉ tra được	số hiệu tiêu chuẩn, điều khoản, bản án, trích dẫn, con số đã công bố	mở nơi giữ bản gốc và đọc chính chỗ đó	dừng lại ở một trang viết lại, tưởng đó là bản gốc
Phép tính trên chất liệu của bạn	tổng, tỉ lệ, so sánh hai phương án, bản tóm tắt một bảng số	chép số gốc sang bảng tính và để máy tính lại, hoặc ước lượng thô	đọc lời giải thấy mạch lạc rồi nhận luôn kết quả

Bảng tổng hợp

Bảng tổng hợp bốn loại khẳng định và cách kiểm từng loại

tiếp

Loại khẳng định	Ví dụ trong một câu trả lời	Kiểm bằng cách nào	Chỗ phép kiểm hay hỏng
Mô tả cơ chế và kiến thức chung	một phương pháp gồm những bước nào, vì sao một hiện tượng xảy ra	đối chiếu với tài liệu chuyên ngành hoặc một người biết việc	nhận phần đúng ở mức chung rồi áp thẳng vào hoàn cảnh riêng
Suy luận và khuyến nghị	nên chọn phương án nào, rủi ro lớn nhất nằm ở đâu	kiểm phần nền ba loại trên mà lời khuyên dựa vào	kiểm lời khuyên bằng cảm giác nghe có hợp lý không

Phần giá trị nhất của bài

Chỗ người ta hay làm sai

- **Dùng chính cái máy làm chỗ đứng để kiểm cái máy**

  Cách tránh là ghi nhớ một ranh giới cứng: mọi phép kiểm đều phải kết thúc ở một chỗ nằm ngoài cuộc trò chuyện, dù chỗ đó là một trang của cơ quan ban hành...

- **Đối chiếu với một bản tóm tắt rồi gọi đó là đã mở nguồn**

  Cách tránh là tự hỏi một câu trước khi đóng việc: trang tôi vừa đọc có phải là nơi chịu trách nhiệm phát hành thứ này không.

- **Kiểm rất kỹ thứ dễ kiểm, rồi bỏ qua thứ đắt nhưng quan trọng**

  Cách tránh là đảo thứ tự: trước khi kiểm bất cứ thứ gì, đọc một lượt và đánh dấu những chỗ mà nếu sai thì hậu quả rơi ra ngoài công ty, rồi bắt đầu từ đúng những chỗ ấy.

Ví dụ để khái niệm sáng ra

Ví dụ minh hoạ

- Ngày 14 tháng 2 năm 2024, toà giải quyết tranh chấp dân sự tỉnh bang British Columbia, Canada, ra quyết định trong vụ Moffatt kiện Air Canada, số hiệu 2024 BCCRT 149.
- Hãng lập luận rằng hãng không chịu trách nhiệm về thông tin do trợ lý trò chuyện đưa ra.

Điều rút ra: với người ở bên ngoài, câu do máy đưa ra và câu do nhân viên đưa ra đều là lời của một nơi có tên, và việc đối chiếu thuộc về nơi phát ra thông tin.

Ví dụ dưới đây là một quyết định công khai của cơ quan phân xử, tra lại được bằng số hiệu.

Bài tập tự chấm

Thử nhìn lại việc của mình

- [ ] Chọn một tài liệu bạn sắp gửi đi trong tuần này, loại có người ngoài nhóm đọc, và trong đó có ít nhất một phần do trợ lý AI dựng nháp.
- [ ] Đi dọc tài liệu và đánh dấu mỗi câu thuộc một trong bốn loại khẳng định ở bảng tổng hợp.
- [ ] Sau đó làm một cột thứ hai: với mỗi chỗ đã đánh dấu, viết một dòng ngắn trả lời câu nếu chỗ này sai thì ai chịu và chuyện gì xảy ra.
- [ ] Xếp các chỗ theo cột thứ hai, nặng lên trên.
- [ ] Bắt đầu kiểm từ chỗ nặng nhất, và với mỗi chỗ thuộc loại có địa chỉ, ghi lại bạn đã mở đúng trang nào của nơi nào.

Cách tự chấm và lời giải mẫu nằm trong bài trên trang web.

m.school

Đọc trọn bài và làm bài tập tại

mschool.meslab.org/lop/AI-02/bai-04.html

m.school by MES LAB · học không mất phí, không cần đăng nhập