

Annexure – Allens AI Australian law benchmark

The Allens AI Australian Law Benchmark explores the intriguing question: can we trust AI tools to provide accurate and reliable advice on matters of Australian law? To cut through the hype and provide concrete insights, Allens with Linklaters LLP has established an ongoing benchmarking process. This allows us to systematically evaluate and track the progress of generative AI in answering legal queries under Australian Law over time. You can view the context and a summary of our results here: [Lawyer or language model?](#)

This annexure accompanies the benchmark. It sets out the scores achieved by each model against each question during our February 2024 test. Each score is the average across three attempts.

Question 1 (Contract law)

In Australia, to what extent is a party's subjective intention relevant when interpreting a contract?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 2.67 | 2.00 | 2.00 | 4.00 | 2.00 |
| Citations | 1.00 | 1.00 | 1.00 | 2.00 | 1.00 |
| Clarity | 1.00 | 1.00 | 1.00 | 2.00 | 1.33 |
| Total | 4.67 | 4.00 | 4.00 | 8.00 | 4.33 |

Question 2 (Contract law)

The customer and supplier are party to a five-year office cleaning contract covering one office in Parramatta. The supplier invoices the customer on a weekly basis. The customer says it will not pay 50% of the supplier's invoice for one week's cleaning because the cleaning was carried out to a poor standard. The customer has no right under the contract to do this. To what extent is the customer in repudiatory breach of contract, and what remedies are available to the customer?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 1.67 | 1.67 | 2.00 | 1.33 | 2.00 |
| Citations | 0.67 | 0.00 | 0.00 | 0.00 | 1.00 |
| Clarity | 1.33 | 1.00 | 1.33 | 1.00 | 2.00 |
| Total | 3.67 | 2.67 | 3.33 | 2.33 | 5.00 |

Question 3 (Contract law)

Describe the key differences between 'reasonable endeavours' and 'best endeavours', and to what extent do those endeavours include a party's own commercial interests?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.67 | 3.00 | 2.33 | 3.33 | 2.33 |
| Citations | 1.00 | 2.00 | 0.00 | 0.00 | 1.00 |
| Clarity | 2.00 | 1.67 | 1.00 | 2.00 | 2.00 |
| Total | 6.67 | 6.67 | 3.33 | 5.33 | 5.33 |

Question 4 (Intellectual property)

Can a pile of bricks be protected by copyright?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.33 | 2.67 | 1.67 | 1.00 | 1.00 |
| Citations | 1.00 | 0.00 | 1.67 | 1.33 | 1.33 |
| Clarity | 1.00 | 1.00 | 2.00 | 2.00 | 1.00 |
| Total | 5.33 | 3.67 | 5.33 | 4.33 | 3.33 |

Question 5 (Intellectual property)

BIGGENE, a biotechnology company, obtained an Australian patent for a genetically engineered organism designed to produce a specific pharmaceutical compound. The patent claim specifies the genetic modifications made to the organism. However, a competitor called MICROGEN has developed an organism, using newer gene-editing technology and has slightly different genetic modifications, and produces a similar pharmaceutical compound. MICROGEN argues that it does not infringe BIGGENE's patented claim. BIGGENE disagrees. If the parties go to court, what principles would the court apply when interpreting the language of the patent claim in light of advancements in gene-editing technology, and what criteria should be used to determine whether the competitor's organism falls within the scope of the patented invention?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 1.67 | 1.00 | 2.33 | 2.00 | 2.00 |
| Citations | 1.00 | 0.00 | 0.33 | 0.33 | 2.00 |
| Clarity | 0.67 | 0.00 | 1.00 | 0.33 | 1.00 |
| Total | 3.33 | 1.00 | 3.67 | 2.67 | 5.00 |

Question 6 (Intellectual property)

I maintain a price index, calculated as a weighted average of the price of 20 consumer products chosen by me. I re-calculate the index value every day and publish it on my website. My website is freely available to the public. It has come to my attention that one of my competitors is copying my index and publishing it on its own website. Is my competitor infringing my intellectual property rights?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 1.33 | 1.67 | 1.67 | 1.00 | 1.00 |
| Citations | 1.00 | 1.33 | 1.67 | 0.67 | 0.67 |
| Clarity | 0.33 | 0.67 | 1.33 | 0.33 | 1.00 |
| Total | 2.67 | 3.67 | 4.67 | 2.00 | 2.67 |

Question 7 (Tax)

An Australian incorporated company undertakes an off-market share buyback. Explain how (i) an Australian company shareholder; (ii) an Australian individual shareholder; and (iii) a foreign resident shareholder would be taxed. Would the position change if (i) the company undertaking the buyback is an Australian 'listed public company' for tax purposes; or (ii) the buyback is an on-market buyback instead?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 2.67 | 1.33 | 1.00 | 1.00 | 1.00 |
| Citations | 1.00 | 1.00 | 1.33 | 0.00 | 1.00 |
| Clarity | 1.67 | 1.00 | 1.00 | 1.00 | 1.00 |
| Total | 5.33 | 3.33 | 3.33 | 2.00 | 3.00 |

Question 8 (Tax)

How does Australian law apply the principle of universal succession in a merger between two foreign groups, one of which has a wholly owned Australian subsidiary that holds underlying land? What are the resulting stamp duty obligations, if any, in each of the Australian states and territories?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 2.00 | 1.67 | 0.00 | 0.33 | 1.00 |
| Citations | 1.33 | 1.33 | 1.33 | 0.00 | 2.00 |
| Clarity | 1.00 | 0.67 | 1.00 | 1.00 | 1.00 |
| Total | 4.33 | 3.67 | 2.33 | 1.33 | 4.00 |

Question 9 (Tax)

An agreement contains the following gross up clause for withholding tax: 'If a Tax Deduction is required by law to be made by an Obligor, the Obligor shall pay an additional amount together with the payment so that, after making any Tax Deduction, the Finance Party receives an amount equal to the payment which would have been due if no Tax Deduction had been required. The Obligor shall make that Tax Deduction and any payment required in connection with that Tax Deduction to the relevant Tax Authority within the time allowed and in the minimum amount required by law'. An interest payment made by the Obligor of \$100 is subject to Australian withholding tax. How much does the Obligor have to pay to the foreign recipient of the payment and how much does it have to pay to the Australian Taxation Office?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 1.00 | 0.67 | 1.33 | 1.00 | 2.67 |
| Citations | 1.00 | 0.33 | 0.33 | 0.00 | 1.00 |
| Clarity | 1.00 | 0.00 | 1.00 | 0.67 | 1.00 |
| Total | 3.00 | 1.00 | 2.67 | 1.67 | 4.67 |

Question 10 (Data privacy)

Describe the situations in which a criminal offence might be committed under the My Health Records Act 2012 (Cth).

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 1.00 | 1.67 | 0.00 | 0.00 | 0.00 |
| Citations | 1.00 | 1.00 | 1.00 | 0.00 | 1.00 |
| Clarity | 1.00 | 1.00 | 1.00 | 0.00 | 1.00 |
| Total | 3.00 | 3.67 | 2.00 | 0.00 | 2.00 |

Question 11 (Data privacy)

I own a shop and have CCTV that captures the footpath outside my shop. The police ask me to disclose CCTV footage on a voluntary basis in connection with a robbery. Would disclosing the CCTV footage be a breach of the Privacy Act 1988 (Cth)?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 1.00 | 1.33 | 1.67 | 1.67 | 1.67 |
| Citations | 1.00 | 1.00 | 1.00 | 0.33 | 1.00 |
| Clarity | 1.33 | 1.33 | 2.00 | 1.33 | 1.67 |
| Total | 3.33 | 3.67 | 4.67 | 3.33 | 4.33 |

Question 12 (Data privacy)

My website uses Google analytics. Write a paragraph to include in my website privacy policy to disclose the fact that Google analytics are used.

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.00 | 1.67 | 1.67 | 2.00 | 0.00 |
| Citations | 0.67 | 1.00 | 0.00 | 0.67 | 1.00 |
| Clarity | 2.00 | 1.00 | 2.00 | 1.00 | 1.00 |
| Total | 5.67 | 3.67 | 3.67 | 3.67 | 2.00 |

Question 13 (Employment)

Is obesity a health condition that is capable of satisfying the definition of a disability under section 4 of the Disability Discrimination Act 1992 (Cth)?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 2.00 | 1.00 | 2.67 | 1.00 | 1.00 |
| Citations | 2.00 | 0.00 | 0.67 | 0.00 | 2.00 |
| Clarity | 1.00 | 1.00 | 2.00 | 1.33 | 2.00 |
| Total | 5.00 | 2.00 | 5.33 | 2.33 | 5.00 |

Question 14 (Employment)

*Explain whether and, if so, why the non-compete restraint set out below is unenforceable in Victoria?
You shall not, without the prior written consent of the Company, directly or indirectly, either alone or jointly with or on behalf of any third party and whether as principal, manager, employee, contractor, consultant, agent or otherwise howsoever at any time within the period of six months from the date of termination of your employment directly or indirectly engage or be concerned or interested in any business carried on in competition with any of the businesses of the Company or any Group Company which were carried on at the date of termination of your employment or during the period of 12 months prior to that date and with which you were materially concerned during such period.*

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.00 | 2.00 | 2.00 | 3.00 | 2.00 |
| Citations | 2.00 | 0.33 | 0.33 | 0.00 | 1.00 |
| Clarity | 1.33 | 2.00 | 2.00 | 1.00 | 2.00 |
| Total | 6.33 | 4.33 | 4.33 | 4.00 | 5.00 |

Question 15 (Employment)

An employer client is considering dismissing a senior employee with over 15 years' service. The employee's contract does not contain an express notice clause. Can the client dismiss the employee with immediate effect or is it required to give notice? If notice is required, what factors should the client consider when determining how much notice to give?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 2.00 | 3.67 | 3.33 | 2.67 | 2.00 |
| Citations | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 |
| Clarity | 2.00 | 1.67 | 1.33 | 1.33 | 2.00 |
| Total | 5.00 | 5.33 | 4.67 | 4.00 | 5.00 |

Question 16 (Corporate)

A private company has 100 ordinary shares and 50 preference shares on issue. The company has five shareholders. Shareholder A holds 5 ordinary shares, shareholders B and C hold 20 ordinary shares, shareholder D holds 25 ordinary shares and shareholder E holds 30 ordinary shares and 50 preference shares. Which shareholders can call a general meeting of the company?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.00 | 2.00 | 0.67 | 1.67 | 1.00 |
| Citations | 1.67 | 1.00 | 1.00 | 0.00 | 1.00 |
| Clarity | 2.00 | 1.67 | 1.00 | 1.67 | 0.00 |
| Total | 6.67 | 4.67 | 2.67 | 3.33 | 2.00 |

Question 17 (Corporate)

The constitution of a small private company states that all shareholders have pre-emptive rights on the sale of existing shares. The shareholders' agreement of the company contains drag-along rights, allowing shareholders who hold a majority of the shares in the company to facilitate the sale of all of the company to a buyer. Shareholders holding 60% of the shares are willing to sell their shares to a buyer but the buyer wants to buy all of the shares in the company. How can they facilitate this?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 2.00 | 1.67 | 2.67 | 3.00 | 2.00 |
| Citations | 2.00 | 1.00 | 1.00 | 0.00 | 1.00 |
| Clarity | 2.00 | 1.00 | 1.67 | 2.00 | 2.00 |
| Total | 6.00 | 3.67 | 5.33 | 5.00 | 5.00 |

Question 18 (Corporate)

A company has suffered a cyber breach incident. Does this create any potential liability for the directors? Is there a fiduciary duty that relates to cyber risks?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.00 | 3.00 | 4.00 | 4.00 | 3.00 |
| Citations | 2.00 | 1.33 | 2.00 | 0.00 | 2.00 |
| Clarity | 2.00 | 1.33 | 2.00 | 1.00 | 2.00 |
| Total | 7.00 | 5.67 | 8.00 | 5.00 | 7.00 |

Question 19 (Real Estate)

What happens to a registered mortgage on title if the relevant title is subdivided with the consent of the mortgagee in NSW?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 1.00 | 0.00 | 1.33 | 4.00 | 1.33 |
| Citations | 1.00 | 0.67 | 0.67 | 1.33 | 0.00 |
| Clarity | 2.00 | 0.00 | 1.33 | 2.00 | 0.33 |
| Total | 4.00 | 0.67 | 3.33 | 7.33 | 1.67 |

Question 20 (Real Estate)

A freehold title owner grants a lease of a property to Tenant A. Subsequently, the freehold title owner grants a lease of the same property to Tenant B, subject to Tenant A's lease. Tenant B therefore has a concurrent lease. If Tenant B's lease is forfeited by the landlord, what happens to Tenant A's lease?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 2.33 | 2.67 | 2.00 | 2.67 | 2.00 |
| Citations | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 |
| Clarity | 2.00 | 1.67 | 1.33 | 2.00 | 1.67 |
| Total | 5.33 | 4.33 | 3.33 | 4.67 | 4.67 |

Question 21 (Real Estate)

How do the provisions relating to an 'existing use' under the Environmental Planning and Assessment Act 1979 (NSW) operate?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.00 | 1.00 | 1.67 | 2.67 | 2.00 |
| Citations | 1.33 | 0.33 | 1.00 | 0.67 | 1.00 |
| Clarity | 1.67 | 0.00 | 1.33 | 1.00 | 1.00 |
| Total | 6.00 | 1.33 | 4.00 | 4.33 | 4.00 |

Question 22 (Dispute resolution)

To what extent can a former shareholder bring a claim against a third party who has breached a contract with the company, and, if so, what are the pre-requisites to such a claim?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 1.00 | 1.33 | 0.00 | 0.00 | 1.00 |
| Citations | 1.00 | 0.00 | 0.33 | 0.00 | 1.00 |
| Clarity | 2.00 | 1.67 | 0.00 | 0.00 | 1.33 |
| Total | 4.00 | 3.00 | 0.33 | 0.00 | 3.33 |

Question 23 (Dispute resolution)

To what extent is a contractual agreement to mediate a dispute enforceable?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 0.67 | 0.33 | 0.67 | 0.00 | 1.00 |
| Citations | 1.00 | 0.00 | 0.33 | 0.00 | 1.00 |
| Clarity | 1.33 | 0.00 | 0.00 | 0.00 | 0.00 |
| Total | 3.00 | 0.33 | 1.00 | 0.00 | 2.00 |

Question 24 (Dispute resolution)

To what extent will communications between a client and a foreign (non-Australian) lawyer attract legal professional privilege?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 2.00 | 1.00 | 1.33 | 0.00 | 1.67 |
| Citations | 1.00 | 0.00 | 1.00 | 0.00 | 1.33 |
| Clarity | 2.00 | 0.67 | 1.00 | 0.00 | 2.00 |
| Total | 5.00 | 1.67 | 3.33 | 0.00 | 5.00 |

Question 25 (Competition)

What is the legal threshold for finding a merger results in a 'substantial lessening of competition' under Australian merger control rules, and what theories of harm may lead to such a finding?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.33 | 3.67 | 2.67 | 4.67 | 1.00 |
| Citations | 2.00 | 3.00 | 1.33 | 2.00 | 1.67 |
| Clarity | 1.33 | 1.67 | 1.67 | 2.00 | 2.00 |
| Total | 6.67 | 8.33 | 5.67 | 8.67 | 4.67 |

Question 26 (Competition)

A manufacturer and reseller enter into a supply contract for football merchandise, whereby the manufacturer supplies football merchandise to the reseller, which the reseller sells to customers in its chain of stores across Australia. The manufacturer also sells football merchandise direct to consumers via its website. The supply contract requires the reseller to provide the manufacturer with certain information every quarter, including information on the reseller's future marketing plans, future pricing plans and historic information about key customer contracts. Describe the Australian competition law risks (if any) with sharing this information.

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.33 | 2.67 | 3.33 | 3.67 | 2.00 |
| Citations | 1.00 | 1.00 | 2.00 | 2.33 | 1.33 |
| Clarity | 1.33 | 1.67 | 1.33 | 1.67 | 1.00 |
| Total | 5.67 | 5.33 | 6.67 | 7.67 | 4.33 |

Question 27 (Competition)

A franchisor sends an email to franchisees with a list of recommended resell prices for products the franchisee sells under licence. The email also asks that franchisees provide the franchisor with details of the actual prices its licensed products are sold for on a monthly basis, and warns that those franchisees that 'fail to comply with the recommended resell prices' will risk their franchise agreement not being renewed. Describe the Australian competition law risks (if any) with this email.

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.00 | 2.00 | 1.33 | 2.00 | 2.00 |
| Citations | 1.00 | 0.67 | 1.33 | 2.00 | 2.00 |
| Clarity | 2.00 | 0.00 | 1.00 | 0.67 | 1.00 |
| Total | 6.00 | 2.67 | 3.67 | 4.67 | 5.00 |

Question 28 (Banking)

Explain whether a supplier of cloud computing services can terminate their agreement with a company that has just gone into administration based on an ipso facto termination clause?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.67 | 0.00 | 1.67 | 1.33 | 2.00 |
| Citations | 1.67 | 0.67 | 2.00 | 1.33 | 2.00 |
| Clarity | 2.00 | 0.33 | 2.00 | 2.00 | 2.00 |
| Total | 7.33 | 1.00 | 5.67 | 4.67 | 6.00 |

Question 29 (Banking)

Can an Australian company give financial assistance to support the acquisition of shares in its Australian holding company?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|--------------|-------------|-------------|-------------|-------------|
| Substance | 5.00 | 4.67 | 2.00 | 2.00 | 2.00 |
| Citations | 3.00 | 2.33 | 2.33 | 2.00 | 3.00 |
| Clarity | 2.00 | 2.00 | 2.00 | 1.00 | 1.67 |
| Total | 10.00 | 9.00 | 6.33 | 5.00 | 6.67 |

Question 30 (Banking)

The PPSA allows for security in relation to certain financial assets such as bank accounts, intermediated securities and investment instruments to be perfected by way of 'control'. How will a security interest over such an investment instrument be perfected by 'control'?

| Model | GPT-4 | Gemini 1 | Claude 2 | LLaMa 2 | Perplexity |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Substance | 3.00 | 2.67 | 1.00 | 4.00 | 3.00 |
| Citations | 3.00 | 2.00 | 0.00 | 0.00 | 1.00 |
| Clarity | 2.00 | 0.67 | 2.00 | 1.00 | 2.00 |
| Total | 8.00 | 5.33 | 3.00 | 5.00 | 6.00 |