

20.10.2025

www.theaiassembly.com



Last Updated: October 2025

Use this checklist to assess your readiness for California's new Al safety requirements and prepare for similar regulations that may emerge globally.

Effective Date: January 1, 2027 (SB243)

## **B SB 243: Al Chatbot Safety Requirements**

### **Does Your Business Use AI Chatbots?**

- We use AI chatbots for customer service
- We use Al chatbots for marketing or sales
- We use Al companion or conversational Al products
- We develop or sell AI chatbot products
- None of the above (skip to next section)

### **Al Disclosure Requirements**

- Our chatbots clearly identify themselves as AI (not human)
- The Al disclosure is shown at the start of every conversation
- The disclosure is clear and understandable (not buried in fine print)
- We have tested the disclosure with actual users to ensure clarity

Action Required: If any boxes are unchecked, update your chatbot interface to include clear Al identification.



## **B SB 243: Al Chatbot Safety Requirements**

# Minor Protection Measures Do you serve users under 18?

- Yes, we knowingly serve minors
- Possibly we don't have age verification
- No, our service is 18+ only

### If you serve or may serve minors:

- We have implemented usage break prompts every 3 hours for users under 18
- We have age verification systems in place to identify minor users
- Our break prompts are clear and actionable (not just a dismissible notification)
- We log when break prompts are shown for compliance documentation

Action Required: Implement usage break functionality and age verification before January 1, 2027.

### Safety Protocol Implementation

- We have systems to detect harmful content (suicide ideation, self-harm, abuse)
- Our Al is trained to recognise crisis situations involving minors
- We have integrated crisis hotline alerts for high-risk situations
- We have documented our safety protocols in writing
- We test our safety systems regularly (at least quarterly)
- We have a process for reporting safety incidents to authorities



### B SB 243: Al Chatbot Safety Requirements

### **Crisis Hotlines to Integrate: New Zealand:**

Lifeline: 0800 543 354

Youthline: 0800 376 633

1737 (Need to talk?)

### International (if serving global users):

US: 988 Suicide & Crisis Lifeline

Australia: 13 11 14 (Lifeline Australia)

**UK: 116 123 (Samaritans)** 

### **Content Restrictions for Minors**

- Our chatbot does not expose minors to sexual content
- Our chatbot does not engage in inappropriate conversations with minors
- We have content filtering specifically for minor users
- We regularly audit chatbot conversations with minors for compliance
   Action Required: Implement age-appropriate content filtering and monitoring.

### **Legal and Compliance Readiness**

- We have reviewed the full text of SB 243
- We have consulted with legal counsel about compliance
- We have assigned responsibility for AI safety compliance
- We have budgeted for compliance implementation
- We have a timeline for implementation
- We understand the penalty structure (\$2,500 negligent / \$7,500 intentional per child)
- We have liability insurance that covers Al-related incidents

Action Required: Complete legal review and assign compliance ownership.





### AB 56: Social Media Mental Health Warnings

### Does Your Platform Qualify as Social Media?

- We operate a social media platform
- We operate a social networking service
- We have user-generated content and social features
- None of the above (skip to next section)

### **Mental Health Warning Implementation**

- We display mental health warnings to users
- Marnings are visible and not hidden in settings or terms
- Warnings link to mental health resources
- We have tested warning effectiveness with users

Suggested Warning Text: "Social media use may affect mental health. If you're experiencing anxiety, depression, or other mental health concerns, please reach out to a mental health professional or contact [local mental health resource]."

Action Required: Add mental health warnings to your platform.





### AB 621 Deepfake Content Policies

#### **Does Your Platform Host User Content?**

- Users can upload images to our platform
- Users can upload videos to our platform
- We use AI to generate images or videos
- None of the above (skip to next section)

### **Deepfake Prevention Measures**

- We have policies prohibiting deepfake pornography
- We have detection systems for synthetic media
- We have a reporting mechanism for deepfake content
- We have a takedown process for violating content
- We document all deepfake incidents for legal compliance
- We cooperate with law enforcement on deepfake cases

Action Required: Implement deepfake detection and removal processes.





### **Are You a Device Maker or App Store Operator?**

- We operate an app store or marketplace
- We manufacture devices with app stores
- We distribute apps through third-party stores
- None of the above (skip to next section)

### **Age Verification Implementation**

- We have age verification tools in our app store
- We verify user ages before app downloads (for age-restricted apps)
- We have parental consent mechanisms for minors
- We document age verification attempts for compliance
- We have tested our age verification system for accuracy

Action Required: Implement age verification before app downloads.





### **Are You Prepared for Similar Laws in Other Jurisdictions?**

- We monitor Al regulations in our key markets
- We have a regulatory compliance team or consultant
- We design Al systems with privacy-by-design principles
- We document our Al training data sources (for copyright compliance)
- We have transparency reports for Al use
- We have an AI ethics policy in place
- We train staff on responsible Al use

Action Required: Adopt proactive compliance practices for future regulations.



### **M** Compliance Score

Count your checked boxes:

- 90-100% checked: Excellent You're well-prepared for California's Al laws
- 70-89% checked: Good Some work needed before January 2027
- 50-69% checked: Fair Significant implementation required
- Below 50% checked: Action Required Start compliance planning immediately





#### **Immediate Actions**

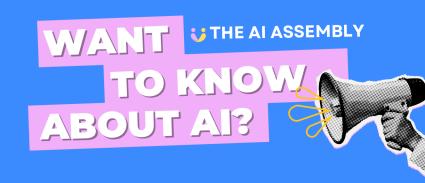
- 1. Assign compliance ownership to a specific person or team
- 2. Conduct legal review of California Al laws with counsel
- 3. Audit current Al systems against this checklist
- 4. Create implementation timeline with milestones

### **Short-Term Actions (Next 3-6 Months)**

- 1. Develop Al disclosure systems for all chatbots
- 2. Implement age verification where needed
- 3. Build safety detection protocols for harmful content
- 4. Integrate crisis hotline alerts
- 5. Train staff on new requirements

### **Long-Term Actions (Before January 1, 2027)**

- 1. Complete all technical implementations
- 2. Test all compliance systems thoroughly
- 3. Document all compliance measures
- 4. Conduct final legal review
- 5. Prepare for ongoing monitoring and reporting





#### **Immediate Actions**

- 1. Assign compliance ownership to a specific person or team
- 2. Conduct legal review of California Al laws with counsel
- 3. Audit current Al systems against this checklist
- 4. Create implementation timeline with milestones

### **Short-Term Actions (Next 3-6 Months)**

- 1. Develop Al disclosure systems for all chatbots
- 2. Implement age verification where needed
- 3. Build safety detection protocols for harmful content
- 4. Integrate crisis hotline alerts
- 5. Train staff on new requirements

### Long-Term Actions (Before January 1, 2027)

- 1. Complete all technical implementations
- 2. Test all compliance systems thoroughly
- 3. Document all compliance measures
- 4. Conduct final legal review
- 5. Prepare for ongoing monitoring and reporting





📞 Resources and Support

### **Legal Resources:**

California Attorney General's Office: www.oag.ca.gov

New Zealand Privacy Commissioner: <a href="www.privacy.org.nz">www.privacy.org.nz</a>

### **Technical Resources:**

Al Safety Guidelines: <u>www.nist.gov/ai</u>

Age Verification Standards: www.w3.org

### **Crisis Support Resources:**

New Zealand Mental Health Foundation: www.mentalhealth.org.nz

Lifeline Aotearoa: 0800 543 354

### The AI Assembly Community:

Access expert guidance and peer support

Stay updated on regulatory changes: www.theaiassembly.com/web-hub

Document Version: 1.0

Last Updated: October 20, 2025 Next Review: January 2026