**AI CHALLENGE Hackathon Rubric**

**Purpose of the Rubric:** This rubric is designed to provide detailed, transparent, and justice-oriented evaluation criteria for the AI version of the CHALLENGE game. It reflects the project’s commitment to critical pedagogy, anti-assimilationist frameworks, and refugee-centered advocacy. Each category below is essential to ensure that your project remains grounded in ethical, transformative, and pedagogically sound practices.

**CATEGORY 1: Ethical and Advocacy Alignment (20 points)**

**Criteria:**

* Does the AI version align with the game’s original mission: to promote justice, inclusion, and refugee rights?
* Does it avoid harmful, tokenistic, or anti-migrant narratives?
* Are AI agents created with nuance and care, avoiding caricature or ideological extremes?
* Is race, ethnicity, gender, or sexual orientation intentionally excluded to reduce stereotyping?
* Does the game provide a space that encourages empathy, critical thinking, and responsibility?

**Outstanding (18-20):** All design elements and AI responses are carefully aligned with the advocacy goals of the CHALLENGE game. No harmful stereotypes, tokenism, or exclusionary views are present.

**Proficient (14-17):** The advocacy principles are mostly respected; some minor ethical or tone inconsistencies may exist but do not undermine the mission.

**Developing (10-13):** Advocacy elements are present but inconsistently implemented. Some features or responses may unintentionally reinforce problematic narratives.

**Inadequate (0-9):** The design promotes deficit views of refugees or allows xenophobic, neoliberal, or tokenistic discourse to flourish.

**CATEGORY 2: AI Agent Design and Dialogue Dynamics (20 points)**

**Criteria:**

* Are the AI agents realistically and respectfully constructed (age, occupation, SES, education, political ideology)?
* Are political ideologies diverse and grounded in real-world complexity?
* Do agents engage with the participant in a meaningful way that supports dialogue and reflection?
* Is voice-entry enabled and functioning effectively?
* Are negotiation dynamics thoughtfully designed to model real policymaking challenges?

**Outstanding (18-20):** AI agents are nuanced, realistic, and ideologically diverse. Voice-entry functions well and interactions feel organic, multidimensional, and reflective.

**Proficient (14-17):** Agents engage the participant effectively, though dialogue may occasionally feel formulaic. Technical aspects are functional.

**Developing (10-13):** Agents lack depth or produce repetitive or shallow responses. Dialogue system may be limited or glitchy.

**Inadequate (0-9):** Agents are one-dimensional or reinforce stereotypes. Conversations lack critical depth. Voice system is not usable.

**CATEGORY 3: Game Structure and Mechanics (20 points)**

**Criteria:**

* Are the three core phases (individual decision, AI-led group discussion, and reflection) clearly implemented?
* Is the 14-unit budget constraint present and visually accessible throughout the game?
* Are policy options accurately presented and functioning?
* Does the system prevent over-budget selections?

**Outstanding (18-20):** The gameplay structure perfectly mirrors the original. Budget constraints are enforced with clarity. Player decision-making flows seamlessly across all three phases.

**Proficient (14-17):** Most mechanics are properly implemented. Minor issues in tracking budget or transitioning between stages.

**Developing (10-13):** Budget enforcement is inconsistent or unclear. Transitions between phases are clunky.

**Inadequate (0-9):** The game lacks the core structure or allows players to bypass major rules like budgeting.

**CATEGORY 4: Reflective and Feedback System (20 points)**

**Criteria:**

* Are reflective prompts integrated at the end of the game?
* Does the AI provide personalized, justice-oriented feedback based on player decisions?
* Does the system evaluate the overall policy package in terms of inclusion, long-term impact, and alignment with justice?
* Are ethical dilemmas acknowledged and discussed in feedback?

**Outstanding (18-20):** Reflection prompts are rich, varied, and tailored. AI feedback is personalized, deeply ethical, and pedagogically sound.

**Proficient (14-17):** Prompts and feedback are thoughtful and mostly accurate, with some room for depth or customization.

**Developing (10-13):** Prompts are generic or repetitive. Feedback lacks depth or clear ties to the player’s decisions.

**Inadequate (0-9):** Reflection is absent or shallow. No meaningful feedback is provided.

**CATEGORY 5: Design, Interface, and User Experience (10 points)**

**Criteria:**

* Is the interface intuitive, accessible, and visually clear?
* Can users easily navigate through all phases of the game?
* Is the game playable on standard platforms?
* Does it accommodate diverse user abilities and needs?

**Outstanding (9-10):** Interface is polished, user-friendly, and supports accessibility. Visual design enhances learning and interaction.

**Proficient (7-8):** Interface is generally smooth with minor usability issues. Accessibility partially considered.

**Developing (5-6):** Design is functional but visually or structurally confusing.

**Inadequate (0-4):** Poor user experience. Confusing layout or inaccessible features.

**TOTAL: 90 POINTS**

**Bonus Points (up to 10):**

* Innovative use of AI to generate dynamic, evolving scenarios
* Ability to analyze emotional tone or hesitation in player input
* Real-time policy consequences during discussion (e.g., population reactions, media coverage)

**Final Notes:** This rubric reflects not just technical execution but the spirit of the CHALLENGE project. You are not being evaluated merely on functionality, but on whether your work helps reimagine refugee education policy through empathy, complexity, and a justice-first mindset. Build responsibly, think critically, and most importantly, advocate boldly.