Historical analysis **Implementation Framework**

To systematically analyze historical patterns of discrimination in technology, follow this structured methodology:

1. **Historical Pattern Identification**
   * Examine historical discrimination in the specific domain where the AI system will be deployed (e.g., healthcare, criminal justice, hiring).
   * Research how previous technologies in this domain have reflected, reinforced, or challenged existing social hierarchies.
   * Identify recurring mechanisms through which bias has persisted across technological transitions in this domain.
2. **Pattern-to-Risk Mapping**
   * Map identified historical patterns to specific components of the ML system under consideration.
   * Determine how historical classification systems might influence feature definitions and encodings.
   * Assess how historical performance disparities might manifest in model accuracy across groups.
   * Analyze how historical optimization priorities might shape evaluation metrics and thresholds.
3. **Prioritization Framework**
   * Assess the strength of the historical connection between identified patterns and the current application.
   * Evaluate the potential harm if historical patterns were to recur in the current system.
   * Determine the visibility of potential bias (some forms are more readily apparent than others).
   * Prioritize which historical patterns require particular attention in subsequent fairness assessments.

Data representation **Implementation Framework**

To systematically examine social contexts in data representation, implement this structured methodology:

1. **Classification Audit:**
   * Document all classification systems used in your dataset, including protected attributes and potential proxies.
   * Research the historical development of these classification systems to identify potential embedded biases.
   * Evaluate whether classifications capture the full diversity of the relevant population or artificially constrain representation.
   * Analyze whether classification boundaries reflect meaningful distinctions for your task or arbitrary divisions with historical baggage.
2. **Representation Gap Analysis:**
   * Compare dataset demographics to relevant population benchmarks to identify underrepresented groups.
   * Document variables that might be systematically missing for certain populations.
   * Examine whether proxy variables have consistent meaning across demographic groups.
   * Identify potential "data deserts" where information is systematically sparse due to historical collection patterns.
3. **Codification Evaluation:**
   * Analyze how social categories are mathematically encoded and what assumptions these encodings embed.
   * Test whether distance metrics in feature space reflect meaningful similarities or arbitrary technical choices.
   * Evaluate whether technical implementations preserve important social nuances or flatten complex identities.
   * Document potential information loss in the translation from social categories to computational representations.
4. **Power Analysis in Data Lifecycle:**
   * Document who designed the data schema and whether diverse perspectives informed these decisions.
   * Analyze who collected the data and under what conditions (mandatory compliance, voluntary participation, etc.).
   * Evaluate who labeled or annotated the data and what guidelines shaped these interpretations.
   * Assess who validates data quality and what metrics they use to determine representativeness.

Ethical Frameworks for fairness evaluation **Implementation Framework**

To systematically apply ethical frameworks to fairness evaluation, implement this structured methodology:

1. **Multi-framework Analysis:**
   * Analyze fairness questions through at least three distinct ethical lenses (e.g., consequentialist, deontological, and virtue ethics).
   * Document how different frameworks would evaluate key decisions about data collection, feature selection, and optimization objectives.
   * Identify areas where frameworks agree and where they suggest different approaches.
2. **Stakeholder-Centered Ethical Mapping:**
   * Identify which ethical frameworks most closely align with different stakeholder groups' priorities and values.
   * Map potential conflicts between stakeholder ethical perspectives.
   * Document how these different ethical perspectives might evaluate system outcomes differently.
3. **Contextual Adaptation:**
   * Analyze how application domain and cultural context should influence which ethical frameworks receive priority.
   * Consider historical patterns of discrimination in the domain and which ethical frameworks best address these patterns.
   * Document domain-specific ethical considerations that standard frameworks might miss.
4. **Trade-off Documentation:**
   * Create explicit documentation of ethical trade-offs when different frameworks suggest conflicting approaches.
   * Develop clear rationales for prioritizing certain ethical considerations over others in specific contexts.
   * Establish processes for revisiting these trade-offs when conditions change.

Modern manifestations of historical biases **Implementation Framework**

To systematically analyze how historical biases manifest in modern AI systems, follow this structured methodology:

1. **Map Historical Patterns to Technical Mechanisms**
   * Identify relevant historical discrimination patterns for your application domain.
   * Document specific technical mechanisms through which these patterns might manifest in modern systems.
   * Analyze which components of your ML pipeline are most vulnerable to these specific patterns.
2. **Conduct Proxy Variable Analysis**
   * Examine your feature set for variables that might correlate with protected attributes.
   * Quantify these correlations using techniques like mutual information or prediction tasks.
   * Document which historical patterns these proxy relationships might reflect.
3. **Perform Feedback Loop Assessment**
   * Identify whether your system's predictions will influence future data collection.
   * Trace potential cycles where initial bias might amplify over time.
   * Develop monitoring approaches to detect emerging feedback patterns.
4. **Implement Domain-Specific Analysis**
   * Apply specialized assessment techniques relevant to your application domain.
   * Reference case studies from similar domains to identify common manifestation patterns.
   * Adapt general fairness approaches to address domain-specific mechanisms.

To agree on which ethical framework: Conduct a multi-framework analysis that documents how different ethical perspectives would evaluate the system, then develop a blended approach that explicitly acknowledges trade-offs and gives priority to addressing the most severe historical health disparities.