Artificial Intelligence for Processing Public Comments

Executive Summary

Public comment processing represents a critical component of environmental regulatory development, with agencies routinely managing thousands of submissions for most policy decisions throughout each year. For example, the Environmental Protection Agency's (EPA) proposed National Primary Drinking Water Regulations (NPDWR) for per-and polyfluoroalkyl substances (PFAS) drew approximately 1,700 unique submissions—each requiring weeks of careful review, categorization, and response.
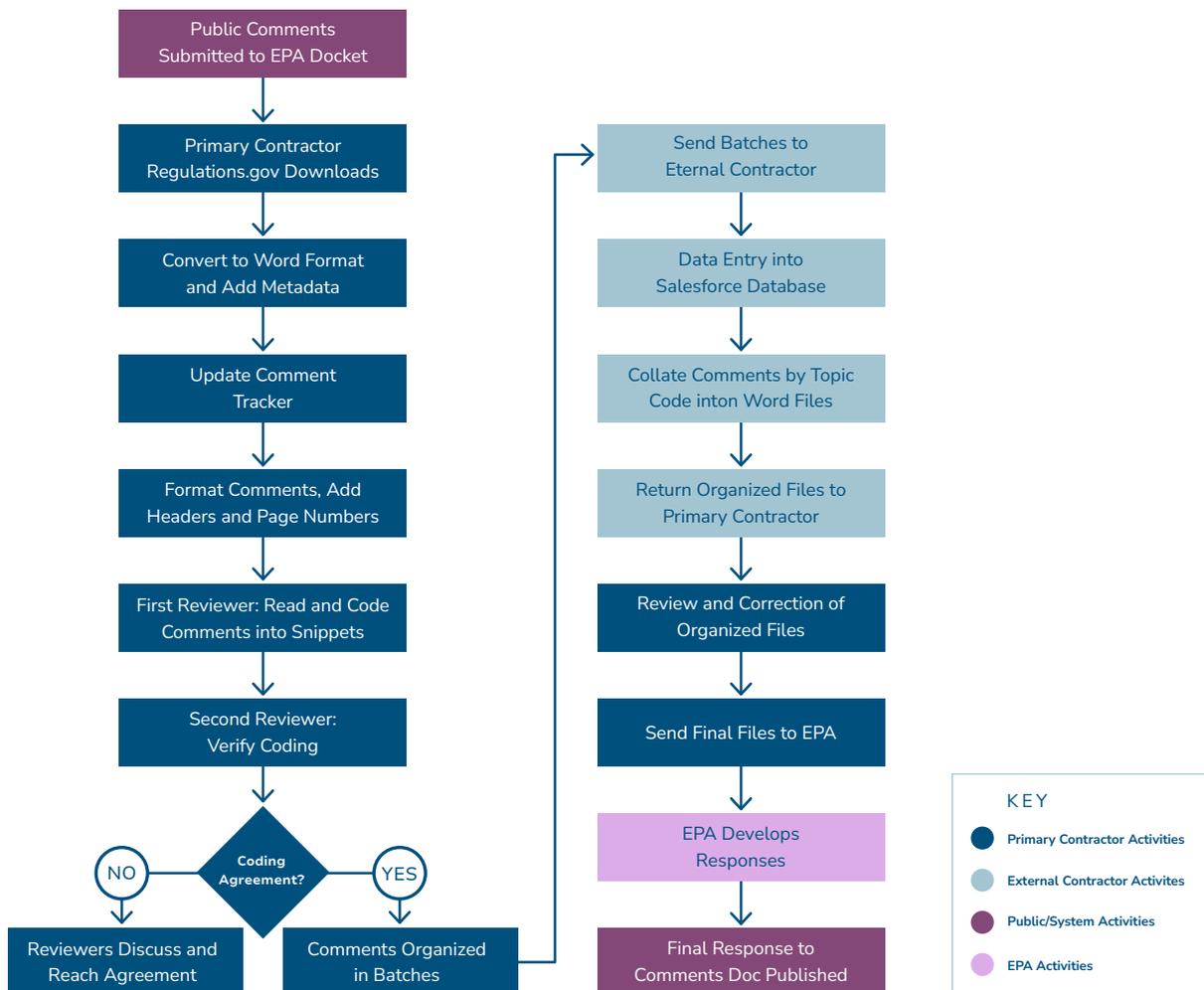
Given the large effort and long processing steps involved in responding to the public comments, it is evident that automation using artificial intelligence (AI) can be strategically integrated into this regulatory workflow, while maintaining human oversight for the accuracy of the critical steps involved.

CADMUS

Current Comment Processing Workflow Analysis

The existing public comment processing system follows a structured workflow beginning with comment submission to the EPA dockets and concluding with published response documents (see Figure 1). Comments are downloaded from regulations.gov, converted to standardized formats, categorized by topic codes, and organized for EPA technical review. Multiple quality assurance steps ensure accurate categorization and comprehensive coverage.

For example, the EPA requested public comment on the proposed PFAS NPDWR in March 2023 and received approximately 122,200 public comments by May 2023. Of those comments, approximately 120,500 were mass mailers from various activist groups and organizations. The comment processing involves downloading individual submissions received in response to the EPA notice, converting to uniform format, extracting metadata, and then systematic coding by trained reviewers. Comments containing attachments, figures, and tables require special handling procedures to preserve content integrity. The process culminates with external contractor support for database management and comment collation by topic areas before final EPA review and response development.

Figure 1. Current Public Comment Processing Workflow

Public Comments Submitted to EPA Docket

Primary Contractor Regulations.gov Downloads

Convert to Word Format and Add Metadata

Update Comment Tracker

Format Comments, Add Headers and Page Numbers

First Reviewer: Read and Code Comments into Snippets

Second Reviewer: Verify Coding

Coding Agreement? — NO → Reviewers Discuss and Reach Agreement

Coding Agreement? — YES → Comments Organized in Batches

Send Batches to Eternal Contractor

Data Entry into Salesforce Database

Collate Comments by Topic Code inton Word Files

Return Organized Files to Primary Contractor

Review and Correction of Organized Files

Send Final Files to EPA

EPA Develops Responses

Final Response to Comments Doc Published

KEY
- Primary Contractor Activities
- External Contractor Activites
- Public/System Activities
- EPA Activities

Current Workflow Roles

The comment processing workflow involves three key stakeholders:

EPA:

- Comment collection setup
- Response development
- Final document publication

Primary Contractor:

- Comment downloading and formatting
- Topic coding and verification
- Quality review and correction
- EPA coordination

External Contractor (Database & Collation Services):

- Database data entry
- Comment organization by topic codes

Current Challenges & Limitations

The current workflow involves significant manual processing which creates several operational challenges. Document format inconsistencies require individual attention for proper conversion and formatting. Review of submissions and topic assignment depend on human reviewers carefully reviewing text from different commenters, creating potential for inconsistent categorization across large comment volumes. Quality assurance protocols requiring dual reviewer verification double the processing time while attempting to ensure accuracy.

In addition, the system treats diverse comment types uniformly, from brief individual submissions to comprehensive technical analyses from industry stakeholders. While thorough, this approach does not optimize processing efficiency based on content complexity or stakeholder category. Scaling challenges become apparent as comment volumes increase for major environmental regulations.

Opportunities for Using AI Technologies

As described in previous sections, the current multi-step process involves document conversion, content review, topic classification, quality assurance, and response coordination—activities that consume significant staff resources and extend regulatory timelines. AI applications can be designed to streamline these workflows through automated document processing, intelligent content classification, semantic analysis, and quality assurance support, while preserving essential human oversight for policy decisions and technical judgments. The following are some of the distinct areas where AI technologies can optimize processes:

Document Processing and Analysis. AI technologies can automate the conversion and initial analysis of comment submissions across various formats. Automated document processing can handle format standardization, extract metadata, and identify document types without manual intervention. Content analysis capabilities can segment longer documents into coherent sections and identify technical versus general content for appropriate routing. Synthesize all scientific and technical findings.

Explore Our People-Driven, AI-Empowered Approach

See how we harness AI to augment processes and workflows, accelerate innovation, drive greater efficiencies, and deliver more value.

Content Classification and Organization. Machine learning systems can learn from historical comment categorization patterns to assign topic codes consistently across large volumes of submissions. These systems can identify recurring themes, cluster similar content, and flag unique technical arguments that require specialized attention. Automated classification addresses consistency challenges that arise from reviewer fatigue while maintaining systematic organization of diverse comment types.

Commenter Analysis. AI systems can automatically categorize commenters by organization type, identify mass submission campaigns, and analyze stakeholder representation patterns. This capability enables agencies to understand the full spectrum of public input and prioritize responses based on commenter categories and content uniqueness.

Quality Assurance Enhancement. Automated validation systems can serve as intelligent reviewers, identifying inconsistencies in classification, flagging content that may require specialized expertise, and ensuring comprehensive coverage of comment themes. These systems can maintain quality standards while reducing the manual verification burden on human reviewers.

Technical Implementation Approach

System Architecture. The technical architecture should maintain EPA data ownership while enabling contractor processing capabilities. Processing systems can be implemented to operate with secure data handling protocols. This ensures no permanent data storage on contractor systems, returning all outputs to EPA-controlled repositories. The architecture must support scalable processing capabilities that can accommodate varying comment volumes across different regulatory proceedings while maintaining consistent processing standards.

Security and Oversight Framework. Implementation requires comprehensive security measures throughout the processing workflow, with particular

attention to data protection during analysis phases. Human oversight should remain central to the system design, with AI providing processing support rather than replacing human judgment on policy matters. Quality control protocols need to ensure appropriate validation of AI-generated outputs before final review.

Integration Strategy. AI capabilities can be integrated with existing regulatory workflows through phased implementation, beginning with routine processing tasks and expanding to more sophisticated applications as organizational confidence develops. The approach should maintain compatibility with current systems while enabling enhanced processing efficiency.

Benefits & Limitations

Operational Benefits. AI integration can significantly reduce processing time for routine comment categorization and organization tasks. Improvements in consistency result from standardized classification approaches that do not suffer from reviewer fatigue or subjective interpretation variations. Scalability enhancements enable agencies to process larger comment volumes without proportional increases in staff resources.

Resource optimization allows technical staff to focus on policy analysis and response development rather than routine processing tasks. Improved tracking and organization of comment themes enables more comprehensive regulatory analysis. Enhanced stakeholder analysis provides better understanding of public concerns and regulatory impact distribution.

Initial cost-benefit estimates suggest that AI implementation could reduce comment processing time by approximately 30–40% while improving classification consistency, leading to potentially large cost savings annually on regulatory proceedings. These efficiencies mainly come from elimination of repetitive manual work, specifically the downloading, converting, and consolidating of comments into a single format for subsequent analysis and categorization. In an AI-augmented world, comments

could be automatically downloaded using triggers, and an AI categorization agent could tag or categorize text based on human-defined categories. Then, human reviewers can verify the AI-generated categorization of comments for accuracy by scanning for keywords and modifying as necessary. Return on investment becomes particularly compelling as comment volumes continue to increase with expanded public engagement in environmental regulations.

Technical Limitations and Risk Management. While automated comment processing significantly reduces manual labor and increases efficiency, AI systems are not without risks. AI systems may struggle with highly specialized technical content or novel regulatory arguments not represented in training data. Content analysis can miss nuanced policy implications that require deep regulatory expertise. Classification accuracy depends on the quality and comprehensiveness of historical training data.

Strategies to mitigate these risks include confidence scoring systems that route uncertain cases to human review, addressing accuracy limitations while maintaining processing efficiency. Starting with small scale pilot implementations on less critical comment categories allow system validation and staff training before full deployment. Ongoing monitoring and model refinement ensure continued accuracy as regulatory contexts evolve.

Strategic Recommendations and Conclusions

Environmental agencies should begin exploring AI applications through targeted pilot projects that demonstrate clear operational benefits while maintaining regulatory standards. Initial implementations should focus on automation of intensive processing steps using AI in the areas described, with clear success metrics and limited risk exposure. Gradual expansion can incorporate more sophisticated capabilities as organizational confidence and technical competence develop.

The regulatory comment processing challenge will continue to grow as public engagement increases and regulatory complexity expands. Agencies that proactively develop AI capabilities will be better positioned to maintain thorough regulatory analysis while managing expanded public participation. Successful implementation requires careful attention to data security, human oversight, and gradual capability development.

The AI technology described in the application areas exists today to begin this transformation while preserving the rigor and transparency that environmental regulations require. Organizations should prioritize pilot projects that deliver measurable improvements to build confidence and expertise before expanding AI applications across broader regulatory processes.

CADMUS