MOATIVE

Applied AI Services & Venture Labs

Computer-Vision Based Asset Registry

Implementation Proposal

Proposed for

Digital Operating Company / Black Oil Partners

23 JanuaryJanuary 2026

Executive Summary

Digital Operating Company (DOC) is an asset-focused operating company that focuses on leveraging artificial intelligence and machine learning to optimize operations, aiming for efficiency and cost-effectiveness.

DOC currently manages over 700 oil and gas wells across Texas, North Dakota, and Montana. While their portfolio is growing rapidly, the process of creating a comprehensive asset inventory has not kept pace. The process remains primarily manual: field teams photograph equipment, hand-record tank sizes and configurations, and enter data into Notion and spreadsheets. This approach takes time and effort, is not regular and is not scalable.

The current process is typical of other players in the market as well. DOC senses an opportunity to build a computer-vision based asset registry that could be monetized.

Based on conversations with DOC, Moative proposes to implement a scalable process to automate asset identification and documentation. This process will leverage computer-vision and AI-ML algorithms to extract information from field images and drone footage. The system will identify equipment types, classify sizes, and populate a structured database that serves as the single source of truth for downstream operations, regulatory compliance, and financial valuation.

Moative proposes a four-week engagement that will deliver a functional version 1.0 of the detection system. This includes a one-week data assessment, model development covering the top equipment types (based on the number and importance of assets), and a working tagging and documentation pipeline with accuracy benchmarks. DigitalOpCo owns all trained models and code outright with no ongoing licensing fees or tooling vendor dependencies.

The work shall be billed on a ‘time and materials’ basis, owing to the inherent nature of having to experiment with model accuracy based on the data availability. We estimate that four-weeks is a reasonable aspiration to begin with and adjust as needed. Assuming full utilization of the team, Moative shall bill every two weeks by submitting a time sheet. We anticipate a weekly burn of $7000. The number of weeks is indicative. As we plan to productionize we may spend time beyond the first four weeks. Such time too will be billed on a pro-rated basis.

Current State Analysis

When acquiring or evaluating oil and gas assets, DigitalOpCo's current process relies heavily on manual effort. Field personnel visit each well site, photograph equipment, and record observations using JotForm with basic image uploads. This data is then manually entered into Notion databases and Excel spreadsheets, with labeling conventions varying across team members and time periods.

DigitalOpCo reports having invested significant effort in data capture and preliminary labeling, including hundreds of images with equipment annotations, drone footage of well sites, and cropped/categorized images organized in OneDrive folders.

We will assess the machine-learning readiness of this data in Week 1 and proceed with training immediately if annotations are usable.

Technical Approach

The solution applies modern object detection techniques to identify and classify oil field equipment from images. We will use a transformer-based detection architecture, leverage public datasets to accelerate training, and deliver a system that DigitalOpCo owns outright.

Detection Architecture

The core detection system will be built on a single multi-class object detection model trained to recognize all equipment types simultaneously. Rather than training separate models for tanks, separators, and wellheads, a unified model learns shared visual features across equipment categories, improving generalization and simplifying deployment.

Moative shall use **RT-DETR** (Real-Time Detection Transformer) as the primary architecture. RT-DETR provides excellent global context understanding, critical for accurately counting multiple tanks in a single frame and understanding spatial relationships between equipment. Its hybrid encoder architecture achieves strong accuracy without the post-processing complexity of traditional detectors.

Since DigitalOpCo's workflow is batch-oriented rather than real-time and the images are uploaded after field visits, not streamed continuously, we optimize for accuracy and robustness over raw inference speed. RT-DETR processes images in under 5 seconds on standard cloud compute, which is more than sufficient for batch processing.

If RT-DETR underperforms on the specific characteristics of DigitalOpCo's data, we shall pivot to an alternative architecture (DINO or Faster R-CNN). **Such a pivot may add approximately one more week to the timeline.**

Leveraging Public Datasets

We may not have to start image recognition from scratch. Several public datasets contain relevant oil and gas infrastructure that can accelerate model training:

Duke/Figshare Above-Ground Storage Tank Dataset: 130,000+ annotated tanks across the continental US, including external floating roof, closed roof, and spherical pressure tanks. This provides an excellent pretraining foundation for tank detection.

Stanford OGNet: 7,000+ images of oil refineries and petroleum terminals with storage tank counts. CC BY 4.0 licensed.

Permian/Denver Well Pad Dataset: 70,000+ well pads and 169,000+ storage tanks from satellite imagery, with high-precision annotations (0.96 precision). Published in Nature Communications.

DIOR and HRRSD: Standard remote sensing benchmarks containing storage tank classes useful for transfer learning.

Our approach: pretrain on these public datasets for common equipment (especially tanks), then fine-tune on DigitalOpCo's domain-specific ground-level imagery. This dramatically reduces the labeling burden and improves accuracy on equipment types that are underrepresented in public data (such as pump jacks, separators, meters, and valves) where DigitalOpCo's proprietary images provide unique value.

Note: We have not done a deep analysis of the data sets. So their relevance will be assessed during the first week.

Training Strategy

Moative team shall fine-tune models pre-trained on large-scale datasets (COCO, Objects365, plus the oil and gas datasets above) that have already learned general visual features like edges, textures, and shapes.

The fine-tuning process works in two stages. First, we freeze the backbone network and train only the detection head on DigitalOpCo's equipment data. This allows the model to learn domain-specific class boundaries without disrupting learned features. Second, we unfreeze later layers of the backbone and continue training with a reduced learning rate, allowing the model to adapt to the specific visual characteristics of oil field equipment.

With this approach and the public dataset foundation, we expect reasonable detection performance with 50-100 labeled examples per equipment type from DigitalOpCo's data.

Size Classification

Determining physical dimensions from 2D images without depth information is fundamentally challenging. Our approach uses categorical classification (small, medium, large) rather than absolute capacity estimation. The model learns size categories from training examples where ground truth is known, providing reliable differentiation between, say, 200-barrel and 400-barrel tanks based on learned visual priors.

Capacity estimates will be approximate and should be validated for critical applications.

Confidence Calibration

Raw model confidence scores often do not reflect true prediction reliability. We will implement temperature scaling during model validation to calibrate confidence scores against observed accuracy. The calibrated confidence threshold for automated acceptance (initially set at 85%) will be tuned on a held-out validation set to achieve the desired balance between automation rate and error rate.

We may need continuous support from the DOC team for such validations.

System Architecture

The production system consists of four components working in sequence:

Ingestion Pipeline: Images from JotForm submissions, OneDrive folders, or direct uploads are normalized and queued for processing. Metadata (well ID, capture date, GPS coordinates if available) is extracted and associated with each image.

Inference Engine: The trained model processes images and outputs detections with bounding boxes, class labels, size classifications, and calibrated confidence scores. The processing time target is under 5 seconds per image.

Routing Logic: Detections above the confidence threshold are automatically accepted. Detections below threshold are flagged for human review. Human review would require participation from the DOC team.

Database Population: Validated detections are written to the asset database with full provenance tracking.

Licensing and Ownership

All model architectures and training tooling are released under the **Apache 2.0 license**, a permissive open-source license with no copyleft obligations. DigitalOpCo receives:

No ongoing licensing fees. The models and code delivered are yours outright with no recurring costs.

No vendor lock-in. You can modify, extend, or redeploy the trained models without restriction.

Clean IP transfer. Trained model weights, inference code, and training pipeline become DigitalOpCo assets upon delivery.

Implementation Roadmap

Week 1: Data Assessment

Before committing to the training schedule, Moative will conduct a thorough assessment of existing data and documentation regarding the assets.

Activities:

1. Access information in DOC's Notion databases, OneDrive folders, and JotForm systems;
2. Examine the documents, analyze how the labels are stored and whether they can be programmatically converted to bounding box format;
3. Review a stratified sample of 50-100 images for quality, consistency and ability to extract information programmatically;
4. Finalize asset taxonomy and priority, classify assets into Tier 1, 2 and 3 based on quality and priority. We will focus on Tier 1 & 2 for this 1-month project.
5. Determine information to be extracted, and identify gaps in coverage or annotation.

Deliverable: Data Readiness Assessment Report with summary of available documentation regarding the assets, annotation format evaluation, finalized equipment taxonomy with priority ranking. We will also provide the final project plan and timelines, based on the data assessment.

If data requires significant transformation or re-annotation, **this adds 1-2 weeks to the timeline** before model development begins. We will scope this explicitly at the end of Week 1 if needed.

Week 2: Model Foundation

In this phase, we set up the infrastructure and begin model training based on DOC's annotated documents, augmented by publicly available annotated datasets. We will validate results on a small test set before expanding to the larger document set

Activities:

1. Set up training infrastructure (cloud compute, experiment tracking).
2. Download and prepare public datasets (Duke AST, OGNet).
3. Prepare DigitalOpCo's training and test dataset based on Assessment findings.
4. Train baseline model using transfer learning from public data.
5. Validate RT-DETR performance on initial test set.

Deliverable:

1. Operational model training and validation infrastructure,
2. validated baseline model with initial metrics on asset detection and tagging,
3. Setup HITL process for tagging/ validation if required.

Week 3: Core Detection

In this phase, we will expand the baseline model to cover all major asset classes. We will also implement confidence scoring for each information extraction. These confidence scores will be validated both algorithmically and HITL calibration. Build database schema and ingestion pipeline. Benchmark detection accuracy across all equipment types.

We will also fine tune the model for Tier 1 assets (such as tanks, separators, wellheads, meters, valves). The goal is to have a largely-automated process that can identify, tag and document information about Tier 1 assets with high confidence.

Deliverables:

1. Expanded model to cover all asset classes, with confidence scoring
2. Automated detection system to process images, identify and extract information and ingest that into a database schema
3. Fine tuned model(s) and pipelines for Tier 1 assets (5-7 core types) that meet expected confidence thresholds.

Week 4: Extended Coverage & Delivery

In this phase, we will focus on fine tuning the model to Tier 2 assets. The objective is to improve identification and tagging to meet the target quality requirements.

In parallel, we will also fine-tune the extraction of some fine grained information such as Implement asset size or model. This size classification (small/ medium/ large) will be implemented for asset types where image-based size differentiation is reliable.

The confidence thresholds will be tuned for Tier 1 & Tier 2 assets on a "golden" validation set to enable accuracy benchmarking.

Other tasks include (i) end-to-end pipeline automation and testing, (ii) user validation and testing, and (iii) Documentation and knowledge transfer.

Deliverable:

Functional v1 system with detection and size classification, along with final accuracy report, and technical documentation.

Note: The v1 system that will be delivered at this point will be functional and usable. But this is not yet a polished product ready for prime time. As a fully functional prototype, the product can be used for internal testing, validations, and external demos.

The V1 System can be used by DOC's internal team on their own for 2-4 weeks. During this phase the team can document feedback about the V1 System's functionality, limitations and any enhancements that may be required. Consider the deliverable for this phase 4 as a 'Handover for testing' version.

We will then expand to a full production-ready system, including feedback from DOC's team, based on discussions with DOC. That will be considered as a separate project/ SOW.

Success Criteria

Metric	Target	Minimum
Equipment types detected	15 types	10 types
Detection accuracy (mAP@0.5)	80%	70%
Size classification accuracy	75%	65%
Processing time per image	< 5 seconds	< 15 seconds

Risks and Contingencies

Computer vision projects carry inherent technical risks. We address these directly with clear contingency plans.

Data Preparation

If existing annotations require format conversion or mapping reconstruction, we will scope the preparation work at the end of Week 1. Minor transformations typically add 1 week; substantial re-annotation with an annotation team adds 2-3 weeks. This cost is only incurred if Assessment findings require it.

Architecture Performance

If RT-DETR underperforms on DigitalOpCo's specific data characteristics (unusual lighting, equipment configurations, or image quality patterns), we will pivot to an alternative architecture (DINO or Faster R-CNN). This pivot adds approximately 1 week to the timeline.

Image Quality Variability

Field images vary in lighting, angle, weather, and resolution. We address this through aggressive data augmentation during training and by flagging ambiguous images for human review rather than forcing incorrect classifications.

Edge Cases

With 700+ wells across multiple states, some equipment configurations will be non-standard. The system flags unknown or low-confidence detections for manual classification rather than guessing. Detection accuracy will naturally be higher for equipment types with more training examples; rare or non-standard configurations may require additional annotation over time.

Team Matrix and Commercials

Team Composition

Role	Allocation	Rate
Sr. AI/ML Engineer	Full-time	$45/hr
AI/ML Engineer	Full-time	$40/hr
Project Manager	20%	$75/hr

Time and Materials Engagement

All work is billed on a time and materials basis at the rates above. Based on the scope described, we estimate the following:

Deliverable	Indicative Duration*	Indicative Cost*
Data Assessment + v1 Detection System	~4 weeks	**~$28,000**

* Timeline and cost variations may happen but in such cases both Moative and DOC shall mutually agree if the additional time/cost are economically worthy.

Additional Effort (If Required)

The following may be required depending on data assessment findings. These are estimated based on the same rates:

Contingency	*Duration*	*Cost*
Data preparation (format conversion)	+1 week	+$6,000
Architecture pivot (RT-DETR → alternative)	+1 week	+$6,000
Significant re-annotation (with annotation team)	+2-3 weeks	+$10-15K

Moative covers cloud compute costs for model training during development. DigitalOpCo is responsible for production inference infrastructure post-deployment.

Assumptions and Dependencies

Data Assumptions

We assume annotated images exist in sufficient quantity (target: 100+ for initial training) and that image-to-label mappings are reconstructible from existing systems. Equipment types are limited to 20-30 distinct categories. These assumptions will be validated in Week 1.

Client Dependencies

Data access must be provided within 3 business days of kickoff. A domain expert (Fred or designee) is available 2-3 hours per week for taxonomy validation and domain questions. Feedback on deliverables is expected within 3 business days.

Responsibilities

Activity	Moative	DigitalOpCo
Provide data access	—	Owner
Data assessment and model development	Owner	Consulted
Equipment taxonomy definition	Owner	Consulted
Domain validation of outputs	Consulted	Owner

Next Steps

To proceed, DigitalOpCo approves the Statement of Work. At an agreed upon date after the SOW execution, DigitalOpCo provides data access to Notion, OneDrive, and JotForm systems. Week 1 begins thereafter.

At the end of Week 1, we will present the Data Readiness Assessment. If the data is ready, we proceed directly into model development. If preparation work is needed, we will scope the additional time and cost before continuing.