



Data wrangling for
effective Prefeed
to FEED projects
in oil and gas



Copious amounts of files and data can be generated during Prefeed (Front Ending Engineering Design) activities. During Prefeed, organizations focus on technical requirements, estimated costs, and potential risk evaluations for a proposed project. Prefeed activities, files and data packages are essential to ensure:

- Design options have been evaluated
- Risks and uncertainties are understood
- Cost implications such as budgets and schedules are understood

Prefeed proves development feasibility, and ultimately narrows down selection and options, which then become the basis of the project FEED.

As a result of Prefeed activities, the development is ready to enter into the FEED phase, with utilization of well-defined Prefeed option(s). Reaching FEED is a major milestone for any project and can also be accompanied by file and data challenges, such as:

- How the organization intends to manage unstructured Prefeed information either by transitioning to FEED or archiving redundant information
- Implementation of numbering schemas such as document numbers and revision codes, etc.
- Selecting a document management system with manages all documents generated for the complete lifecycle of the project
- Defining system and numbering/coding protocols and processes before entering FEED is crucial to clearly establish robust methods for all files and data generated and managed throughout the life of the project. During this time, organizations need to clarify:
 - Identification of true source information while removing duplicates
 - How numbering and coding will be applied to Prefeed files

- How to migrate Prefeed files and data from unstructured sources into the project information management system
- What secure methods are in place to receive and share project files and data during and after FEED
- Who needs access to Prefeed information

Organizations can tackle the challenges which encompass Prefeed migrations into FEED project environments by following the steps outlined below.

Step 1 Identification: Prefeed files and data for migration to FEED

Firstly, organizations need to identify Prefeed files and metadata within structured and unstructured environments which will be progressed into the FEED phase of the project. This includes identification and quarantine of highly sensitive information and archival of non-critical information - , thus maintaining access should that information be required at a later date. Organizations should also consider and manage the following:

- Identify disparate information sources
- File deduplication ensuring that one true source file is identified and maintained
- Identification of redundant, obsolete and trivial information to ensure ROT is quarantined for archival, or removal
- Compliance with personal and sensitive information requirements

Step 2: Target System Requirements

Once files have been identified for migration to FEED, the next step is to identify the target system where these files will reside.

Organizations should consider the following before migrating files into a new system:

- What are the mandatory system attribution requirements, and how does that compare to what metadata is currently available?



Secure methods to receive and share project files and data during and after FEED

- Will new numbering schemes and taxonomies be applied to Prefeed files, if so:
 - Apply new numbering schemas, while retaining Prefeed history and document origin information
 - Populate supporting system attributes with Prefeed data to improve file and data findability
- Does the target system functionality meet project requirements?
 - Can file renditions be automatically populated from native files?
 - Can the enterprise search engine search content within files, or only document profile attribution?
 - Are all PDFs fully OCR'd (optical character recognition), enabling recognition of text within digital images to maximize data findability by end users?
 - What security and access restrictions are available, and how is that applied during bulk upload?
- Extract metadata from documents and images (document numbers, revisions, authors, well names, well identifiers, dates, etc.)
- Extract complex metadata using client defined taxonomies
- Identify sensitive and legal files so they can be located and appropriately classified to avoid data leakage
- Verify content to ensure scanned information is of high enough quality—such as removal of background noise and de-skewing of images

Different document types have different attribution requirements and some content may need to be converted or repurposed from one format to another before migration, such as:

- Foreign language text converted into an English word file (Note: images/pictures are not converted and transferred to word)
- Identification and removal of illegal characters such as (*%\$£"!@&) within titles and file naming conventions
- Creation of controlled value lists to allow for metadata format standardization (dates, text cases, formats)
- Title description acronyms removed and replaced by whole words, ensuring descriptive titles reflect file contents

Step 3: Populating system attributions, and applying security protocols

Organizations need to identify migration tools which enable the migration of files and data from structured, unstructured environments and from system to system while performing the following during migrations to systems:

Data-wrangling service benefits:

Organizations need to collaborate with IT partners that can offer global best practices, innovative automated AI solutions, and domain expertise to protect and retain the integrity of critical data and files during migration.

Data wrangling services are ideally situated to assist with file and data interrogation, which include the identification, deduplication, metadata extraction, metadata transformation and migration of files and data to target systems while adhering to system attribution alignment requirements.

We have teamed our data-wrangling technologies and resources with our Global IM. SME services to manage the following during migration:

- Offer services for Prefeed to FEED information migration
- Consolidate and create numbering and coding schemas
- Apply global best practices to safeguard the integrity of the information
- Assist in the creation of information management plans, processes and procedures



About the author

Janine Murray

Principal Consultant, Wipro Limited.

Janine Murray is an IM Consultant with over 15 years of experience in the O&G industry. She has extensive FE/Operations and Major Capital Project (MCP) Information Management experience. She has deep experience with IM brownfield modifications, greenfield

enhancements, MCP joint ventures, closeout, and MCP handover to Operations. Additionally, she is experienced with document cleansing and data extraction techniques for digitizing O&G legacy assets. She can be reached at **janine.murray@wipro.com**



Wipro Limited

Doddakannelli, Sarjapur Road,
Bangalore-560 035,
India

Tel: +91 (80) 2844 0011

Fax: +91 (80) 2844 0256

wipro.com

Wipro Limited (NYSE: WIT, BSE: 507685, NSE: WIPRO) is a leading global information technology, consulting and business process services company. We harness the power of cognitive computing, hyper-automation, robotics, cloud, analytics and emerging technologies to help our clients adapt to the digital world and make them successful. A company recognized globally for its comprehensive portfolio of services, strong commitment to sustainability and good corporate citizenship, we have over 175,000 dedicated employees serving clients across six continents. Together, we discover ideas and connect the dots to build a better and a bold new future.

For more information,
please write to us at
info@wipro.com

