

Table 3.1: A summary of publicly available maintenance work order datasets, describing size, unstructured text fields, metadata, and annotated data.

Dataset	Domain	Number of Records	Field for Work Description (unstructured text)	Examples of Fields/Metadata Included (format and description in brackets)	Is the dataset annotated?	Annotation Information Examples
Handyman Work Order (HWO) Charges ^a	Housing and Development	92.4K	HWODescription	ChargedAmount, HWOCreateDate, WorkTypeGeneral, HWOStatusReason	No	N/A
Initiating Events ^b	Unexpected reactor trips at commercial power plants	3,285	IPFShortDesc	EventDt (date), Abstract (long text), VENDOR, UnitMode	No	N/A
WSSC Completed Service Alert Work Orders ^c	County Water Department	260K	N/A	Date Received, Date Completed, WORK DESCRIPTION (text, desc: Description of Work Performed)	No	N/A
MaintNet (Vehicle Historical Maintenance Dataset) ^d	Automotive Maintenance	200	Notes	JobDate, JobTime, jobno (numerical), Reason (code)	No	N/A
MaintNet (Aircraft Historical Maintenance Dataset) ^e	Aviation Maintenance	6,169	Problem, Action	IDENT (code)	No	N/A
MaintNet (Facility Historical Maintenance Dataset) ^f	Facility Maintenance	200	DESCRIPTION	SITE_ID, DATE_REQUESTED, Date Completed, Problem Type (text, example: "BUILDING INTERIOR INSPECTION")	No	N/A
MaintIE (Fine-Grained Expert-Annotated Corpus) ^g	heavy mobile equipment, rail, mineral processing and infrastructure industries	1,076	text	N/A	Yes	Entities include: PhysicalObject and subclasses based on inherent function, State, Process, Activity, Property Relations include: Mereological, Property, Type, Participatory
MaintIE (Coarse-Grained Large-Scale Corpus) ^h	heavy mobile equipment, rail, mineral processing and infrastructure industries	7,000	text	N/A	Yes	Entities include: PhysicalObject, State, Process, Activity, Property Relations include: Mereological, Property, Type, Participatory
Excavator Maintenance Work Orders ^h	Mobile Mining Equipment	5,486	OriginalShorttext	BSCStartDate, Asset (code), Cost	Yes	MajorSystem, Part, Action, Variant, FM, FunctionalLocation
Asset Management Parks System (AMPS) ⁱ	Park assets	1.97M	EVT_DESC	EVT_TYPE (text, desc: work order type), EVT_DATE, EVT_SAFETY (text, desc: Indicates if the work order is safety related)	Yes	EVT_REQM (desc: problem code), EVT_PRIORITY (desc: Assigned priority of the work order), EVT_STANDWORK (text, desc:work order standard type)
MaintNorm ^j	Industrial Maintenance (Heavy Mobile Equipment)	12,000	No field name	Normalised short text description in WNUT standard normalisation format	No	N/A

^aNYC OpenData, 2024, "Handyman Work Order (HWO) Charges", Department of Housing Preservation and Development (HPD)[Online]. Available: https://data.cityofnewyork.us/Housing-Development/Handyman-Work-Order-HWO-Charges/sbnd-xujn/about_data

^bUnited States Nuclear Regulatory Commission, 2024, "Initiating Events" [Online]. Available: <https://nrc.nrc.gov/InitEvent/>

^csvc dmesb, 2023, "WSSC Completed Service Alert Work Orders", dataMontgomery [Online]. Available: https://data.montgomerycountymd.gov/Community-Recreation/WSSC-Completed-Service-Alert-Work-Orders/4tjk-62h7/about_data

^dConnecticut Open Data, 2020, "Vehicle Historical Maintenance Dataset", MaintNet [Online]. Available: <https://people.rit.edu/fa3019/technical/vehicle.html>

^eMaintNet, 2020, "Aircraft Historical Maintenance Dataset", MaintNet [Online]. Available: <https://people.rit.edu/fa3019/technical/aircraft.html>

^fBaltimore City Maryland Data, 2020, "Facility Historical Maintenance Dataset", MaintNet [Online]. Available: <https://people.rit.edu/fa3019/technical/facility.html>

^gT. Bikaun, T. French, M. Stewart, W. Liu, and M. Hodkiewicz, "MaintIE", UWA NLP-TLP Group [Online]. Available: <https://github.com/nlp-tlp/maintie>

^hM. Hodkiewicz, 2024, "Excavator Maintenance Work Orders", UWA System Health Lab [Online]. Available: <https://prognosticsdl.systemhealthlab.com/dataset/excavator-maintenance-work-orders>

ⁱNYC OpenData, 2024, "Asset Management Parks System (AMPS) - Work Orders", Department of Parks and Recreation (DPR)[online]. Available: https://data.cityofnewyork.us/Environment/Asset-Management-Parks-System-AMPS-Work-Orders/8sdw-8vja/about_data

^jT. Bikaun, W. Liu, and M. Hodkiewicz, "MaintNorm", UWA NLP-TLP Group [Online]. Available: <https://github.com/nlp-tlp/maintnorm>