

ONTOHUMAN : ONTOLOGY-BASED INFORMATION EXTRACTION TOOLS WITH HUMAN-IN-THE-LOOP

Kobkaew Opasjumruskit

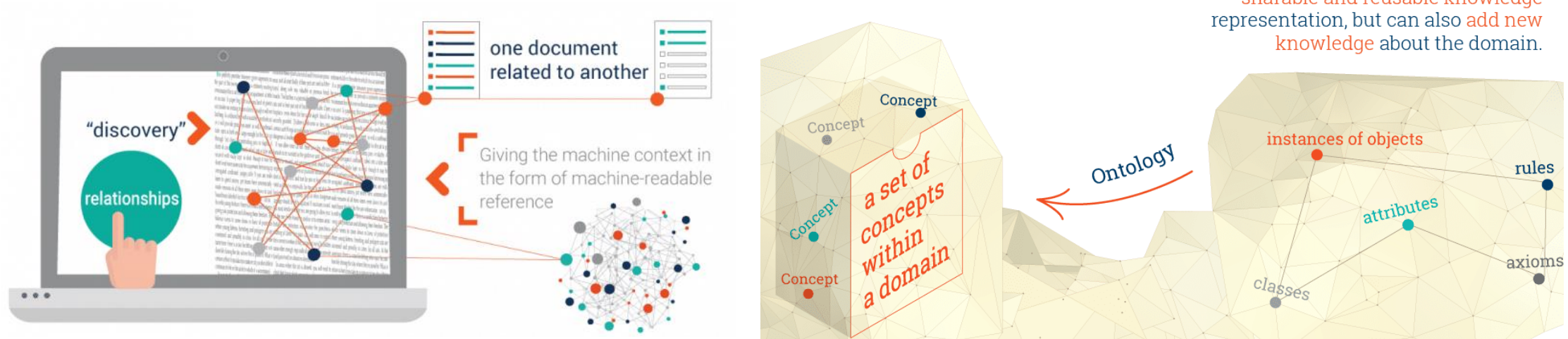
Information Extraction and Interoperability DLR Institute of Data Science

NFDI4Ing Conference 2022

Partially funded by Deutsche Forschungsgemeinschaft as part
of NFDI under project number 442146713.



Solution: Information Extraction x Ontology

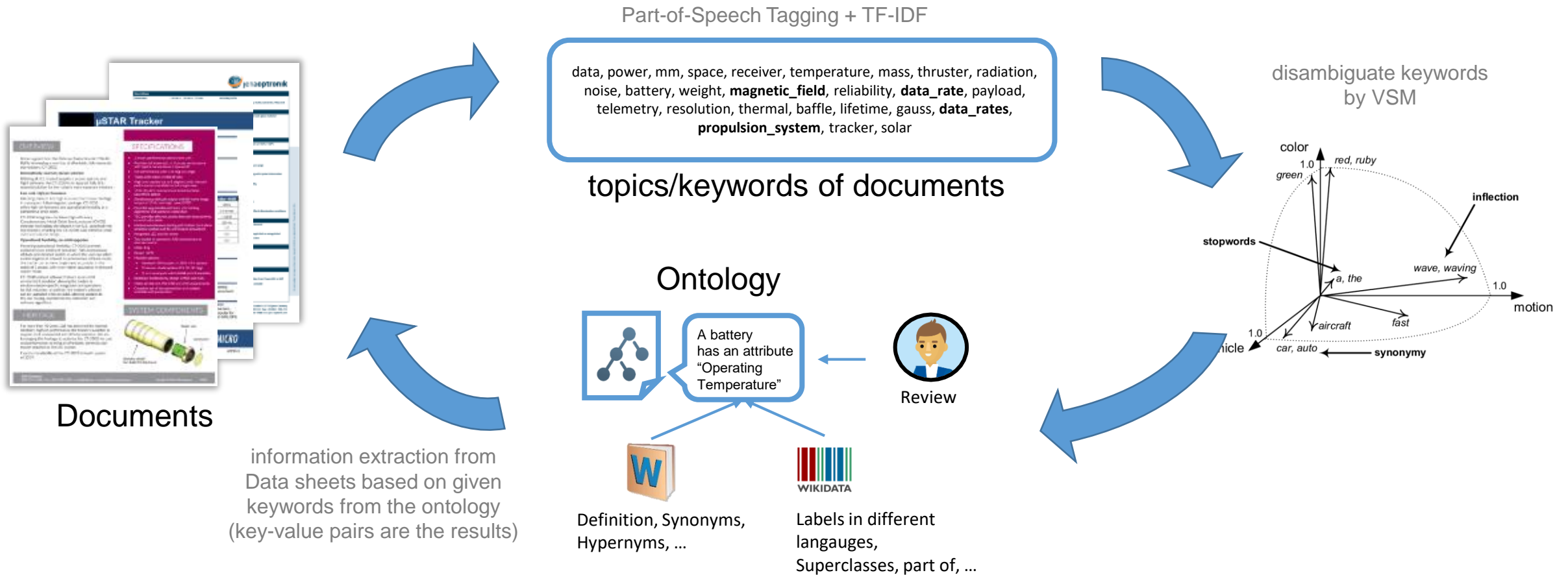


Ontologies do not only introduce a sharable and reusable knowledge representation, but can also add new knowledge about the domain.

<https://www.ontotext.com/knowledgehub/fundamentals/what-are-ontologies/>

- Natural Language Processing
- Semantic Knowledge
- Ontology

Information Extraction + Ontology + Human = OntoHuman



System Overview

DSAT

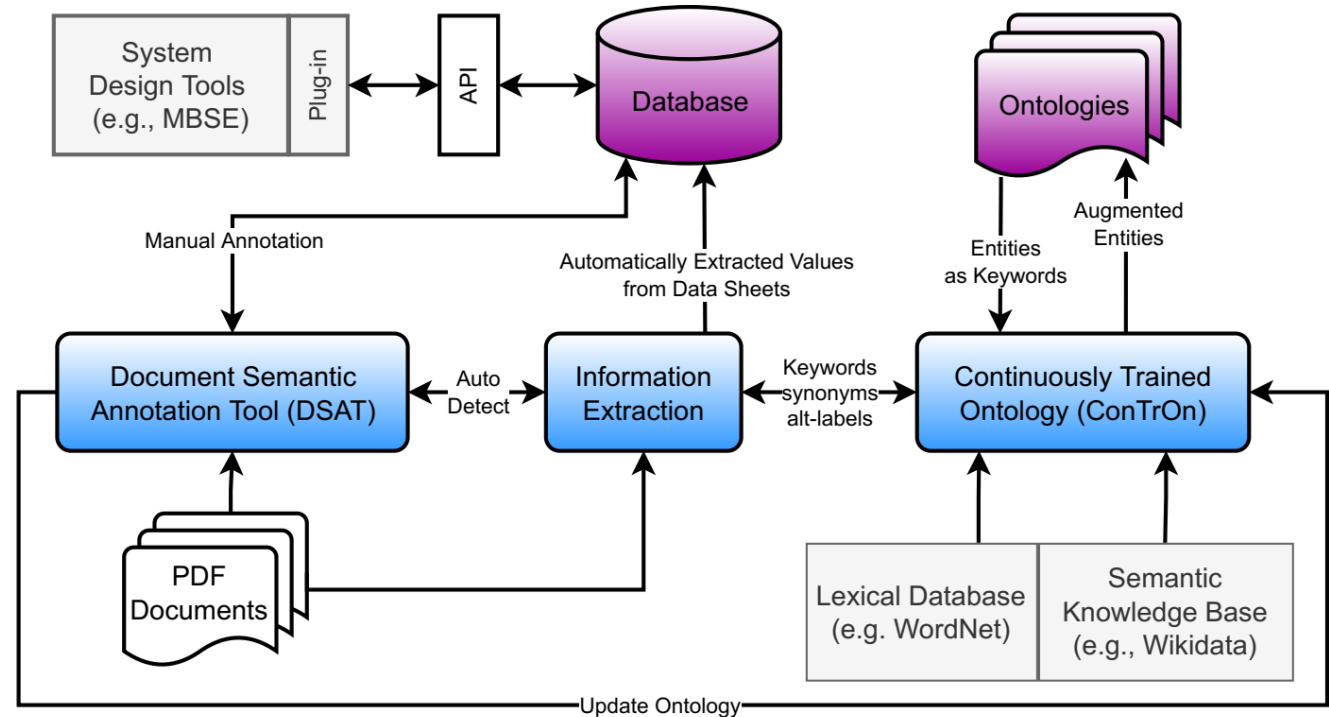
- Intuitive way for annotating data
- **Auto-Detect** feature
- Annotations revision & **feedback** collecting

Information Extraction (PLIX)

- Ontology-based **key-value** pairs extraction

ConTrOn

- **Enrich ontologies** based on
 - extracted information from data sheets
 - external semantic knowledge base



DSAT Screenshot

The screenshot displays the DSAT interface for a document titled "Eagle Picher: Battery-15 EP-SAR-10197-DATA-SHEET". The interface is divided into three main sections:

- Left Panel (II):** Contains a PDF view of the document. It features a blue callout box with the text: "Lithium-ion provides higher energy levels and longer cycle life, at a lower weight and smaller volume, with less maintenance than lead acid, Ni-Cd, or Ni-MH batteries." Below this is a "Specifications" table and a "Features & Benefits" list.
- Top Panel (I):** Shows the document title and a search bar.
- Right Panel (III):** Displays an "Auto-Detect" table with 1-2 items. The table lists various keys and their values, such as Weight (63.5 kg), Dimensions (10.6"W x 29.5"L x 10.4"H), and Operating Voltage (27.0V to 35.8V).

Specifications Table:

Part Number	SAB-10197
Weight Not to Exceed	140 lb/63.5 kg
Nominal Voltage	33.3V
Operating Voltage	27.0V to 35.8V
Beginning of Life Capacity/Energy	200Ah/6660WH @ 20°C
Maximum Dimensions	10.6"W x 29.5"L x 10.4"H
Specific Energy	104.9 Wh/kg
Operating Temperature	-5°C to 35°C
Survival Temperature (non-op)	-15°C to 40°C
Charge Current (Max)	100A
Maximum Continuous Discharge	154A
Maximum Discharge Pulse	400A for 1 sec
Random Vibe Levels	14 G RMS
Sine Vibe Levels	15G
Shock Level	1135G

Features & Benefits:

- + Autonomous cell bypass capability
- + Primary and redundant heaters
- + Back-up temperature and voltage telemetry
- + Built-in cell safety protection
- + Current sense capability
- + Operational for 15-year mission life (GEO)
- + Space flight heritage
- + High reliability ≥ 0.99
- + Built-in lifting points with detachable lifting fixture.
- + Designed for radiation total-dose exposure of $3.11E6$ rads
- + Connector savers for spacecraft integration and test bracket
- + Custom connectors are keyed and clocked per customer specification; MIL-DTL-38999 and/or NASA-S-311-P-768 connectors available upon request
- + Subjected to vibratory and thermal-vacuum testing

Applications:

- + Military communications and surveillance
- + Commercial communication and broadcasting
- + NASA research
- + Environmental monitoring
- + Global navigation and tracking

Auto-Detect Table:

Key	Value
Weight	63.5 kg
Dimensions	10.6"W x 29.5"L x 10.4"H
Weight	Not to Exceed 140 lb/63.5 kg
Shock	Level 1135G
Operating Temperature	-5°C to 35°C
reliability	> 0.99
Nominal Voltage	33.3V
Operating Voltage	27.0V to 35.8V
Survival Temperature	(non-op) -15°C to 40°C
Capacity	/Energy 200Ah/6660WH @ 20°C
Charge Current	(Max) 100A

- I. Select document
 - update metadata
 - update ontology
- II. PDF view
- III. Annotation list
 - edit item
 - copy
 - delete
 - feedback

DSAT : Ontology preview & update

The screenshot displays the DSAT ontology viewer interface. At the top, a title bar reads "Eagle Picher: Battery-15 EP-SAR-10197-DATA-SHEET". Below this, the main area is titled "View Ontology" and contains a network graph. The graph features a central node labeled "Class" with numerous other nodes connected to it, including "DischargePuls", "ChargeRate", "EnergyDensity", "DischargeRate", "EndOfDischarge", "DischargeCurrent", "EndOfChargeVoltage", "ContinuousDischargeCurrent", "Battery@capacity", "Literat", and "Battery capacity design for el.". A popup window is overlaid on the left side of the graph, displaying metadata for the selected node:

creator
license <http://creativecommons.org/licenses/by/3.0/>
subject E-Commerce, E-Business, GoodRelations, ...
title ACCOM: The Accommodation Ontology for ...
versionInfo V 1.0, Release 2020-05-05

Select predicates for graph:

- owl#imports
- rdf-schema#isDefinedBy
- rdf-schema#subClassOf
- rdf-schema#subPropertyOf

An "Apply" button is located at the bottom of the popup. Below the graph, there is a "Browse" button and a text input field containing "Or IRI, <http://purl.example.org/ex.owl> (no preview)". At the bottom right of the interface, there are "Reset", "Save", and "Cancel" buttons.

Auto-detect feedback

Eagle Picher: Battery-15 EP-SAR-10197-DATA-SHEET

Lithium-Ion provides higher energy levels and longer cycle life, at a lower weight and smaller volume, with less maintenance

Features & Benefits

- + Autonomous cell bypass capability
- + Primary and redundant heaters
- + Back-up temperature and voltage telemetry
- + Built-in cell safety protection
- + Current sense capability
- + Operational for 15-year mission life (GEO)
- + Space flight heritage
- + High reliability > 0.99
- + Built-in lifting points with detachable lifting fixture
- + Designed for radiation total-dose exposure of 3.11E6 rads
- + Connector savers for spacecraft integration and test bracket
- + Custom connectors are keyed and clocked per customer specification; MIL-DTL-38999 and/or NASA-5-311-P-768 connectors available upon request
- + Subjected to vibratory and thermal-vacuum testing

Auto-detected Annotation

nominal voltage 33.3V

Key Value Remark

was detected because
nominal voltage is a label of **supply voltage**

Correct
 Incorrect, because ...

- the label should belong to
non operating shock
- something else:
different reason

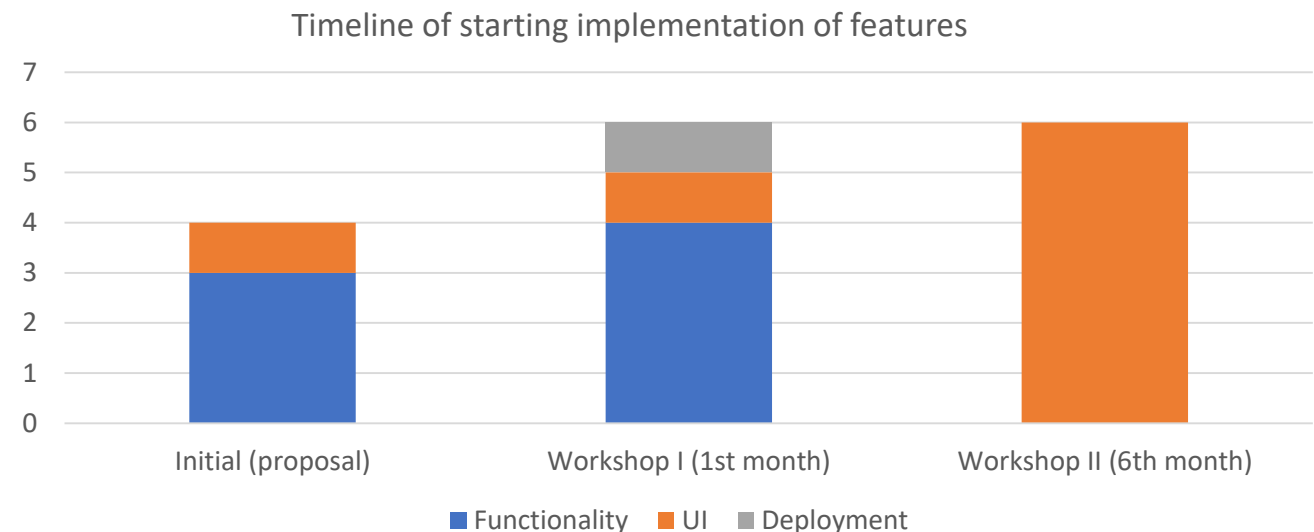
Related terms: nominal voltage, voltage, DC Supply Voltage, DC Voltage, Supply Voltage,

Auto-Detect 1-2

Key	Value
Weight	63.5 kg
Dimensions	10.6"W x 29.5"L x 10.4"H
Weight	Not to Exceed 140 lb/63.5 kg
Shock	Level 1135G
Operating Temperature	-5°C to 35°C
reliability	> 0.99
Nominal Voltage	33.3V
Operating Voltage	27.0V to 35.8V
Survival Temperature	(non-op) -15°C to 40°C
Capacity	/Energy 200Ah/6660WH @ 20°C
Charge Current	(Max) 100A


New functions implemented in OntoHuman

- Correction of results from automatic information extraction (via feedback UI).
- Export of annotated information.
- Upload other ontologies for annotating documents.
- Support spreadsheets
- Option for processing specific pages instead of the whole document.
- Preview and display metadata of ontologies.
- Using multiple ontologies.



Live-Demo

Eagle Picher: Battery-15 EP-SAR-10197-DATA-SHEET



Lithium-ion provides higher energy levels and longer cycle life, at a lower weight and smaller volume, with less maintenance than lead acid, Ni-Cd, or Ni-MH batteries.

Specifications

Part Number	SAR-10197
Weight Not to Exceed	140 lb/63.5 kg
Nominal Voltage	33.3V
Operating Voltage	27.0V to 35.8V
Beginning of Life Capacity/Energy	200Ah/6660WH @ 20°C
Maximum Dimensions	10.6"W x 29.5"L x 10.4"H
Specific Energy	104.9 Wh/Kg
Operating Temperature	-5°C to 35°C
Survival Temperature (non-op)	-15°C to 40°C
Charge Current (Max)	100A
Maximum Continuous Discharge	154A
Maximum Discharge Pulse	400A for 1 sec
Random Vibe Levels	14 GRMS
Sine Vibe Levels	15G
Shock Level	1135G

Applications

- + Military communications and surveillance
- + Commercial communication and broadcasting
- + NASA research
- + Environmental monitoring
- + Global navigation and tracking

Auto-Detect 1-2

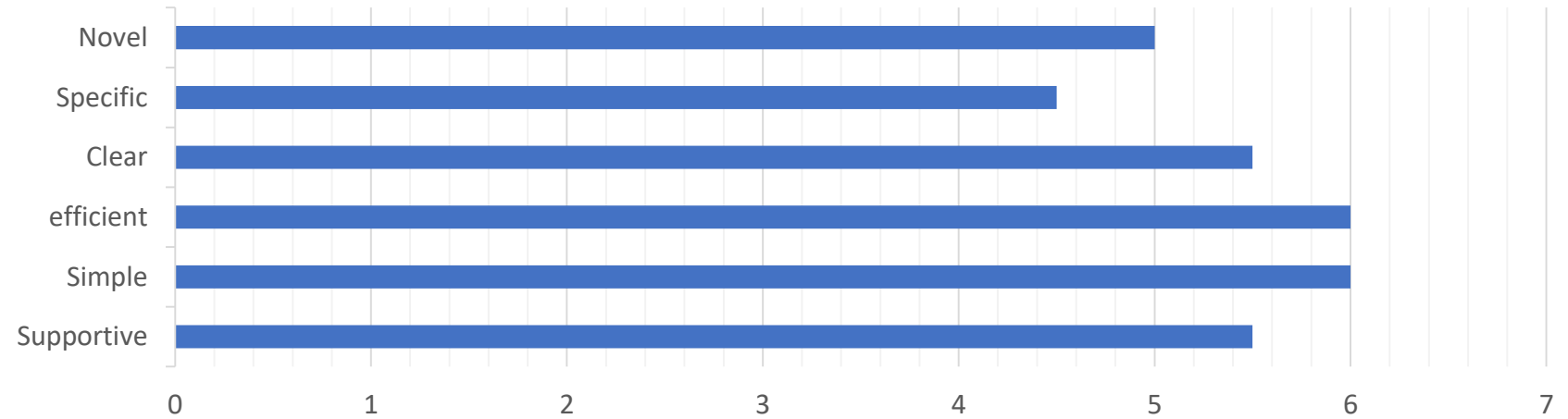
Key	Value
Weight	63.5 kg
Dimensions	10.6"W x 29.5"L x 10.4"H
Weight	Not to Exceed 140 lb/63.5 kg
Shock	Level 1135G
Operating Temperature	-5°C to 35°C
reliability	> 0.99
Nominal Voltage	33.3V
Operating Voltage	27.0V to 35.8V
Survival Temperature	(non-op) -15°C to 40°C
Capacity	/Energy 200Ah/6660WH @ 20°C
Charge Current	(Max) 100A

Features & Benefits

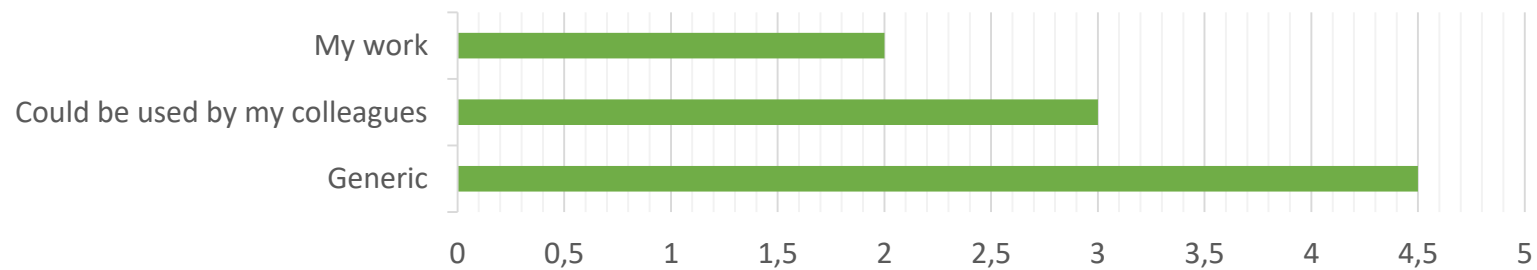
- + Autonomous cell bypass capability
- + Primary and redundant heaters
- + Back-up temperature and voltage telemetry
- + Built-in cell safety protection
- + Current sense capability
- + Operational for 15-year mission life (GEO)
- + Space flight heritage
- + High reliability > 0.99
- + Built-in lifting points with detachable lifting fixture
- + Designed for radiation total-dose exposure of 3.11E6 rads
- + Connector savers for spacecraft integration and test bracket
- + Custom connectors are keyed and clocked per customer specification; MIL-DTL-38999 and/or NASA-5-311-P-768 connectors available upon request
- + Subjected to vibratory and thermal-vacuum testing

DSAT : Users Survey

Tool rating



Domain of Usage

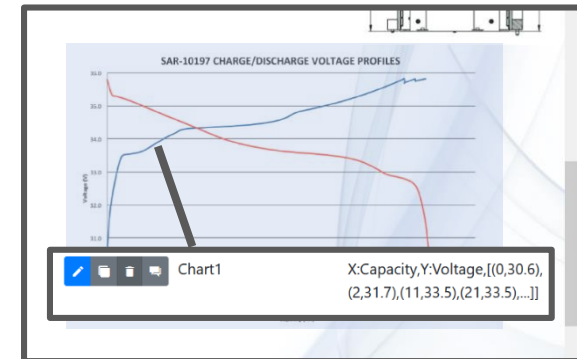


Publications & Deliverables

- 2 workshops (July 2021, February 2022) with participants from NFDI4Ing community
- 1 NFDI Tool talk
 - <https://nfdi4ing.de/tooltalk-dsat>
- CDVE (Cooperative Design, Visualization and Engineering) paper (published in process)
 - September 2022
- Sourcecode & installation packages
 - <https://zenodo.org/record/6783007>

Outlook

- Published tutorials & source code
 - <https://nfdi4ing.de/tooltalk-dsat>
 - <https://zenodo.org/record/6783007>
- Extend the ontology preview with editing mode
- Currently detecting numerical values, but will include text value later
- Extend the auto-detect results with graph digitizer
- Ontology supporting multiple-languages keys



Frequency band	400.15 - 406 MHz
Plage de fréquences	400,15 - 406 MHz
Frequenzband	400.15 - 406 MHz