SAWE Paper No. 3694
Category No. 12

USING KNOWLEDGE ANALYTICS TO SEARCH AND CHARACTERIZE MASS PROPERTIES AEROSPACE DATA

Jeffrey A. Cerro - NASA Langley Research Center, Hampton VA
Theodore D. Sidehamer - Craig Technologies, Cape Canaveral FL
Dorothy L. Notarnicola - NCI Incorporated, Reston VA

For Presentation at the
77th SAWE International Conference
on Mass Properties Engineering
May 6-10, 2018

Fort Worth, Texas

Using Knowledge Analytics to Search and Characterize Mass Properties Aerospace Data

Jeffrey A. Cerro - NASA Langley Research Center, Hampton VA
Theodore D. Sidehamer - Craig Technologies, Cape Canaveral FL
Dorothy L. Notarnicola - NCI Incorporated, Reston VA

Abstract

There is growing capability in the field of “Big Data” and “Data Analytics” which Mass Properties Engineers might like to take advantage of. This paper utilizes an implementation of the IBM Knowledge Analytics and Watson search capabilities to explore a corpus of material developed primarily with the interests of Mass Properties Engineers and vehicle concept developers at its forefront. The full collection of SAWE Technical Papers from 1939 thru 2015 is a major portion of the knowledge content. Additional aerospace vehicle design information includes metadata from AIAA (American Institute for Aeronautics and Astronautics), and INCOSE (International Council on Systems Engineering) as well as author provided personal search material. This data is processed with respect to certain expected content, data taxonomies and key words to become the core data in NASA Langley Research Center’s “Vehicle Analysis Analytics”, IBM Watson Content. Processed data becomes the corpus of information which is interrogated to provide examples of finding data for mass regression analysis, technology impacts on MPE, mass properties control, standards, and other aspects of interest.

Acronyms

AIAA	American Institute of Aeronautics and Astronautics
BoE	Basis of Estimate
COTS	Commercial Off The Shelf
CSV	Comma Separated Variable
DOI	Digital Object Identifier
FBS	Functional Breakdown Structure
IBM	International Business Machines
ICA	Ibm Content Analytics
INCOSE	International Council on Systems Engineering
LaRC	Langley Research Center
LaSSe	Launch and Space Systems e-library
MER	Mass Estimating Relationship
MPE	Mass Properties Engineering
NASA	National Aeronautics and Space Administration
NLP	Natural Language Processing
NTRS	NASA Technical Report Server
OCIO	Office of the Chief Information Officer
OCR	Optical Character Recognition
RIS	Research Information System
SAWE	Society of Allied Weight Engineers
SME	Subject Matter Expert
STI	Scientific Technical Information
WBS	Weight Breakdown Structure
WCA	Watson Content Analytics
WEX	Watson EXplorer

Introduction

A large portion of the capabilities which define good mass properties engineering involves having a knowledge of the past and using that knowledge to predict the future. Success in the bid phase of competitive vehicle development is enhanced by being able to provide viable mass properties estimates quickly and with reasonable certainty to the proposal team. Success in the vehicle development phase by defining accurate mass targets and correctly characterizing driving vehicle synthesis requirements means early vehicle mass estimations will be more stable through a programs development lifecycle and variations from expected growth will be minimized. In the parlance of “Big Data” this general capability is characterized as performing accurate “predictive analytics”. Due to the early and constant need for accurate mass properties data in vehicle development and use, Mass Properties Engineers are quite regularly engaging in predictive analytics. Whether it’s a corporation, government organization, or individual, the entity that has the best capability in utilizing predictive analytics for mass properties control stands the best chance of continued program success. As the engineering world becomes more digitized in terms of information availability and process execution, having the capability to develop accurate predictive analytics is now enhanced by perusal and use of the growing number of digitized information sources. We may characterize the general repositories of such data as “Big Data”, and the efficient mining and processing of Big Data provides mass properties predictive capabilities. Airframe companies have great historical sources of information which are becoming increasingly digitized. The NASA Langley Research Center has advanced along this digital trajectory as well and this paper provides an example of how useful it can be to utilize emerging capabilities in the area of big data interrogation with regards to Mass Properties Engineering (MPE).

Big data may be either structured or unstructured. Structured data is well categorized, and tabulated as one thinks of how information in a spreadsheet or database is maintained. Unstructured data are the more humanly readable but not necessarily rigorously machine understandable types of data. Primarily this is in the form of papers, journals, presentations, and web information. In this paper aerospace vehicle design related unstructured data in the form of papers, magazine articles and other human readable text comprise a corpus of information which was processed by the IBM Watson Content Analytics computer program. A corpus of information to interrogate was built up from the author’s private hardcopy data repository as well as leased access to full content of all SAWE technical papers, download of INCOSE metadata with abstract information, and similarly downloaded AIAA metadata. Example topical data querying and establishing of data for predictive analytics is shown. Quality Mass Properties data for success in future projects stems from having a solid historical basis-of-estimate using organized, searchable historical data. The capabilities provided by Big Data search techniques are shown to be useful in creating knowledge from an initially unorganized set of multi-sourced textual information.

General Watson Explorer Content Analytics Capabilities

This paper utilized “Big Data” interrogation software from the IBM corporation, specifically IBM’s Watson Explorer Content Analytics [1], version 11.0.2. Note that the IBM terminology for this data search tool has evolved thru IBM’s development process from being termed Ibm Content Analytics, (ICA), to Watson Content Analytics (WCA) to Watson Explorer (WEX). WEX is the state at which the queries in this paper were performed. Full explanation to utilizing the capabilities of Watson Explorer Content Analytics is provided in concise detail via video training [2]. These capabilities are best summarized in the data views provided by the software’s user interface, but before we can present the data views the user must first be familiar with some content analytics terminology:

Correlation: This calculated value is a measure of how strongly facet values are related to the set of documents determined by the current query. If for example a value is associated more with one particular author than another it will have higher correlation to that first author, despite possibly less frequency of use. It is typically used to build a search query, by drilling into a query based on correlation values. Only documents found that resulted from high correlation of facet value to the search context need be considered for examination. Using correlations, a user can also take a "guided" tour of the data without knowing a lot about it.

Dictionaries: A dictionary is a SME defined set of synonyms for a particular term. In that manner a search can include all items of the dictionary term without specifying common other nomenclature for the same item:

Ex: engine may have dictionary synonyms; propulsion, power, motor

Facet: A facet is a categorization of information, such as by date, author, or by user defined facets such as a subsystem type. Facets may be hierarchically defined, as in a mass properties Functional Breakdown Structure. In this corpus a hierarchical facet structure is defined based on the FBS presented in [3]. Facets are made up of values and allow a user to interact with the data by drilling down into it. Some available facet categories are automatic, NLP (Nouns, verbs, phrase, etc.) items and metadata are of this type. Drilling into facets shows not just selected terms but other terms found of the same category.

Frequency: a count of the number of documents that contain a value.

Parsing Rules: Adding context by looking for specific SME defined terms, phrases and patterns across sentences, paragraphs and documents.

Parsing rules can be set such that facet terms may be required to be contained within a smaller context than a full document. For example, the terms mass and the term figure can be restricted in a parsing rule to be found within the context of a paragraph, or a sentence. Mass with Figure in the context of a paragraph returns 153 documents, in our corpus, whereas in the context of a sentence returns 110. Note that mass and figure together in the context of a whole document returns 1047 documents.

Taxonomy: A taxonomy is a hierarchical categorization of terms defined by the SME. Like a dictionary both a taxonomy and a dictionary should be defined prior to the Watson crawling process which builds the processed searchable corpus. MPE has always required well defined taxonomies in terms of Functional Breakdown Structure, and Work Breakdown Structure (FBS, WBS) to enable comparisons of MPE data across historical, product type, and derived product variations. This project has a focused interest in mass properties for Missile and Space Systems so a taxonomy was defined based upon the FBS provided in [3].

Query building:

To systematically review a large amount of data in search of particular aspects of interest requires querying the data. Watson Analytics provides a hierarchical query state that is defined by user interaction in two principal manners. First is by knowing the rules for creating queries [4] and simply keying in search criteria. Searches generate queries which can be interactively or built manually by the user. Standard querying techniques are available like ~, +, - and * for wild card searches. Other features include type ahead, boosting, similarity, keyword, facet name and value searches.

The following sample query structures are used to investigate the Vehicle Analytics corpus:

Free style: Ex: multiple words, [vehicle mass cg] returns all documents with all three terms contained in them. Includes word synonyms if they are defined. Phrases contained in "__ ", are exact phrase searches.

mass:	frequency = 1340 documents
vehicle mass:	frequency = 849 documents
vehicle mass cg:	frequency = 245 documents
satellite mass:	frequency = 240 documents
moment of inertia	frequency = 460 documents
"moment of inertia"	frequency = 270 documents

Fuzzy searches allow users to search for a phrase like IBM analytics~0.5. Such a search will return document that contain IBM and analytics, IBM and analyze, IBM and analysis and so on.

=prefix for only that exact term to appear (not case sensitive)
~prefix, includes word synonyms
-prefix, to exclude an item from a match

=mass:	frequency = 1330 documents
~mass:	frequency = 1340 documents
-mass:	frequency = 8975 documents

The second means for building queries and thus defining the current query state is via dynamic interactions provided by the user interface within the various data views. If for example a time series view of the term "product of inertia" is set for the initial query state, 121 documents are returned, Fig 1. Two years stick out as being more active in terms of documents related to

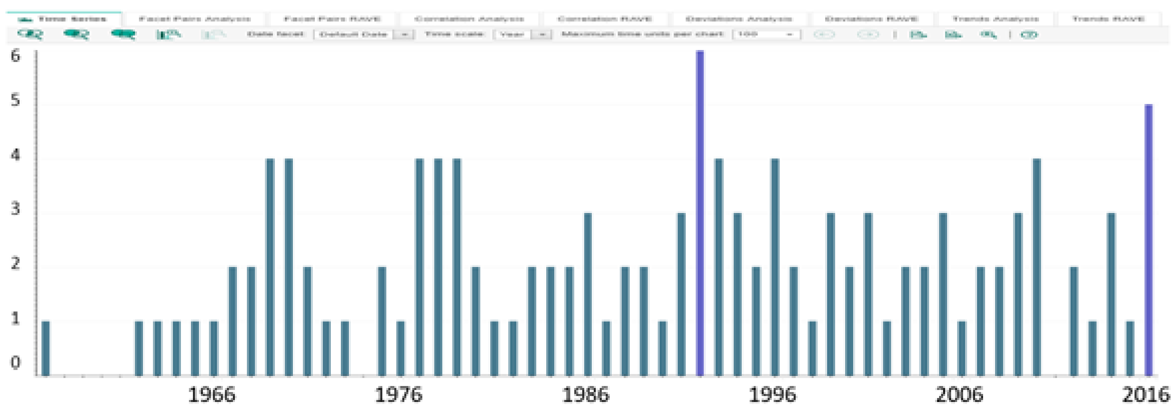


Fig. 1 - Time series view of "product of inertia" with selected time periods used to refine the query state

product of inertia. Interactively from this view the years 1992 and 2016 are selected and added to the query state with a Boolean "AND" selection. The additional query tree items reduce the number of applicable documents down to 11. The documents view is then a likely view to select for detailed investigation into publications of those two years. The query tree created by this selection process is shown in Fig. 2. The Advanced Search tab of the user interface shown in Fig 2. is also a means for users to modify an existing query tree with prompted text fields. Note also that all search states can be named, stored and recalled as needed.

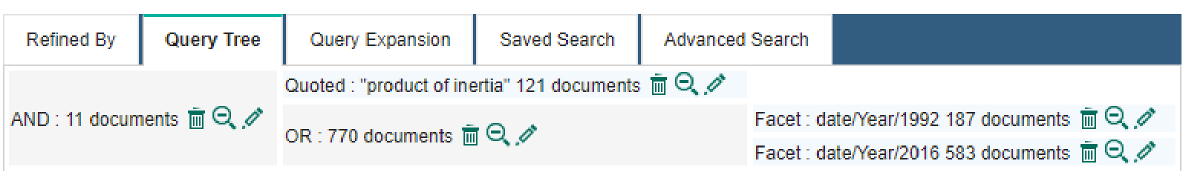


Fig. 2 - Interactively created query tree, resulting from drilling down with additional search criteria

With an understanding of analytics terminology some of the most utilized means to display search results via data views can be explained:

Facets view: Shows the frequency of facet values in the corpus

Time series view: A bar chart view of the variation in values over a selected time period.

Deviations view: To examine a variation in how expected behavior over time may be deviated from, for example if an author publishes a paper every year you may see a deviation from that which could require investigation into why such over or under productivity may have occurred.

Documents view: This view provides a listing of all data items which pass the current search criteria. It is from this view that items with full text content, such as SAWE technical papers in the LaRC Vehicle Analysis Analytics Watson project are instantly available to the investigator. If full text content is not available, metadata and additional linkages may direct the investigator to where full text information on a selected document may be found. A filter on the documents view also permits screening of information without modifying the full content query state.

Facet Pairs view: To compare the frequency and correlation of one facet with respect to another the facet pairs capability is used. This is a powerful query approach for understanding interrelationships in data. A “row” selected facet is compared against a “column” selected facet in a two dimensional representation of the items intersecting features. Frequency and correlation values of intersecting row/column items are shown in the intersected cell.

Rows: Author	Columns: Keywords	Frequency	Correlation
University, California State	student papers	13	309.6
Tellet, David	13. weight engineering - marine	15	66.4
MacConochie, Ian O.	18. weight engineering - spacecraft design	11	35.4
Boynton, Richard	03. center of gravity	14	30.6
Boynton, Richard	06. inertia measurements	13	21.7
McCarty, J R	01. aircraft loading - general	7	19.8

Fig. 3 - Facet pairs view

In Fig. 3 we join author to keywords and sort based on correlation. This images show how strong a relationship there is between an author and a particular keyword.

Connections view: Shows relation linkages between values in two facets. Links between values are color coded by degree of correlation. Here again we join author to keywords and show the connections. An important note is the connection view also “joins” author to author and keyword to keyword. This images show how strong a relationship there is between an author and a particular keyword. Note that two of the authors, Mark Lewis and Ryan Starkey are highly correlated. Likely they are collaborating researchers.

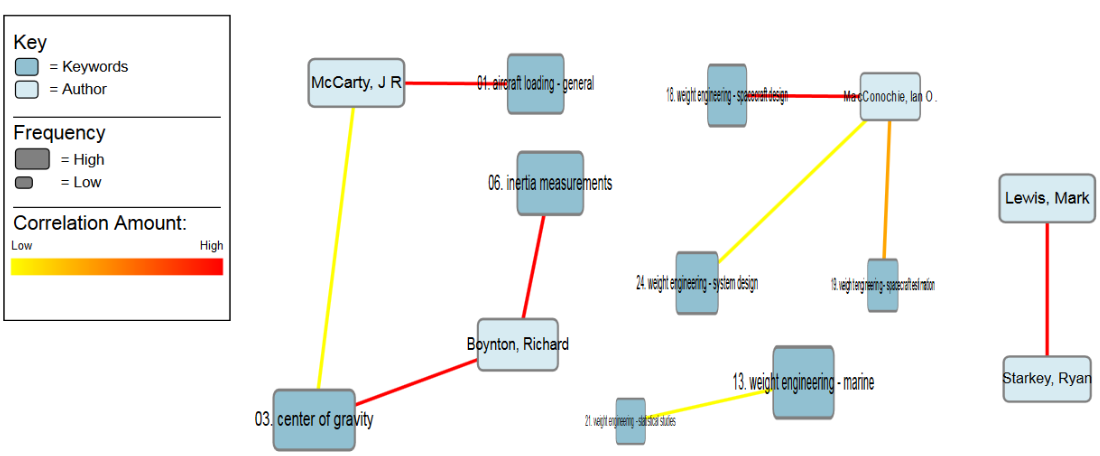


Fig. 4 Connections

Fig. 5 - Rave View Keyword Facet Joined to Author Facet

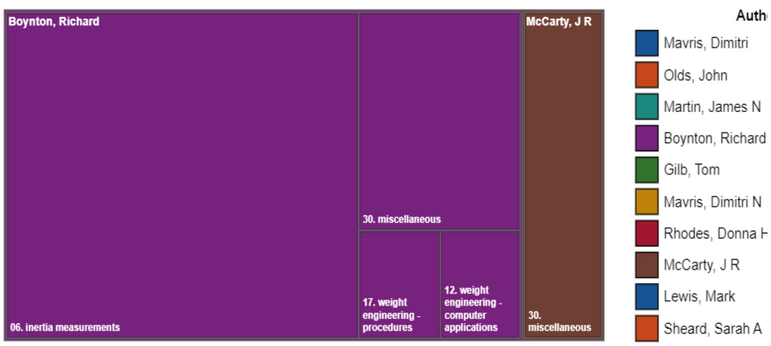


Fig. 6 - Rave View Author facet joined to keyword facet

Facet Pairs RAVE view: Facet Pairs Rapid Adaptive Visualization Engine (RAVE) Chart. A RAVE chart joins two facets to each other based on frequency and correlation and can be visualized as a heat map. In Fig. 5 we see the Keyword Facet joined with the Authors facet, showing the area of interest in MPE that authors are most associated with.

We can reverse the order, Fig. 6, and join the Authors facet to the Keyword facet, showing authors and their associated keywords. The size of the boxes is based on correlation and these authors are likely subject matter experts for particular MPE topics.

Dashboard view: An administrator selection of windows may be created for a Watson project to have a default set of information easily accessible to data investigators. This project did not define such views.

Reports view: Appropriate text reports of currently displayed information are available from several of the views by selecting a reports icon. Tabular CSV files with frequency and correlation values are typical report outputs.

Further data interrogations in context with the subject NASA compiled Vehicle Analysis Analytics corpus is provided after an explanation of the corpus raw data sources and its processing by required preparatory steps to formulate rapidly searchable WEX data.

Corpus Content

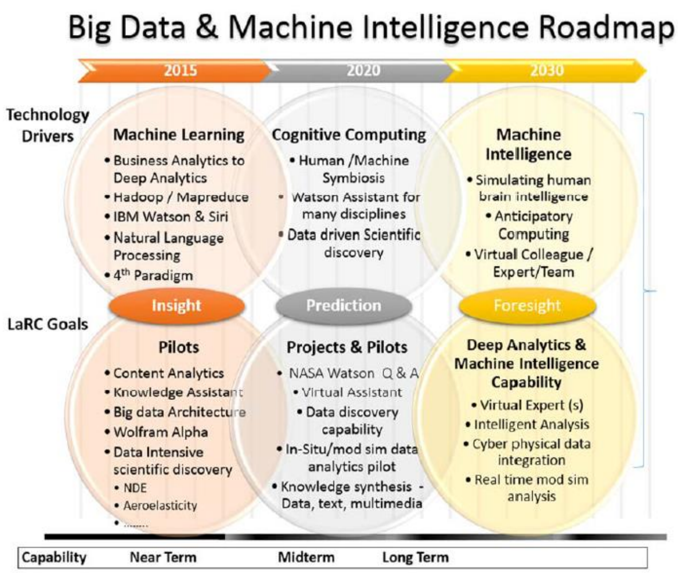


Fig. 7 - Big Data and Machine Intelligence Roadmap [5]

The NASA Langley Research Center's OCIO, and associated personnel involved in LaRC digital transformation activities have implemented a concerted effort into investigating the capabilities which "Big Data & Machine Intelligence" may bring to projects and discipline expertise at the center. As part of that initiative IBM Watson Content Analytics / Watson Explorer (WCA/WEX) was instantiated from several center areas of interest Fig. 7. Instantiated projects involved in the LaRC WEX initiative include aspects of autonomous flight, nano materials, vehicle and mission design, space radiation, uncertainty, human machine teaming, and model based approaches. For this paper a database of information, a WEX Corpus, specific to aerospace vehicle design with emphasis in mass properties characterization was created. This project was termed the "Vehicle Analysis Analytics" and its intent is described in [5].

> *Working with SMEs in the Systems Analysis and Concepts Directorate (SACD), more than 20,000 reports from databases including AIAA, NTRS, SAWE, and hundreds of scanned historical documents are being ingested into the WCA tool. This will give SMEs the power to more easily find mass and weight properties used for various vehicle concepts and designs, while leveraging the corpus of research and design reports from the past fifty years. The SMEs envision a resource that will improve their ability to accelerate the design process, predict 'what ifs' of future vehicle designs, and better leverage the work they have done over the past few decades. To facilitate SME goals for the collection, a large taxonomy based on SME-provided technical content was built to facilitate efficient faceted search.*

One of the major collections for this project was obtained through leased access to the complete holdings of SAWE papers from 1939 through 2015. The LaRC OCIO worked with the Executive Board at SAWE to obtain a Partnership that entitled NASA Langley Research Center to have two years of access to the Paper Database as an SAWE Corporate Partner. SAWE technical paper abstracts are publically downloadable in BibTeX, RTF, XML format. Along with ingesting such publically available metadata the lease agreement meant that the WEX corpus could also include full text searchable documents for each of the SAWE papers. Additional searchable full text document content was provided for 11 boxes, 432 papers, which were OCR scanned in by the OCIO from the author's personal collection. These scanned documents were loaded into a reference management system, Mendeley [6] and exported for WEX ingestion. By using Mendeley all of the PDFs are gathered into one folder, then an exported XML file with links to PDF files is provided along with file metadata.

Using Mendeley to store and export all records provides an end result of similarly formatted records. This ensures uniform metadata fields to make ingestion, searching, and discovery of documents as consistent and equal as possible in WEX. Once in PDF form Mendeley, Fig. 8, uses a three step process to extract details. First, the client looks for embedded metadata. However, many PDFs don't contain this information, particularly older ones which are scans of print copies. The next thing Mendeley does is to look within the text of the document to see if it can find an identifier string, such as a DOI. If it finds one, it looks up the metadata in the appropriate database (Ex: Crossref for DOI's [7]). There are still a significant fraction of papers (maybe as many as 20% of papers uploaded to Mendeley) that don't have DOI's, For this, Mendeley uses some machine learning algorithms to identify the title, author, and publication year and uses that in a Google Scholar query to try to retrieve the information. In the future, Mendeley will be able to use its own records about which metadata is associated with which files to automatically fill in these details. Even after those steps have been taken, there will still be papers without proper metadata. That's when the manual process begins. In Mendeley, you can view the PDF and the metadata at the same time. Then you are able to delete, edit or add to the metadata fields in order to include as much information as necessary.

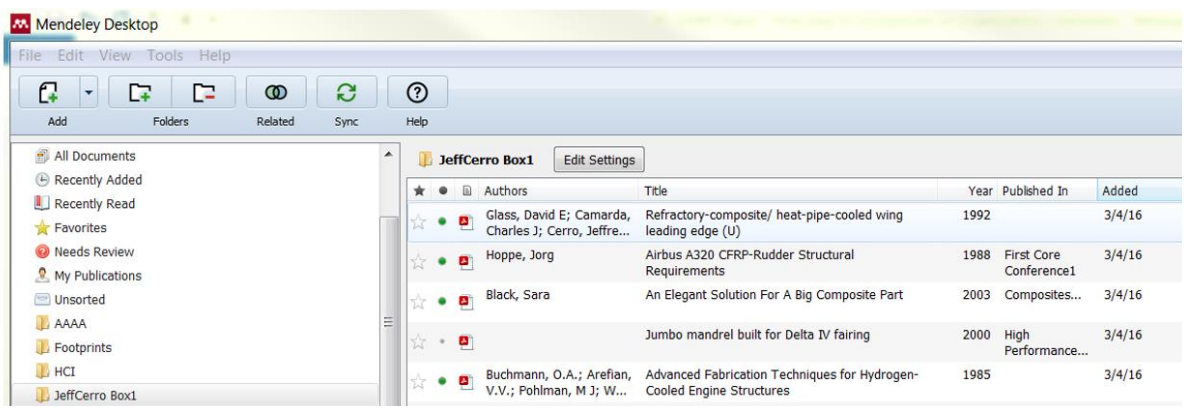


Fig. 8 - Mendeley Graphical User Interface

For OCR' d documents, after software attempts to extract metadata, a manual clean-up of information is required to assure correct metadata extraction. This required labor is on the order of one to two hours for each of the first three boxes of data. Eight of the boxes were processed thru a high-speed scanner with associated software that performed OCR and PDF creation. The scanning operator creates the metadata by manually filling in fields in a spreadsheet.

Through this corpus building period it was noted that most online database vendors offer record downloads in RIS or BibTeX format. In order to best load the metadata into WEX, XML was determined to be most effective. RIS documents were imported into Mendeley and then exported as XML to be loaded into the corpus. Metadata was also taken from the open AIAA and INCOSE organization technical paper access areas, processed via Mendeley and added to the

Vehicle Analysis Analytics WEX corpus. Note that WEX installations may be charged by the volume of material in a corpus, so a corpus of only metadata would be less costly to maintain than a corpus of full text material. The utility of having full text documents access within the WEX dashboard is very desirable but it's value may have to be traded off if a large corpus of source material is desired over a focus on detailed document content. A large corpus of metadata with smaller corpus sets in focused areas of interest may be a means to control user costs.

There are two more very notable pertinent sources of NASA information which ideally would be part of a corpus for Vehicle Analysis Analytics. First is the general public publishing of NASA papers available from the NASA Technical Report Server NTRS [8]. There are currently 1,002,666 documents listed in NTRS. Full text incorporation is beyond the scope of LaRC's WEX installation, though a separate corpus of NTRS information was created to investigate how such material could be included in future search projects. It is planned that metadata with abstracts from the NTRS database will be included in the future pending working with NASA STI to incorporate a large XML export. The second source of un-accessed information is a set of approximately 70,000 documents maintained at the NASA Marshall Space Flight Center. This is called the Launch and Space Systems e-library, LaSSe [9]. The material in this database would be very pertinent to Vehicle Analysis Analytics, but as it was noted to be on the order of 75% restricted distribution it was not considered for the current WEX corpus build. Note in general though that non-government personnel working with NASA may obtain LaSSe access if requested access permission is granted.

Having "clean" and complete metadata is extremely important when building the corpus of one's research. When the metadata is consistent across the documents, then each item carries similar weight when performing data mining. This is as opposed to when there is missing and inaccurate metadata such as journal, date, and author information. Also it should be noted that the database must be understood and sometimes interrogated in multiple ways by a user to assure his final citations are as expected. For example, an SAWE technical paper from 1932 was discovered in a query. As SAWE papers only data back to 1940 the paper in question was found on further investigation to be from 1982. This error being the result of OCR error from author provided hardcopy material. The question of duplicate document management could also be an issue as corpus material loaded over time from multiple sources are used to grow a database. However, WEX eliminates this problem by providing a flag such that duplicate documents may be treated as one when performing data interrogation. Having a robust corpus of documents is key to providing useful knowledge. During this corpus building process, we have learned that having detailed metadata with a well-written abstract is near as valuable as having a full-text

Analytics Facet | Metadata Facet | **Fields** | Categories | Rules

Name	Value
secondarytitle	34th AIAA Fluid Dynamics Conference and Exhibit
modifieddate	Jan 1, 2004, 12:00:00 AM
date	Jan 1, 2004, 12:00:00 AM
sourcetype	Metadata/XML
documenttype	VAB/AIAA
author3	Zoby, Ernest
author2	Thompson, Richard
author1	Wurster, Kathryn
title	Aeroheating Design Issues for Reusable Launch Vehicles -- A Perspective

Fig. 9 - Sample Metadata returned, for use in finding full text documents from an outside vendors database

document, especially of course if the WEX user has subscribed access to the results of queries where only metadata is returned. This was the case for AIAA and INCOSE information in the corpus. In that case the metadata provided by a search easily lets the user find the information from the associated organizations database, Fig. 9.

A summary of the final Vehicle Analysis Analytics corpus content is as follows, prior to loading of the additional OCR' d data:

- 10,170 characterized by "Content Type" facet
    - 7867 Metadata/XML metadata records.
    - 1969 Metadata/XML with Full Text PDF – The ingest process stripped a pdf of its text and "joined" with its corresponding metadata record.
    - 334 – Full Text only. (No Metadata)
- 9836 characterized by "Content Source/Collections" facet
    - 4420 AIAA
    - 3435 INCOSE
    - 1981 SAWE Technical Papers – Full Text

It is interesting to note that there are actually 10177 documents in the corpus. The above two tallies are characterized by two different defined facets. There remained 7 documents that did not get associated with a particular source format, or a particular supplier source.

Corpus Organization, and Preparation

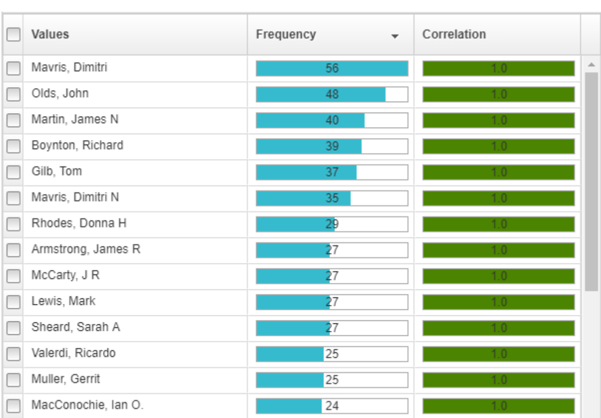
To prepare the mix of datafiles for incorporation into a WEX project several steps occur. Close interaction between the project SME and the WEX project developer on the information technology side of the task is required. For efficient searching numerous project facets need to be defined. The WEX developer will get keyword facets from available document metadata if so defined for a document. Also NLP facets and other document metadata type facets can be included in the projects facet list as determined by the WEX project developer. Typically, available metadata includes items like document title, author, and date. To focus the data interrogations on features of interest to the SME, the SME defines domain specific dictionary terms and taxonomies. Parsing rules are also defined prior to data ingestion to speed data retrieval in some general areas of a known generic search goal. Multiple taxonomies may also be defined. In our case vehicle performance based, vehicle design, and discipline taxonomies were created to capture traditional engineering tasks in Structures, Aerodynamics, Aerothermodynamics, Trajectory and Mass. Of primary concern and utilization in this paper is the use of the Mil 38310B [3] Functional Breakdown Structure for missile and space vehicles. This taxonomy is summarized in Appendix A – FBS Taxonomy based on Mil 38310B.

With the data files created in an ingestible format, and with the metadata, dictionary, parsing rules, and taxonomies set, the data can be intelligently incorporated into a WEX corpus via a document crawling process. A document can be just metadata, or metadata and the full text. Metadata joined with a full text searchable document is called a logical document. A document can also be just the searchable document itself without any metadata. Though as noted earlier this precludes the document from being included in any search which depends on using a metadata value. There are several ways to crawl the data [10] which we will not delve into detail on in this paper, but rather just note that this crawling process takes on the order of 30 minutes for a corpus of approximately 2000 full text and 8000 metadata only documents. Recrawling is required when new rules, facets, taxonomies and the like are defined in WEX. When adding additional data it may be done by crawling only the new data content or by a crawling of the old and new data together. As time was not overburdened in crawling of our Vehicle Analytics, all data was combined for any crawling instance. Large facet structures also significantly impact data crawl time.

Using Search Capabilities to Build MPE Knowledge

FBS Level I		Source				
		SAWE	AIAA	INCOSE	Misc.	All
1	Aerodynamic Surfaces	1417	329	140	73	1959
2	Body Structure	1628	546	358	78	2610
3	Induced Environment Protection	1206	84	577	66	1933
4	Launch, Recovery and Dockng	1556	423	782	77	2838
5	Main Propulsion	1604	555	1147	78	3384
6	Orientation Controls, Separation and Ullage	1677	848	2450	78	5053
7	Prime Power Source	1530	345	465	75	2415
8	Power Conversion and Distribution	1581	389	1048	75	3093
9	Guidance and Navigation	1383	66	772	62	2283
10	Instrumentation	1641	358	1056	78	3133
11	Communication	1615	712	3103	78	5508
12	Environmental Control	1187	57	581	62	1887
13	Armament	1367	55	625	62	2109
14	Personnel Provisions	1262	71	684	57	2074
15	Crew Station Controls and Panels	1416	377	1139	69	3001
16	Range Safety and Abort	1529	394	969	72	2964
17	Personnel	1185	97	846	60	2188
18	Cargo	697	72	85	34	888
19	Ordinance	888	1022	91	57	2058
20	Ballast	171	-	-	1	172
21	Residual Propellant and Service Items	1350	370	159	67	1946
22	Reserve Propellant and Service Items	-	-	-	-	0
23	Inflight Losses (Liftoff to Separation)	224	7	3	17	251
24	Thrust Decay Propellant	1220	161	128	59	1568
25	Full Thrust Propellant	918	84	45	58	1115
26	Thrust Propellant Buildup	736	49	20	41	846
27	Pre-Ignition Losses	756	49	20	41	866

Fig. 10 Mil38310B Taxonomy Characteristics

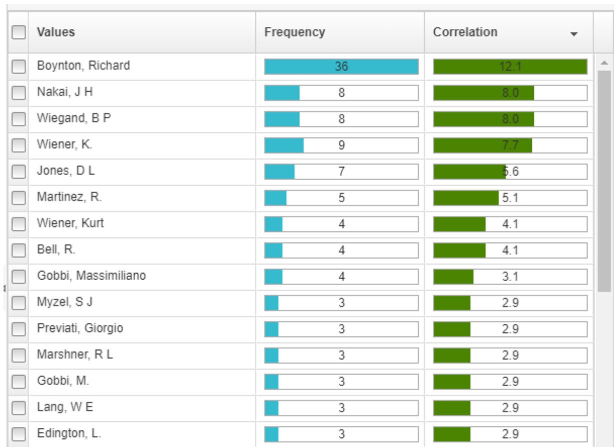


Values	Frequency	Correlation
Mavris, Dimitri	56	1.0
Olds, John	48	1.0
Martin, James N	40	1.0
Boynton, Richard	39	1.0
Gilb, Tom	37	1.0
Mavris, Dimitri N	35	1.0
Rhodes, Donna H	29	1.0
Armstrong, James R	27	1.0
McCarty, J R	27	1.0
Lewis, Mark	27	1.0
Sheard, Sarah A	27	1.0
Valerdi, Ricardo	25	1.0
Muller, Gerrit	25	1.0
MacConochie, Ian O.	24	1.0

Fig. 11 Frequency count on values for the "Author" facet

Using WEX query procedures a chart which provides insight into the source contributions of documents likely associated with the Appendix A FBS terminologies can be created. This count of document content over the four main sources of information and broken down by Mil 381310B FBS level 1 is given in Fig. 10.

Simple search strings are used to provide general knowledge of the type of MPE information which comprises the Vehicle Analysis Analytics corpus. Fig. 11 shows a simple single facet view frequency count by author regarding the full corpus set of 10177 reports. In this view there is no search query in place and so all values of the author facet being shown are given a correlation value of 1. The list is sorted by frequency where the top author Mavris has 56 documents. In Fig. 12 we place the inertia term in the query string and show that now there is a high degree of correlation for author Boynton with 36 documents relevant to the term inertia. In all, not shown in this figure is that the documents with relevancy to the term inertia reduces the full set of 10305 down to 546. At this point the search is defined but the continued search context is not yet restricted to the 546 documents. That can be done by selecting "all" author values and adding that restriction with a Boolean AND selection to the search context. If only the documents from Boynton were of interest, then only the value for that author could be selected and added to the general search context with the AND operation. A limited set of 36 documents would result. It is then useful to use the documents view to examine the full text content of the found items.



Values	Frequency	Correlation
Boynton, Richard	36	12.1
Nakai, J H	8	8.0
Wiegand, B P	8	8.0
Wiener, K.	9	7.7
Jones, D L	7	5.6
Martinez, R.	5	5.1
Wiener, Kurt	4	4.1
Bell, R.	4	4.1
Gobbi, Massimiliano	4	3.1
Myzel, S J	3	2.9
Previati, Giorgio	3	2.9
Marshner, R L	3	2.9
Gobbi, M.	3	2.9
Lang, W E	3	2.9
Edington, L.	3	2.9

Fig. 12 Correlation of Author value to Inertia search string

Search Term (and synonyms)	Frequency
MPE	
mass , weight	2394
center of gravity, cg	773
moment of inertia, moi	100
product of inertia, poi	149
Product Development	
cost	1968
safety	814
reliability	577
management	1764
SAWE Industry Sector	
airline / "airline affairs"	232 / 3
aircraft / "military aircraft"	1908 /146
missiles / "missiles and space"	560 / 57
marine	223
ground / ground systems	1045 / 30
offshore	35
Standards	
standard, std	1597
recommended practice, rp	288

Fig. 13 Frequency of common MPE terms

Also of general interest is searching on some standard MPE terms as shown in Fig. 13. The fact that the terms for mass and its synonyms only turns up 2394 of 10177 documents (23%) is consistent with the database makeup in that there are on the order of 2000 SAWE documents in the corpus, but 8000 more from AIAA, INCOSE and other sources. In terms of general project development, reliability, and safety concerns there is a strong representation in cost, almost as many documents as indicated by mass. One might think that this aligns with the fact that mass properties are tightly related to product costing estimations. SAWE has a long history in this cross reference of terms by its past work regarding the "Value of a Pound" and the many associated technical sessions related to costing across all transportation sectors, not just aerospace. But one may also suspect a high contribution in this area from INCOSE as cost is a predominant characteristic in Systems Engineering analyses. Further search characterization by document source does return the fact that there is a high degree of SAWE papers in the cost area, 53%, with INCOSE following up with an additional 29%.

Searches can be optimistically misleading if not well enough defined. The term; standard/std, produces 1597 frequency hits. But if one is looking for mass standards the search would have to be more restrictive as just "standards" provides reference to a document any time any Military or other organizational standard is referenced in text, as well as the general English language use of the word standard. Note that the often utilized SAWE term Recommended Practice / RP supplies a more selective response.

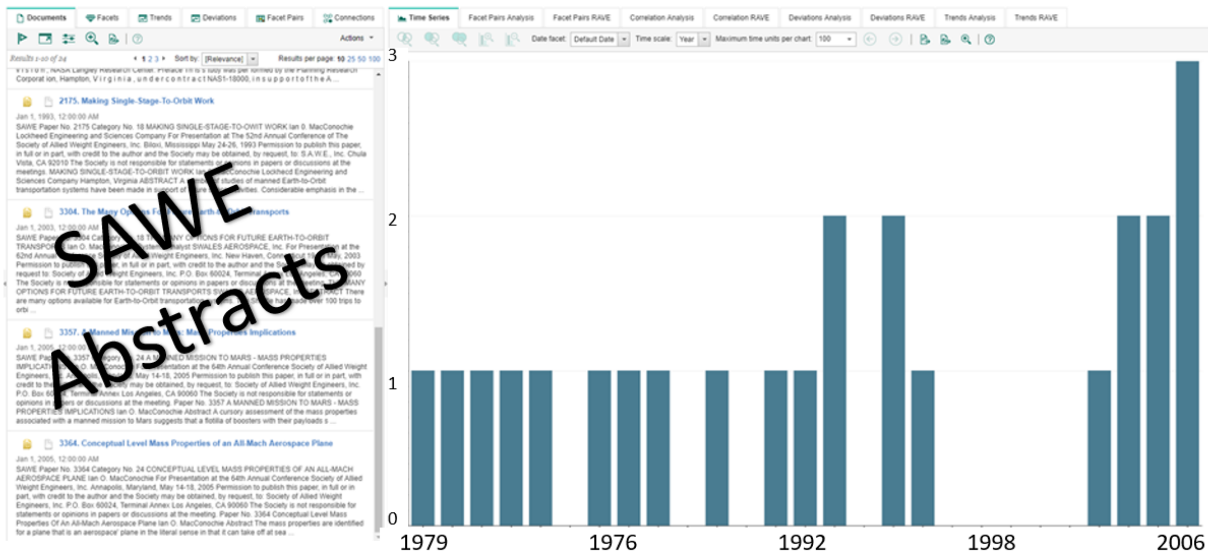


Fig. 14 Time series view of a single author's corpus contributions

Fig. 14 shows a time series view of of the author facet and facet value MacConochie with results representing the frequency hits of data items associated with this facet value. Also in this view is a documents pane which provides a summary excerpt from each of the found documents. In the case of an SAWE paper there is direct linkage to the document full text, other source material may show up as metadata which would have to be accessed outside of the WCA environment. Specific AIAA or INCOSE information for example is easily accessible if the user has appropriate search information for that suppliers publishing platform. Fig. 15 shows WEX provided metadata that can be used as input information for procuring such alternately published material.

As mentioned earlier using facet pairing is a powerful method to obtain relationships between search items and associated correlations. Facet pairing analysis returns frequency and correlation

Name	Value
Content Source/Collections	VAB/AIAA
Title	X-37 Flight Demonstrator Project: Capabilities for Future Space Transportation System Development
Journal	55th International Astronautical Congress of the International Astronautical Federation, the International Academy of Astronautics, and the International Institute of Space Law
Content Type	Metadata/XML

Name	Value
date	Jan 1, 2004, 12:00:00 AM
documenttype	VAB/AIAA
secondarytitle	55th International Astronautical Congress of the International Astronautical Federation, the International Academy of Astronautics, and the International Institute of Space Law
modifieddate	Jan 1, 2004, 12:00:00 AM
sourcetype	Metadata/XML
title	X-37 Flight Demonstrator Project: Capabilities for Future Space Transportation System Development

Fig. 15 Metadata access for query input to other source provider databases

	ersonnel isions	21. Residual Propellant and Service Items 2288	1. Aerodynamic Surfaces 2286	3. Induced Environment Protection 2242	12. Environmenta Control 2193	24. Thrust Decay Propellant 1881	25. Full Thrust Propellant 1379	18. Cargo 1130	27. Pre-Ignition Losses 1065	26. Thrust Propellant Buildup 1032	23. Inflight Losses (Liftoff to Separation) 391
Olds, John 48		5 0.1	0 0.0	0 0.0	0 0.0	4 0.1	1 0.0	1 0.0	0 0.0	0 0.0	0 0.0
Martin, James N 39		1 0.0	1 0.0	5 0.1	5 0.1	1 0.0	0 0.0	1 0.0	0 0.0	0 0.0	0 0.0
Boynton, Richard 38		34 2.6	38 3.0	30 2.3	31 2.4	29 2.6	22 2.4	38 6.2	10 1.1	8 0.7	3 0.3
Gilb, Tom 36		0 0.0	1 0.0	4 0.1	4 0.1	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0
Mavris, Dimitri N 35		2 0.0	6 0.2	2 0.0	2 0.0	1 0.0	1 0.0	1 0.0	1 0.0	1 0.0	0 0.0
Rhodes, Donna H 29		0 0.0	2 0.0	4 0.1	4 0.1	0 0.0	0 0.0	2 0.0	0 0.0	0 0.0	0 0.0
Sheard, Sarah A 27		0 0.0	1 0.0	3 0.1	3 0.1	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0
Armstrong, James R 27		0 0.0	1 0.0	1 0.0	1 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0
Lewis, Mark 26		5 0.2	4 0.1	1 0.0	0 0.0	5 0.2	0 0.0	1 0.0	0 0.0	0 0.0	0 0.0
Valerdi, Ricardo 25		1 0.0	0 0.0	3 0.1	3 0.1	1 0.0	0 0.0	1 0.0	0 0.0	0 0.0	0 0.0
Muller, Gerrit 25		2 0.0	2 0.0	4 0.1	4 0.1	2 0.0	0 0.0	1 0.0	0 0.0	0 0.0	0 0.0
Hause, Matthew 23		4 0.1	0 0.0	3 0.1	3 0.1	4 0.2	1 0.0	0 0.0	0 0.0	0 0.0	0 0.0
Ferris, Timothy L J 22		0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	1 0.0	0 0.0	0 0.0	0 0.0	0 0.0
MacConochie, Ian O. 21		21 2.7	20 2.5	20 2.5	19 2.4	20 3.0	18 3.5	18 4.3	17 4.2	17 4.3	9 4.6
Chudoba, Bernd 21		1 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0
Jacob, Jamey 21		1 0.0	2 0.0	0 0.0	1 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0
Forsberg, Kevin 21		2 0.0	0 0.0	1 0.0	1 0.0	2 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0
Haskins, Cecilia 20		0 0.0	0 0.0	3 0.1	3 0.1	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0
McCarty, J R 20		16 1.9	18 2.3	6 0.4	10 1.0	15 2.1	14 2.6	13 2.9	14 3.4	14 3.5	2 0.2 Bird's-eye View
Carson, Ronald S 19		0 0.0	0 0.0	1 0.0	1 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0
Ring, Jack 19		1 0.0	0 0.0	1 0.0	1 0.0	1 0.0	0 0.0	1 0.0	0 0.0	0 0.0	0 0.0

Fig. 16 Birds Eye view of a facet pairing between Author and FBS

values for the paired items. Frequency count provides insight into the number of documents occurring in the pairing and correlations provide understanding in the strength of relationships between paired items. Earlier we explained the concept of facet pairs and the viewing of facet pairing with the pairing between Author and Keywords. Now we look at a slightly narrower, but more organized view of comparing Author against the SME defined hierarchical Functional Breakdown Structure. Fig. 16 is what is called a "Birds Eye" view of facet pairing. We have paired any author, which includes some organizations as authors, against the mil38310 FBS. Documents from all levels of the hierarchical FBS taxonomy will be included in this analysis. Such a broad search results in a large number of both authors (rows) and breakdown structure terms to level III (columns). Indeed, this birds eye view is so large that it cannot be fully viewed at one time, it must be scanned by the user for areas of high interest. Scanning is implemented via the grey box in the lower right quadrant in Fig. 16. This region encompasses the full result of the facet pairing. Within the grey box is another box, a blue box in the upper right quadrant of the grey box. The blue box is a zoom region. Data shown in this figure which is the displayed text readable data is only the blue box partitioning of the larger full grey-box result. Users scan the grey box by panning the blue box to look for areas of high correlation. The two numbers in each row/column intersection of the displayed section of data represent frequency and correlation values, in this case the author frequency and correlation result with respect to the column FBS taxonomy term.

Note also that in a hierarchically defined facet, such as FBS all of the actual sub-facet values, not just the three level hierarchy names, can be used as column display values. This represents again a large number of terms, over 200, for our Level III FBS. A portion of this data view is provided in Fig. 17.

	system	vehicle	control	case	support	structure	air
Mavris, Dimitri	11/0.2	11/0.4	1/0.0	0/0.0	0/0.0	0/0.0	1/0.0
Olds, John	6/0.1	9/0.4	0/0.0	0/0.0	0/0.0	0/0.0	0/0.0
Martin, James N	37/1.2	0/0.0	2/0.0	2/0.0	5/0.2	6/0.3	1/0.0
Boynton, Richard	38/1.3	24/1.8	31/2.7	36/3.4	29/2.9	25/2.4	32/4.0
Gilb, Tom	16/0.4	0/0.0	7/0.3	1/0.0	5/0.2	2/0.0	0/0.0
Mavris, Dimitri N	8/0.1	15/1.0	0/0.0	0/0.0	1/0.0	0/0.0	5/0.2
Rhodes, Donna H	28/1.2	2/0.0	2/0.0	4/0.1	2/0.0	2/0.0	1/0.0
Armstrong, James R	21/0.9	0/0.0	1/0.0	2/0.0	0/0.0	2/0.0	1/0.0
Sheard, Sarah A	23/1.0	0/0.0	1/0.0	5/0.2	2/0.0	1/0.0	1/0.0
McCarty, J R	18/0.7	0/0.0	17/1.8	12/1.1	3/0.1	3/0.1	17/2.5
Lewis, Mark	4/0.1	9/0.7	0/0.0	0/0.0	0/0.0	0/0.0	2/0.0

Fig. 17 Partial keyword labeled view of facet pairing between Author and FBS taxonomy

Another example area of interest to mass properties engineers might be in knowing the history and currency of a particular subsystem technology. For example, how might materials dominance have changed over time and are there papers in existence which represent future expectations of materials yet in development. Or perhaps to investigate when does a new technology become prominent, such as when did winglets come to dominance in aircraft design. Big Data data interrogation of a corporate corpus and subsequent document review answers questions like these and provides the reference backup information.

Mass properties predictions for new vehicle designs greatly depend on a corporate knowledge derived Basis of Estimate for related vehicle types. Having a large and well-documented BoE is of paramount importance in coming up with reliable mass properties early in a project's life cycle. WEX can assist in building historically based mass estimations with a rigorous BoE behind them, often a required artifact in contracted acquisition work. Note too that "history" can be not only from production vehicles but also from concept studies where project maturity state could be incorporated as a field in the BoE database. With only the use of WEX search capabilities, and not follow on capabilities in Artificial Intelligence and Machine Learning, the knowledge building process is still human in loop intensive. Progressing from WEX search capabilities to then understanding of retrieved data thru AI techniques, to transform such into mass properties knowledge, should be a future goal of BoE building procedures. To illustrate a path that can be taken using current LaRC WEX capabilities and user intensive data processing we consider how building Mass Estimating Relationships for cryogenic liquid hydrogen fuel tanks could be performed.

Table 1 is an almost direct output from WEX called a summary report, created from the document view. This report can be made in MS/Excel format as was done here, the heading row and column 2 flag are user added. This table becomes the data source reference link to a database of human processed information useable for BoE and MER building activities.

Table 1 - Example WEX document report data

Summary	Included in database
SAWE Paper No. 1317 MASS PROPERTIES O F FLUIDS IN LARGE CONTAINERS Lewis T. Hassman, Member of Technical Staff North American Space Operation Rockwell International For Presentat ion a t the 38th Annual Conference of the Society of Allied Weight Engineers, Inc. New York, New York 7-9 May …..	no
COMPOSITE STRUCTURES ON THE DC-XA REUSABLE LAUNCH VEHICLE Michael J. Robinson McDonnell Douglas Aerospace Huntington Beach, California 92647 ABSTRACT In the summer of 1993, under the sponsorship of the Ballistic Missile Defense Organization (BMDO), …………	yes
Microsoft Word - IEEE_2011-1044_final.docx 1 Simple, Robust Cryogenic Propellant Depot for Near Term Applications IEEE 2011-1044 Abstract 1â€" The ability to refuel cryogenic propulsion stages on-orbit provides an innovative paradigm shift for space transportation supporting National Aeronautics and Space Administration ……………	no
Continued rows for all papers in current query…..	**yes/no**

That "knowledge" database is of the form shown in Table 2. Fields of interest for component characterization and scaling are defined in this table and at this point it is up to the WEX user to read the references discovered in Table 1 and extract information to fill in a Table 2 type component characterization of knowledge. Information of this type is also valuable for loading into a corporate [11] process, or a commercial weight management tool enabling COTS provided weight management capabilities in for example: "Statistical regression on historical data for parametric estimation" [12], or "Weight Analytics" [13].

Table 2 - Human Processed Data Supplementation from Summary Report Items

title	vehicle	sub description	capacity, gal	loadpath integration	liner	structural material	reusable	areal weight, lbm/ft^2	area ft^&2	tank weight, lbm
See summary report for list of included references		summaryReport_85resultsLh2TankSearch.xlsx								
query state		mass + "liquid hydrogen" + tank								
COMPOSITE STRUCTURES ON THE DC-XA REUSABLE LAUNCH VEHICLE	DC XA	reusable launch vehicle	6000	integral	no	graphite composite	yes			
COMPOSITE STRUCTURES ON THE DC-XA REUSABLE LAUNCH VEHICLE	DC XA	reusable launch vehicle	8x16 cyl		no	graphite composite	yes			
AIAA-80-0407 Thermostructural Analyses of Structural Concepts for I Hypersonic Cruise Vehicles	Mach 5 Cruiser	all metal multiwall external foam internal zees		integral				6.98		
AIAA-80-0407 Thermostructural Analyses of Structural Concepts for I Hypersonic Cruise Vehicles	Mach 5 Cruiser	TG-15000 multiwall external foam internal zees		integral				5.92		
AIAA 2000-3141 Mass Breakdown of the Saturn V	Saturn Stage II	tanks and insulation mass		integral	no	aluminum				
AIAA 2000-1043 Reusable Launch Vehicle Tank/Intertank Sizing Trade Study	SSTO X-33	LL401							9439	12971
AIAA 2000-1043 Reusable Launch Vehicle Tank/Intertank Sizing Trade Study	SSTO X-37	LA1							9304	13328
AIAA 2000-3141 Mass Breakdown of the Saturn V	Saturn Stage II	tanks and insulation mass		integral	no	aluminum				29723.4

Summary

A capability such as provided by WEX type data interrogation has shown itself to be extremely useful for various aspects of Mass Properties Engineering. It is noted in general that having electronic search access to the authors personal material which was OCR scanned into the corpus is far superior than the prior file cabinet based system. Documents of interest to a topic were often unveiled via electronic search which would not have been found through a hardcopy file based approach. Several questions and areas of potential future capabilities can be formulated based upon this trial experiment in creation of the Vehicle Analysis Analytics corpus and using it with the IBM WEX software.

- Direct access to WEX loaded full text information is much more user friendly and productive to use than having to subsequently use WEX found metadata to later obtain full text files from a professional society, academic or corporate web based data provider.
- Obtaining access to full text document data is critical to developing rigorous vehicle knowledge bases. The SAWE material obtained for this project was done so on a leased basis and so will not be available after project completion. Occasional access to subscription based material should be planned for by the WEX user if it is desired to maintain a corpus over time.

- Users of WEX need to assure they can obtain source material and utilize it in accordance with provider access terms. Perhaps supplier organizations may consider the advantages of having a Big Data computer accessible public interface to their metadata and full text data.
- Future computing trends such as cloud based services and the pricing flexibilities available for providing analytics as a service [ex: IBM Blue Mix https://console.bluemix.net/catalog/] could assist in providing Big Data capabilities to a larger user group.
- Can "personalized versions" of an analytics site be created for individual users, ex: a WEX user has purchased access to AIAA documents, and SAWE documents, can they be incorporated into a personally tailored version of an employer's larger Big Data infrastructure?
- There are additional sources of publically available vehicle design information such as NASA's NTRS and LaSSe systems. Incorporation of such data should be included in a future version of the "Vehicle Analysis Analytics" corpus.
- OCR scanned data maybe software processed to a format suitable for loading as WEX content but this is not without considerable human interface activity to clean up the data mapping process. Therefore, paper authors and their publishing organizations should always enforce the filling out and tracking of standard metadata information.
- Though not tried out in this project the inclusion of web crawling features is expected to be very useful as well as just static corpus sourcing of material.
- Application of Machine Learning and Artificial Intelligence are the next logical steps to take in creating a knowledge base from big data content inquiries. In this manner Mass Estimating Relationships and other features of Mass Properties Engineering could be developed more easily and provide intuitions a manual build process might not discover.

Conclusions

In conclusion it can be stated that an organization such as NASA will derive great utility for its various project and discipline engineers by providing the capability for those persons to use WEX type document search capabilities and to subsequently build and maintain knowledge bases. This project permitted the removal of several office file cabinets, replacing them with online data which is more easily accessed and searched than the hardcopy file system. Many times information was recalled thru WEX inquiries which would have been missed had the author relied upon looking thru hardcopy material. Building of the WEX corpus facet values, parsing rules and taxonomies is an iterative process between the discipline expert and the information technologist developing WEX implementations. For a discipline like Mass Properties Engineering having broadly accepted breakdown structures is very useful in being able to correlate values sourced from a variety of data sources. This well recognized aspect of MPE shows its utility by easing the currently manual process of turning mass properties data into mass properties knowledge. For example, by the creation of estimating relationships at different levels of an FBS taxonomy. The next logical step would be in using machine intelligence to assist the MPE knowledge builder to create their desired mathematical regressions.

References

[1] IBM Corporation, "IBM Watson Analytics," [Online]. Available: https://www.ibm.com/watson-analytics.

[2] IBM corporatio, "Content Analytics with Enterprise Search (video training)," [Online]. Available: https://mediacenter.ibm.com/channel/t/33942902.

[3] United States Air Force, "MIL-M-38310B Mass Properties Control Requirements For Missile and Space Vehicles," 1971.

[4] IBM Corporation, "IBM knowledge center, Query Syntax," [Online]. Available: https://www.ibm.com/support/knowledgecenter/SS5RWK_3.5.0/com.ibm.discovery.es.ap.doc/iiyspqysyntx.htm.

[5] Y. A. Ambur, J. J. Yagle, W. Reith and E. McLarney, "Big Data Analytics and Machine Intelligence," in *NASA/TM–2016-219361*, Dec. 2016.

[6] Mendeley, Elsevier Group, [Online]. Available: https://www.mendeley.com/.

[7] "Crossref," [Online]. Available: https://www.crossref.org/dashboard/.

[8] NASA Technical Report Server, [Online]. Available: https://ntrs.nasa.gov/.

[9] NASA, "Launch and Space Systems e-Library (LaSSe)," [Online]. Available: https://explornet.nasa.gov/docs/DOC-23876.

[10] International Business Machines, IBM, "IBM Watson Content Analytics - Discovering Actionable Insight from Your Content," July 2014. [Online]. Available: https://www.redbooks.ibm.com/redbooks/pdfs/sg247877.pdf. [Accessed 17 01 2018].

[11] A. Walker, "Development of a conceptual flight vehicel design Weight Estimation Library," Denver CO, 2016.

[12] BAS Engineering, "ShipWeight - Weight Engineering Software," [Online]. Available: http://shipweight.com/. [Accessed 18th January 2018].

[13] Altair Corporation, "Weight Analytics," [Online]. Available: http://www.altairproductdesign.com/Page.aspx?category=PSO&item=Weight+Analytics. [Accessed 18th January 2018].

Appendix A – FBS Taxonomy based on Mil 38310B

Level I

1	Aerodynamic Surfaces
2	Body Structure
3	Induced Environment Protection
4	Laundh, Recovery and Dockng
5	Main Propulsion
6	Orientation Controls, Separation and Ullage
7	Prime Power Source
8	Power Conversion and Distribution
9	Guidance and Navigation
10	Instrumentation
11	Communication
12	Environmental Control
13	Armament
14	Personnel Provisions
15	Crew Station Controls and Panels
16	Range Safety and Abort
17	Personnel
18	Cargo
19	Ordinance
20	Ballast
21	Residual Propellant and Service Items
22	Reserve Propellant and Service Items
23	Inflight Losses (Liftoff to Separation)
24	Thrust Decay Propellant
25	Full Thrust Propellant
26	Thrust Propellant Buildup
27	Pre-Ignition Losses

Level II

1. Aerodynamic Surfaces
   1.1 Fixed Surfaces
   1.2 Movable Surfaces
   1.3 Fairings and Associated Structure (Including Insulation)
2. Body Structure
   2.1 Structural Fuel Tank
   2.2 Structural Oxidizer Tank
   2.3 Structural Propellant Tank (Common Bulkhead)
   2.4 Structural Solid Propellant Case or Tank
   2.5 Structure Enclosing Nonintegral Tanks
   2.6 Structure Forward of Integral Tanks
   2.7 Structure Between Integral Tanks
   2.8 Structure Aft of Integral Tanks
   2.9 Thrust Structure
   2.10 Interstage/Spacer/Vehicle instrument Unit Structure
   2.11 Pressurized Compartment
   2.12 Nonpressurized Compartment
   2.13 Ballast
   2.14 Multipurpose Equipment-Containers, Panels Supports
   2.15 Exterior Finish and Sealer
   2.19 Contingency
   2.20 Growth
3. Induced Environment Protection
   3.1 Thermal Protection (Active)
   3.2 Thermal Protection (Passive)
   3.3 Noise Protection
   3.4 Meteorite Protection
   3.5 Radiation Protection
   3.19 Contingency
   3.20 Growth
4. Launch, Recovery and Docking
   4.1 Launch Gear
   4.2 Deployable Aerodynamic Devices
   4.3 Landing Gear
   4.4 Flotation Gear
   4.5 Docking Structure
   4.6 Recovery Aids
   4.19 Contingency
   4.20 Growth
5. Main Propulsion
   5.1 Liquid Rocket Engine and Accessories
   5.2 Solid Propellant Engine and Accessories
   5.3 Nuclear Power Plant and Accessories
   5.4 Ion Engine and Accessories
   5.5 Photon Engine and Accessories
   5.6 Air Breathing Engine and Accessories
   5.7 Purge System for Stage Chilldown
   5.8 Fuel Container (Nonstructural)
   5.9 Fuel System
   5.10 Pressurization, System - Fuel
   5.11 Oxidizer Container (Nonstructural)
   5.12 Oxidizer System
   5.13 Pressurization System - Oxidizer
   5.14 Auxiliary Fluids System
   5.15 Propellant Utilization System
   5.16 Lubricating System
   5.19 Contingency
   5.20 Growth
6. Orientation Controls, Separation and Ullage
   6.1 Thrust System (Main)
   6.2 Thrust System (Auxiliary)
   6.3 Aerodynamic Control
   6.4 Spatial Control
   6.5 Separation
   6.6 Ullage (Separate from 6.4, 6.5)
   6.7 Fuel Containers
   6.8 Oxidizer Containers
   6.9 Pressurization
   6.10 Distribution and Control - Fuel
   6.11 Distribution and Control - Oxidizer
   6.12 Thrust Structure
   6.13 Artificial Gravity
   6.19 Contingency
   6.20 Growth
7. Prime Power Source
   7.1 Power Source
   7.2 Power Source - Fuel Cell
   7.3 Power Source - Batteries
   7.4 Power Source - Solar Cell
   7.5 Power Source - Nuclear
   7.6 Power Source - Gas Generator
   7.19 Contingency
   7.20 Growth
8. Power Conversion and Distribution
   8.1 Power Conversion - Electrical (AC)
   8.2 Power Conversion - Electrical (DC)
   8.3 Power Conversion - Hydraulic/Pneumatic
   8.4 Power Distribution - Electrical (AC)
   8.5 Power Distribution - Electrical (DC)
   8.6 Power Distribution - Hydraulic/Pneumatic
   8.7 Utility Provisions - Electrical
   8.8 External Service Provisions (Type Power)
   8.19 Contingency
   8.20 Growth

9 Guidance and Navigation
   9.1 Guidance Source - Inertial/Stellar/Planetary/Relative Reference
   9.2 Guidance Evaluation
   9.3 Output
   9.4 Spares
   9.19 Contingency
   9.20 Growth

10. Instrumentation
   10.1 Sensors
   10.2 Signal Conditioning
   10.3 Signal Transmission
   10.4 Electrical Coupling
   10.5 Support Items
   10.6 Spares
   10.19 Contingency
   10.20 Growth

11. Communication
  11.1 Intercommunication
  11.2 Near Earth Communication
  11.3 Deep Space Communication
  11.4 TV Systems
  11.5 Tracking System
  11.6 Spares
  11.7 Electrical Coupling
  11.8 Racks & Supports
  11.19 Contingency
  11.20 Growth
12. Environmental Control
  12.1 ECS - Equipment
  12.2 ECS - Personnel
  12.3 Coolant System
  12.4 Multipurpose Equipment - Containers, panels, etc.
  12.19 Contingency
  12.20 Growth
13. Armament
  13.1 Safety and Arming
  13.2 Detonating
  13.3 Warhead Provisions
  13.19 Contingency
  13.20 Growth
14. Personnel Provisions
  14.1 Accommodations for Personnel
  14.2 Fixed Life Support Equipment
  14.3 Cargo Handing
  14.4 Furnishings
  14.5 Emergency Equipment
  14.19 Contingency
  14.20 Growth
15. Crew Station Controls and Panels
  15.1 Pedestal
  15.2 Control Stands
  15.3 Instrument Panels
  15.4 Crew Station Controls (Flight)
  15.19 Contingency
  15.20 Growth
16. Range Safety and Abort
  16.1 Sensors
  16.2 Signal Conditioning
  16.3 Signal Transmission
  16.4 Signal Evaluation
  16.5 Destruct System
  16.6 Electrical Coupling System
  16.7 Service Items
  16.8 Spares
  16.9 Debris, Re-entry
  16.19 Contingency
  16.20 Growth
17. Personnel
  17.1 Crew
  17.2 Personal Gear
  17.3 Life Support
  17.4 Crew Accessories
18. Cargo
  18.1 Scientific instruments
  18.2 Experiments
  18.3 Cargo
19. Ordnance
  19.1 Propulsion Ordnance
  19.2 Orientation Controls, Separation and ullage
  19.3 Prime Power Source
  19.4 Armament
  19.5 Range Safety and Abort
20. Ballast
21. Residual Propellant and Service Items
  21.1 Fuel Pressurizing Gas - As Container Residual
  21.2 Oxidizer Pressurizing Residual - As Container Residual
  21.3 Fuel Trapped - Main Engine
  21.4 Oxidizer Trapped - Main Engine
  21.5 Fuel - Outage Main Engine
  21.6 Oxidizer - Outage Main Engine
  21.7 Fuel Trapped - Auxiliary Propulsion System
  21.8 Oxidizer Trapped - Auxiliary Propulsion System
  21.9 Fuel - Outage Auxiliary Propulsion System
  21.10 Oxidizer - Outage Auxiliary Propulsion System
  21.11 Fuel Trapped - Electrical Power
  21.12 Oxidizer Trapped - Electrical Power
  21.13 Service Items Trapped - Including Non-expendables
22. Reserve Propellant and Service Items
  22.1 Fuel Pressurizing Gas Reserves
  22.2 Oxidizer Pressurizing Gas Reserves
  22.3 Fuel - Main Engine Reserve
  22.4 Oxidizer - Main Engine Reserve
  22.5 Fuel Pressurizing Gas Reserves Auxiliary Propulsion Systems
  22.6 Oxidizer Pressurizing Gas Reserve - Auxiliary Propulsion Systems
  22.7 Fuel Reserves - Auxiliary Propulsion Systems
  22.8 Oxidizer Reserves - Auxiliary Propulsion Systems
  22.9 Fuel - Electrical Sewer Reserve
  22.10 Oxidizer - Electrical Power Reserve
  22.11 Fuel - Atmospheric Cruise
  22.12 Oxidizer - Atmospheric Cruise
  22.13 Service Item Reserves
23. Inflight Losses (Liftoff to Separation)
  23.1 Vented Pressuring Gas-Fuel System
  23.2 Vented Pressuring Gas-Oxidizer System
  23.3 Fuel - Auxiliary Propulsion System
  23.4 Oxidizer - Auxiliary Propulsion System
  23.5 Fuel - Electrical System
  23.6 Oxidizer - Electrical System
  23.7 Service Items
24. Thrust Decay Propellant
  24.1 Fuel - Consumed During Thrust Decay
  24.2 Oxidizer - Consumed During Thrust Decay
  24.3 Solid Propellant - Consumed During Thrust Decay
  24.4 After-Cooling Propellants (Nuclear)
  24.5 Service Items

25. Full Thrust Propellant
    - 25.1 Fuel
    - 25.2 Oxidizer
    - 25.3 Solid Propellants
26. Thrust Propellant Buildup
    - 26.1 Fuel
    - 25.2 Oxidizer
    - 26.3 Solid Propellants
    - 26.4 Auxiliary Propulsion System
    - 26.5 Service Items
    - 26.2 Oxidizer
27. Pre-Ignition Losses
    - 27.1 Fuel
    - 27.2 Oxidizer
    - 27.3 Prime Power Source
    - 27.4 Service Items

Jeff Cerro is a registered Professional Engineer with a Masters in Mechanical Engineering from the Rensselaer Polytechnic Institute. He is currently a vehicle systems analyst and mass properties engineer at the NASA Langley Research Center in Hampton Virginia. His current interests are in Model Based vehicle concept development approaches with increased dependence on digital data and collaboration techniques. Utilizing "Big Data" search techniques is a portion of that activity that is recognized as useful to assure burgeoning as well as historical knowledge is not lost in the quest for new vehicle designs. Mr. Cerro is currently Chairman of the Standards Review Board for the Society of Allied Weight Engineers (SAWE).

Theodore D. Sidehamer Jr. is an applications developer with a Bachelor's Degree in Business Administration (Old Dominion University). He currently serves as the Watson Explorer (WEX) application administrator and developer, employed by Craig Technologies, Inc., at the NASA Langley Research Center in Hampton Virginia.

Dorothy L. Notarnicola is a librarian with a Master's Degree in Library Science (University of Pittsburgh) and a Master's Degree in Business Administration (Penn State). She currently serves as the lead contract librarian, employed by NCI, Inc., at the NASA Langley Research Center in Hampton Virginia.