FROM NACA TO NASA

Visualizing the Research Paper Library of the
National Advisory Committee for Aeronautics
(1915-1958)

Antonie Christoffel Dreyer

May 2020

Submitted in partial fulfillment of the requirements for the degree of Master Science in Data Visualization at Parsons School of Design.

Acknowledgements

"Good design reminds you to think of things you didn't even know you were asking for."
—Daniel Sauter

Firstly, I would like to thank my sponsor; the Fulbright program which made this study possible, as well as the Institute of International Education (IIE) who facilitated the process. I would then like to thank my Professors and supervisors, Prof. Daniel Sauter, Prof. Aaron Hill and Prof. William Bevington for valuable inputs on design, their kindness and firm, but gentle guidance through the difficult transition period when the global pandemic struck. I especially appreciate their understanding in allowing me to complete my projects during my repatriation process.

I would also like to thank my classmates, firstly for companionship and support, and secondly for supplying valuable insights into aspects of visual designs, styles, and honest and helpful critiques.

Thanks also to the management of the 92Y Residence (a Jewish community center) who were greatly empathetic to the plight of international students when University accommodation had to be evacuated and they renegotiated my accommodation contracts on my behalf which allowed me to stay in my dorm room.

Then I would like to thank my parents and girlfriend Eunice for continued support across time zones, many days of work and asynchronous communication.

Finally, I want to thank God for supplying in my earthly needs, giving me a peace and hope for a future in a world where it might feel like everything is falling apart.

Abstract

During World War I aircraft technology was in its infancy, which prompted the United States to establish the National Advisory Committee for Aeronautics (NACA) in 1915. Research, testing, and development were undertaken by scientists and aviators and "the NACA"[1] significantly advanced commercial and military aviation during the Golden Age of Flight. By the time World War II ended, the NACA was well established as a leading research institution. When the space race began in earnest in the late 1950's, the NACA's proven research model was ideal for advancing technologies from the field of aeronautics toward aerospace[2]. In 1958 the National Aeronautics and Space Administration (NASA) was created; this was when NACA became NASA.

This thesis project visualizes the entire NACA research paper collection with document metadata to add context into a comprehensive knowledge system. The NACA research model originally exhibited successful creation and sharing of complex knowledge and with the current University of North Texas (UNT) NACA database is ideally suited for exploring this data set using relationships that were meticulously catalogued and classified by its creators.

Original scans of Table of Contents and Index records from the NACA Research Division were digitally processed (using existing OCR text) to allow visual exploration of the papers binned into subject headings. A flattened hierarchy combined with a basic network is used to alleviate the tension between top-down and bottom-up data structures used to link information. The NACA metadata illustrates how the use of classic hierarchical classification can co-exist with network-like links to manage and explore complex multi-disciplinary subject knowledge to gain deeper insight into a complex field.

---

[1] Originally and formally the Committee was referred to as "The N.A.C.A."; pronounced with individual letters ("en-ay-see-ay"). However, later it colloquially started to be pronounced as an acronym ("NACKA"). Its successor NASA (National Aeronautics and Space Administration) was formally pronounced as an acronym since its inception in 1959. The forms "NACA" and "The NACA" are used interchangeably in this thesis, depending on its context in a sentence.
[2] "Aeronautics" refers to atmospheric flight and "aerospace" is a broader term that includes both air and space.

Table of Contents

Acknowledgements 1

Abstract 2

Table of Contents 3

Table of Figures 4

1 Introduction 5

2 The Nature of Knowledge 7

2.1 Data, Information, Knowledge & Wisdom 7

2.2 Some Limitations of Texts 8

2.3 Tables of Contents vs. Indexes 9

2.4 Knowledge Management and Data Mining 10

2.5 Knowledge Trees and Networks 12

3 How NACA became NASA 16

3.1 NACA Research History 16

3.2 The NACA Library Metadata Exploratory Analysis 19

3.3 Metadata Structure 24

3.4 Visual mapping 32

3.5 Implementation 34

3.6 Visual Style 38

4 Conclusion 41

5 References: 43

Appendix A: Web Page Abstract 46

Appendix B: NACA Report Distribution 47

Appendix C: Full Subject list 48

Table of Figures

Figure 1: A metaphor of Trees engulfed in webs (Watkins,2010 UK DFID) 13

Figure 2: Langley Research Center Growth...16

Figure 3: Numbers of NACA documents by Type..19

Figure 4: Average Report Subject Distribution (see Table 1)..20

Figure 5: Subject Coverage by Year 21

Figure 6: Subject Heading Word Map (horizontal) vs. Time (vertical) 23

Figure 7: Extract from NACA Index of Technical Publications (NACA, 1954) 25

Figure 8: NACA Index Record 26

Figure 9: Document Example (NACA-TR-1107) 27

Figure 10: One document (leaf) in many subject headings (branches).......................................28

Figure 11: One document (leaf) in many subject headings (branches).......................................28

Figure 12: UNT API XML OUTPUT 29

Figure 13: Data Wrangling and Visualization Process Logic..31

Figure 14: D3 Force Layout of all records as points..36

Figure 15: D3 Force Layout Closeup showing point cluster spacing...37

Figure 16: The NASA Logo (Note the font)..38

Figure 17: Web Site Opening Section 39

Figure 18: Example State of Main Visualization 40

1 Introduction

'There is no "I" in "knowledge worker.'
—A. McFay

A defining characteristic of many modern societies is the use of the mind's power more than the power of the muscle (Powell, Snellman, 2004). The rise of knowledge economies imply that accumulating knowledge becomes a central concern for workers whose livelihoods rely on obtaining, maintaining, and transferring knowledge. Subsequently knowledge management is becoming crucial in both industrializing and post-industrialized societies.

Through accumulating and adapting knowledge, humans have managed to rise from the stone age into today's information age. Although the elements of progress might have changed with methods of production, human ingenuity and reliance on pre-existing knowledge has remained, often hidden in the background of success. We might be able to expertly use various technologies in our daily lives, but we will not be able to recreate these technologies in isolation, nor fulfill our daily roles without these technologies. We keep relying on the knowledge of others to survive and thrive.

John F. Kennedy famously said, "*it will not be one man going to the moon—if we make this judgment affirmatively, it will be an entire nation*." This aspect of teamwork and sharing both muscle power and brain power is partly what landed man on the moon. Without the ability to store and navigate complex knowledge, humanity would not be able to reach the skies, physically or proverbially.

Greater need for knowledge management also illustrates a wider characteristic of industrialization; that we might have been able to leverage the ability to develop technologies at greater tempos, but it often makes it difficult for people to keep up physically, cognitively or ethically. We have entered a time where people now need to learn to swim and keep from drowning in a deluge of information. If we want to make technological and societal progress, it requires distributed understanding of complex issues. Therefore, communication of complex subject matter is important to advance modern society generally and collectively, not just to get to the Moon or to Mars.

To explore the visual representation of complex interdisciplinary knowledge, this thesis visualizes distinct areas of aerospace engineering developed over the past century and integrates this complex interdisciplinary knowledge into one comprehensive knowledge system.

Drawing inspiration from the model research conducted at the National Advisory Committee for Aeronautics (NACA) from 1915-1958, a broad history of aircraft research and development is presented, and meaning is extracted by interconnecting various topics in aerospace engineering. Topics are related through intra- and inter-disciplinary connections.

To further explore the aspect of knowledge transfer, a brief overview of knowledge management will be given to supply some perspective on historical knowledge management and library cataloguing. Thereafter a brief historical introduction is given to the NACA, after which the NACA research repository is presented as the primary data source for the main data visualization plot. Attention will be given on extracting the metadata in a way that is meaningful to supply context and focus. Finally, some aspects of implementation about the online visualization will be discussed.

2 The Nature of Knowledge

“Knowledge is knowing that a tomato is a fruit, but wisdom is knowing not to put it into a fruit salad”
—Miles Kington (Kington,2003)

2.1 Data, Information, Knowledge & Wisdom

The Data Information Knowledge and Wisdom Hierarchy (DIKW) fits into a pyramid model which is extremely useful to explain complex multidisciplinary topics, with data at the bottom of the pyramid and wisdom (or “insight”) at the top (Roberts,2015). The DIKW hierarchy is centered around the notion that to gain wisdom (insight), one must first gain knowledge. Knowledge in turn, consists of contextualized information, and information is in turn, data with added meaning (Ackoff,1989). This logic of using data to obtain insight about a certain topic is at the core of the scientific process: to do experiments, gather “low-level” data and draw conclusions in order to accumulate knowledge and insight (i.e. “wisdom”).

However, in the context of knowledge transfer, mentorship and education, the pursuit of wisdom is not always a straight-forward task, especially considering traditional teaching methods. Often traditional teaching is focused on information transfer with “data'' curated by an instructor. At best it could become knowledge transfer with “information” curated by an expert. Hence, methods of self-exploratory study are increasingly popular and education reform is a hot topic. However, information curation always seems to be partially required, because students often face the problem that they “don’t know what they don’t know” (Ackoff,1989). Students without any guidance might find it difficult to know where to start searching for information about a specific topic, or rather, what keyword to type into the Google search bar. In theory, given enough time, anyone could find the information they need, but when time is money, especially in context of Knowledge economies, time is often a luxury few can afford.

Knowledge management systems aim to solve the problem of information curation, even though it implies that a search for insight often consists mainly out of information processing. Information processing then easily devolves into actions that resemble text mining more than a search for wisdom (Geisler,2008,p.259). In contrast, effective knowledge management is more concerned by answering the questions of “what”, “how”, “why”. Therefore, to gain deeper levels of insight (given limited time), some information needs to be curated and/or constructed through automated searches.

Although digital information tools might not necessarily hold the promise of instant insight or wisdom with automated "wisdom searches", some levels of automation might certainly help users to smooth the process of digesting large amounts of information. Searches often require traversing through layers of information that begs the need for being contextualized; a task not always possible with traditional cataloguing systems such as libraries or archives.

2.2 Some Limitations of Texts

When Gutenberg invented the printing press it revolutionized learning because information could be transferred much more effectively than ever before. Knowledge could persist after its creator perished and people would be able to learn subjects without the presence of a teacher or a tutor. However, books have an inherent limitation of presenting information in a linear and mostly text-based way; and more importantly, books are limited by their size and number of pages between the covers. To contextualize complex subjects in a field such as Aerospace engineering, a whole library of various types of media would be needed to explain all aspects of a specialized topic such as Aeroelasticity[3].

Books and papers (even electronic ones) are generally a good way of presenting stories or concepts that advances along a certain train of thought, however when dealing with complex system topics, non-linear presentation of information becomes more relevant. Sometimes multiple books or papers are required to explore and highlight different aspects of a specific field, or alternatively some types of information need to be communicated using diagrams, graphs, maps or other visualization tools. It would hardly be feasible to describe a star chart or a world map using only words.

To extract meaning from various sources of information, relationships and context need to be drawn in non-linear ways and teachers, experts, librarians, and archivists typically link or catalogue those connections. Libraries have therefore played a major role in the curation of knowledge throughout history, but these increasingly face the challenge of keeping up with the times. Even though the Dewey Decimal library cataloguing system was widely used in the 20th century, it became apparent that it could not keep up with the demands of increasingly disparate informational topics (Weinberger,2008).

[3] Aeroelasticity is the multi-disciplinary field that concerns interaction between airflow and elastic structures such as airplane airframes, an example being "Wing Vibration and Flutter".

The Library of congress has a much more expansive cataloging system, yet remains limited (Bair,2005). Even though cataloguers have done commendable jobs in creating meaningful informational structures, some level of subjectivity is always present and even master librarians "don't know what they don't know".

It would seem as if knowledge cataloguing systems are forever bound to the process of lumping information topics into categories or splitting them into sub-categories which ultimately result in a hierarchical tree-like structure of knowledge; a structure that might not be sufficient when previously unknown types of information are discovered or created. With the sheer amount of information being generated on the internet, and traditional top-down hierarchical classification systems running into limitations, non-hierarchical bottom-up tagging or indexing seems to be the next logical alternative.

2.3 Tables of Contents vs. Indexes

Top-down Tables of Contents have long graced the front pages of books, but *bottom-up indexes* have also been used (and often overlooked) at the back of some textbooks for centuries (Knight,1979). A similar manifestation of bottom-up navigation found its way to the internet as tagging on media platforms, which might be likened to indexes in books, but having the advantage spanning across multiple sources. The introduction of the hashtag has become the primary way to navigate archives of images that remain mostly unclassified from a traditional top-down perspective (e.g. Instagram).

Tagging or indexing has subsequently become a fundamental way we navigate information in social media networks and collections of miscellaneous pictures on the internet. It should come as no surprise, since direct linking of information is broadly in line with the philosophy of hypertext and the original inspiration of what the internet should be; a web of information linked without a central hierarchical classification (Nelson,1965). The traditional library catalogue of parent-child relationship has therefore been mostly replaced with linking peer-to-peer siblings in networks[4].

The irony seems to be that in solving the problem of the top-down hierarchy of catalogues by creating links and tagging indexes in networks, the big picture might have been lost and everything is now classified under "miscellaneous". No longer is the structure clear

---

[4] The idea of a network of information gave rise to many different browsers during the time of the dot com boom. Some of these browsers were visual, such as the Xerox hypertree browser.

by looking at the table of contents or library catalogue, which does not exist for these "miscellaneous" items that are interconnected in a web. Subsequently, the age of the Internet, and going into the future, "the task of knowing is no longer to see the simple. It is to swim in the complex." (Weinberger,2008). As Weinberger puts it: "In the world after the Enlightenment, the cultural task was to build knowledge. In the miscellaneous world, the task is to build meaning." (Weinberger,2008).

Hence, we might have obtained the knowledge that a tomato is a fruit, but we still need to teach each other about not using it in a fruit salad. Therefore, to build meaning out of knowledge, it is deemed worthwhile to further explore what knowledge actually is. This topic is explored is in the following section, after which it can be further expanded to show how it relates to hierarchical classification and/or indexing in networks.

2.4 Knowledge Management and Data Mining

A more general way of viewing the nature of knowledge is to understand that knowledge is more than just bits of data, but rather it could be considered as nuggets of information clustered or linked in some way. These nuggets might be paintings on a cave wall, a chapter in a textbook, a paragraph in a research paper, some marks on a diagram or a set of neuron signals in someone's brain.

Even though no known quantified standard of units of knowledge exists, the use of knowledge nuggets seem effective in the context of "data mining" where a higher level of insight could be obtained if one were able to do "knowledge mining". However, further abstraction would be required to explain efficient storage and transfer of knowledge (Geisler,2008).

A key aspect of knowledge is that it can be divided into "tacit" vs. "explicit" knowledge. Tacit knowledge is considered knowledge that someone has in their mind, whereas explicit knowledge can be documented in text, images, or other formats. To obtain and retain expert knowledge, we often need to gain access to the tacit knowledge in the mind of a subject specialist and convert it to explicit knowledge to a format that can be stored in an archive. Unfortunately, the efficiency of encoding tacit knowledge into explicit knowledge and decoding it back into the mind of a learner is often dependent on technology. We cannot yet download or upload information to people's brains as Elon Musk intends to do with Neuralink (Markman,2019) and our eyes and ears are still the primary means of absorbing knowledge.

Knowledge transfer is also hampered by individual communication and learning styles, which is especially evident when certain types of communication technology favors certain communication and learning styles. This is a primary reason for education reform, since some people learn more effectively reading texts (i.e. traditional teaching methods) whereas others learn more effective visually (with diagrams), auditory (with sound) or kinesthetically (with movement). This multimodal learning is often employed at museums, with physical installations, interactive touch areas and various interactive sounds. These museums are often more memorable and engaging than textbooks because they engage multiple learning styles and intelligence types.

Gardner put forth the typology that there are up to nine different intelligence types: Interpersonal, Mathematical-logical, Musical, Spatial, Naturalistic, Intra-personal, Linguistic, Body-Kinesthetic and Existential (Gardner, 2006). Differing strengths and weaknesses in these intelligence types often result in challenges or inefficiencies when people learn or teach complex subject matter. For example: often subject experts (e.g. award winning physicists) might be more talented at doing math than teaching others how to do math (if they fall into the stereotype of someone who's Mathematical-logical intelligence might overshadow their Linguistic intelligence). Note that a pitfall of the intelligence type model is that people don't have just one, two or three of these types of intelligence types, and misunderstanding the model could easily lead to discrimination. However, one should understand that every person has a spread of the 9 intelligence types and some types might just be more dominant than others.

Hence to unlock more potential in the minds of all types of learners, we could be tempted to imagine that advances in data and information visualization techniques can improve upon the traditional methods of knowledge transfer using mainly text. We can also imagine that knowledge transfer would improve to increase insight into complex subject matter by sourcing the strengths of both top-down classification hierarchies and bottom up tagging networks. If this could be done in a visual and/or spatial way, various disciplines might be accessible to wider audiences through connecting multiple levels of thought. Ideally this would be achieved by engaging multiple intelligence types through text, visual, spatial, and other means (Victor, 2011). To explore some of this complexity, it would be necessary to revisit the topic of organizing information in hierarchical trees (such as library catalogues) and enter the world of network graphs.

2.5 Knowledge Trees and Networks

A key difference between libraries, museums and archives has historically been the way users interact with all their contents. Libraries allow the user to freely browse the shelves that are laid out according to the cataloguing hierarchy (and the architecture of the building), whereas museums and archives might traditionally be more restrictive. The structures of hierarchical catalogue trees and natural biological trees give further insight into each of their respective functions because their structures are due to the dual need for connectivity, and to optimize space.

Natural biological trees need to be connected through their roots to the soil for nutrients yet maximize their leaf surface area to absorb sunlight[5]. The analogy of branches spanning from a central structure allow the hierarchical tree structures in libraries to be easily navigable and traceable, like chapters and subsections in a book, but with some freedom to expand (to grow). Data trees in graphics or digital data visualization share these common traits; that the branches of classification structures often optimize the surface area used on a page or a screen.

An open hierarchical structure might not always supply a sufficiently complex framework to represent complex discipline specific knowledge (recall the Dewey Decimal system), but at least classification structures allow users to navigate content and make their own associations (links) between nuggets of information in multiple classification categories or disciplines. Hence knowledge networks can freely be created in the minds of users based on the rigid navigation structure and new research conclusions can be made unhindered (they are free to browse and make their own links).

A good example of this is Hegel and Luhmann’s card index system (called the “Zettelkasten”) to document the thoughts and connections they made while reading books and papers (Gitelman,2013). Using the card indexes, they were able to each spin their own web of knowledge on the leaves of the information trees they have read. The analogy of a spider’s web spun in the leaves and branches of a tree seems to be most an effective to illustrate how thoughts, ideas and knowledge can be generated from the hierarchical tree-like information structure from the shelves of a library and the chapters in a book (see Figure 1).

[5] Tree roots have a similar function of maximizing surface area, both to absorb nutrients from soil and to provide structural stability through friction in the ground.

Figure 1: A metaphor of Trees engulfed in webs (Watkins,2010 UK DFID)

Whereas Libraries are designed to be navigable and open, museums traditionally only allow users to view small subsets of their collection. The content is typically curated and displayed in ways to give context, with visitors often being oblivious to the number of items hidden away in storage. Similarly, archives are mostly accessible by appointment, with the archivist providing focus and access toward an area the researcher might be interested in through an appointment (Besser, 1987).

An increasing shift toward digitization and open data in some public and private institutions meant that the method of sharing knowledge has steadily been changing, not only for museums, but also online libraries and archives. With libraries, museums and archives opening their collections to more users, this allows more knowledge connections to be made (Fitzpatrick,2017). Hence it seems unsurprising that increasingly complex information visualizations have been applied to open data, since linked networks are necessary to visualize and navigate the underlying complexities. This is especially relevant where classification and the search for meaning becomes the responsibility of the user without guidance of an expert. Considering a knowledge visualization; it implies that ideally the information interface should provide a service of navigation, curation, and archivist support to provide customizable context and focus, so that someone can use the interface to teach themselves something new.

A primary reference repository for many types of network diagrams is located on the website *Visualcomplexity.com*, created by Manuel Lima (a Parsons Alumni), who beautifully highlights the changing representation of knowledge from trees toward networks. His books about mapping information and visualizing knowledge contain multiple examples throughout human history (Lima, 2013; Lima, 2014). One example included a classification tree of macroscopic species, which microbiologists have discovered to be cross-linked across its branches by a network of bacteria found across various types of animal species. This example of a bacterial network on animal species can also be extended to the idea of library catalogue systems and hashtags in media archives, where the catalogue system is analogous to the tree of animals and hashtags or indexes are analogous to bacteria.

The importance of both the top-down animal species classification and bottom-up bacteria network identification results in an interesting philosophical conundrum; that the world is indeed very complex and without our full understanding of the invisible links we often tend to classify (or possibly misclassify) the things we don't fully understand. These classifications might be completely valid, but often oversimplifies the existence of the items in the process if we think the hierarchical tree or catalogue is all there is to it. With increased prevalence of data visualization, a primary aspect is that we now have better tools to explore and communicate some of this complexity, which Manuel Lima calls an "emergent taxonomy"; the "Syntax of a new language".

Even though our limited understanding of the natural world (such as biology and physics) have often resulted in us missing the invisible things that make them work, human curiosity has always led us to discover a greater understanding of the world's underlying principles. We know that some bacteria can be good (even essential, such as species found in the digestive tract) or bad; those that cause disease. More importantly, we can also document these complex links more meticulously to develop technology to improve our lives.

Concerning the complexities in the field of Aerospace Engineering; which, just like Microbiology, contain various invisible phenomena that are interconnected in a complex way that experts and students still sometimes grapple with. Not only did scientists and engineers have to invent tools and methods to observe the invisible airflow around bodies flying through the atmosphere, they also had to solve microscopic problems with structural materials in order to manufacture metals that are strong and light enough to make powered flight of heavier-than-air machines possible (e.g. Aluminum, Alloy Steel and Titanium).

What makes the quest to reach the skies in the fields of Aeronautics and Aerospace different from Microbiology might be the scales and speeds that the natural laws occur at, but the hierarchical catalogue of information (or network of indexes and tags) of all these fields share very much the same fate of all previous human attempts to gain insight and meaning out of complexity; Splitting, Merging and Indexing subject matter to try and make sense out of the complexity.

Fortunately, we have the "Syntax of a new language" to explore, document and transfer these types of multi-disciplinary knowledge more efficiently than ever before. If effective non-linear knowledge documentation can be achieved, no longer would the complexities and knowledge of supersonic flight be limited to the minds of inaccessible experts or textbooks that are difficult to read, but it would be possible to share the webs of knowledge for anyone who wants to look, even if it is from a vantage point where the intricacies of airflow and airframes aren't needed to appreciate the human endeavor that got us into the skies and out of the earth's atmosphere.

3 How NACA became NASA

"It was a very lonely feeling when we began to run out of data" [when approaching the sound barrier]
—Walt Williams (NACA engineer)

3.1 NACA Research History

During World War I, the strategic advantages of the emerging aircraft technology became apparent and it prompted the United States to establish the National Advisory Committee for Aeronautics (NACA) in 1915. During the interwar period great advances were made by scientists and aviators and NACA promoted commercial and military aviation in the USA through research and development activities, as well as the building of wind tunnels that were crucial for advancing research on aircraft. By the time World War II broke out, NACA was well established as a leading aerospace research institution (Bilstein, 1989; Roland,1985). The NACA was spread across various Committees and facilities such as Langley, Ames, and Lewis Research Centers. Langley was the organization's original facility and its growth over time is illustrated by the number of employees that were based there (see Figure 2 below).

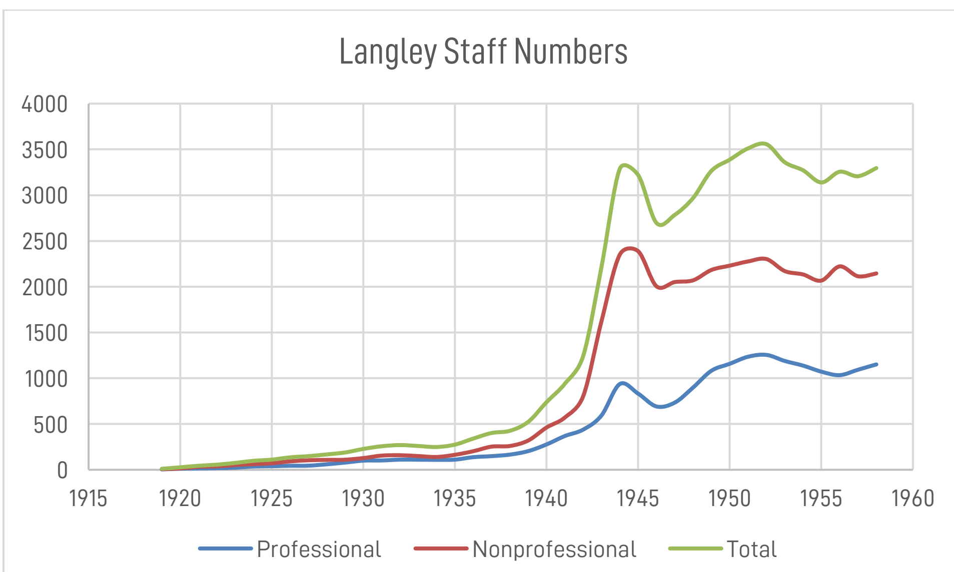


Figure 2: Langley Research Center Growth

Significant contributions were made to commercial aerospace in the early 1920's and aspects of safety and operational efficiency were a central focus. Aircraft like the Douglas DC-3 partly developed at NACA became the mainstay of the early airline industry and operated on international routes and even in service of foreign airlines. Other civilian aircraft development included personal aircraft like the Piper Cub, Stinson and Cessna. These foundational strong ties between the NACA and civilian industries would prove to be mutually beneficial to both well into the future. The sharing of knowledge bred success and increasingly connected people.

Most research at NACA consisted of developing and testing aerodynamic theories on topics such as wing outer shapes (airfoils) and aircraft nose shapes (e.g. cowlings). A key aspect of research was that theoretical underpinnings were confirmed with wind tunnel tests (i.e. scientific data). For example, NACA research studies included topics that showed aircraft landing gear contributed up to 40% of total drag (with significantly reduced fuel efficiency and performance), which resulted in industry wide adoption of retractable landing gear. Other research in the 1930's indicated that mounting engines in front of the wing instead of above and below the wing (using struts) also increased efficiency; design aspects we might take for granted today. Aircraft skin turbulence was also studied, and it was found that protruding rivets on the airframe skin also had significant effects on drag. Aspects like these produced knowledge to re-design and optimize aircraft development and construction methods to make airframes sleeker and more fuel efficient.

Advances such as laminar flow (as compared to turbulence) reduced drag and resulted in the wing designs of famous WW2 aircraft such as the P51-Mustang and Boeing's B-17 bomber. Both these aircraft were crucial for American victory in WW2. Often the research and optimization to reduce drag of equipment sticking out into the airstream could result in up to 10% improvement in speed just with minor tweaks, a margin that often-meant the difference between victory or defeat in combat.

During the war new types of aircraft engine development was also conducted after it became clear that European countries placed emphasis on liquid cooled engines with greater performance in speed at high altitudes, whereas American designs at the time consisted of air-cooled engines that were designed for airlines requiring better range and fuel efficiency. Post-war criticism that NACA faced was that they did not focus on jet engines during the war, however on the positive side, they did much research on swept wings which were later incorporated into much of Boeing's aircraft designs during the Jet age.

Post-war research continued on aerodynamics and airflow transitioning from subsonic to supersonic flight speeds, as well as topics aimed toward astronautics (i.e. getting into space). Collaboration with the Bell aircraft company and the Air Force led NACA to research on the first aircraft to break the sound barrier; the rocket powered Bell-X1. Approaching the sound barrier was a daunting task at the time and NACA engineers like Walt Williams grudgingly admitted it was "*a very lonely feeling as we began to run out of data*." (Bilstein,1989)

The sound barrier was broken in 1947 and continuous development allowed aircraft to fly beyond Mach 2, confronting issues like aerodynamic heating and the use of Titanium in aircraft structures; aspects that would be required later for space flight. Much of the data generated in these types of flight tests were fundamental to the design of modern airliners in the cruise speed regime just below the sound barrier. Multiple studies on the aerodynamics and materials of re-entry vehicles were also conducted, since these space vehicles had to survive the heat generated when entering the earth's atmosphere at high speeds.

Other research aspects included topics such as measurements of atmospheric turbulence and cabin pressurization, especially in light of catastrophic decompression that happened with the British De Havilland Comet airliner, which had 26 hull-loss accidents, with 13 fatal crashes and 426 casualties due to structural fatigue issues. Much of the successes at NACA propelled the organization increasingly into the public eye and made share the limelight with advances in spaceflight. Its aerodynamic and material science research proved invaluable and success led to more success.

NACA's public-facing image and its ties with civilian industrial and military research made it the obvious choice for Congress when the space race began in earnest with the launch of the Russian Sputnik satellite. The NACA agency's operational research model had to represent the American Space program with unique requirements on the dissemination of information and in 1958, The National Aeronautics and Space Administration (NASA) was created, and NACA officially became the center of NASA.

NASA has ever since been a primary source of knowledge about space and a proponent of issues in the public's eye, including issues such as diversity in engineering[6]. Even though much of current commercial and military aerospace knowledge have been

[6] During WWII, The NACA's workforce had a ⅓ female employees, fulfilling various administrative, computing, technical and engineering roles.

privatized or classified, the legacy of NACA and NASA remains through sharing knowledge and educating the next generation of scientists and engineers. The continued importance of safety in aerospace established through a culture of open knowledge is tightly linked with progress in the industry. Various modern aircraft designers and engineers still refer to the knowledge generated by NACA[7]. The following subsection will elaborate on how the NACA research archive is structured and optimized for visualization in a hybrid hierarchical tree and knowledge network.

3.2 The NACA Library Metadata Exploratory Analysis

Various types of documents were generated at the NACA, with the total research document collection amounting up to 13,801 research papers; taken from the Online Archive at University of North Texas[8] (UNT,2020). These documents consisted of various levels of formality, detail, and classification levels (Roland, Alex,1985). The distribution of each type and number of documents are shown in Figure 3. The "Technical Reports" were the most prestigious, but most of the documents consisted of less formal and more quickly distributed "Research Memorandums" and "Technical Notes".

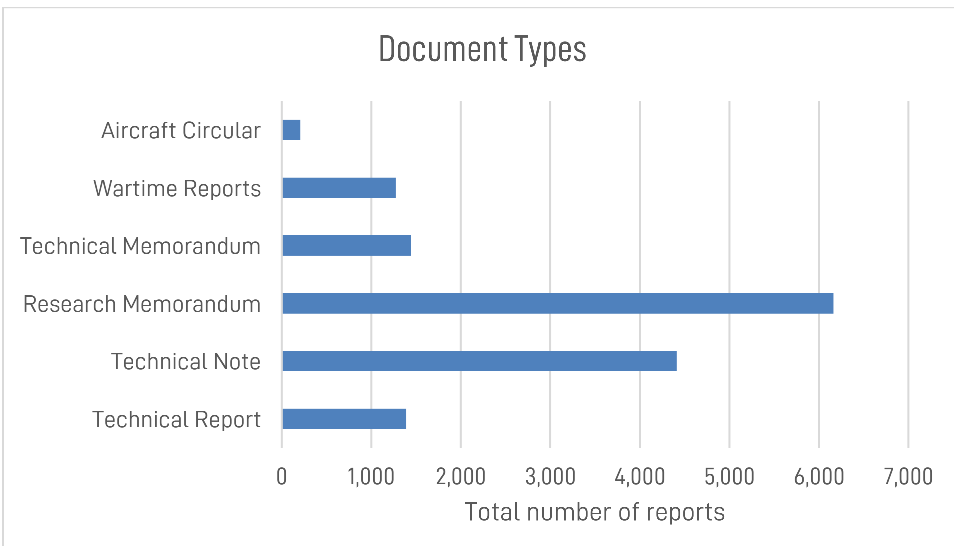


Figure 3: Numbers of NACA documents by Type

[7] For example, the NACA airfoil series (wing profiles) remain a core aspect of modern wing design.
[8] Various collections of NACA repositories exist (Smith,1994), but UNT was selected because of its comprehensive, clean database and API.

A wide variety of research topics were addressed, with the vast majority (59%) of these topics being in the field of Aerodynamics. Other major topics included Propulsion (14.7%), Structures (9.3%) and Materials (5.4%). A breakdown of the main subject headings are given in Table 1 and Figure 4, with two categories observed to contain two headings with almost negligible percentages (i.e. the Nomenclature and Technical Summaries). A full list of subject headings is given in Appendix C: Full Subject list.

	Subject Heading	Total report distribution (%)
1.	Aerodynamics	59
2.	Hydrodynamics	2.2
3.	Propulsion	14.7
4.	Structures	9.3
5.	Materials	5.4
6.	Meteorology	1
7.	Operating Problems	2.5
8.	Instruments	1.5
9.	Research Equipment Techniques	4.3
10.	Nomenclature	~ 0
11.	Bibliographies and Indexes	0.2
12.	Technical Summaries	~ 0

Table 1: Average Report Subject Distribution

The document subject distribution needs additional consideration because it informs the creation of a fully contextualized, yet balanced visualization of the collection. Because the Aerodynamic field dominates the number of records, focus will invariably be drawn to that area of the chart if all the records are shown at once. Such an unbalanced or lop-sided distribution might be sub-optimal and probably will not be aesthetically pleasing. Hence some strategies are required to appropriately space out the subjects.

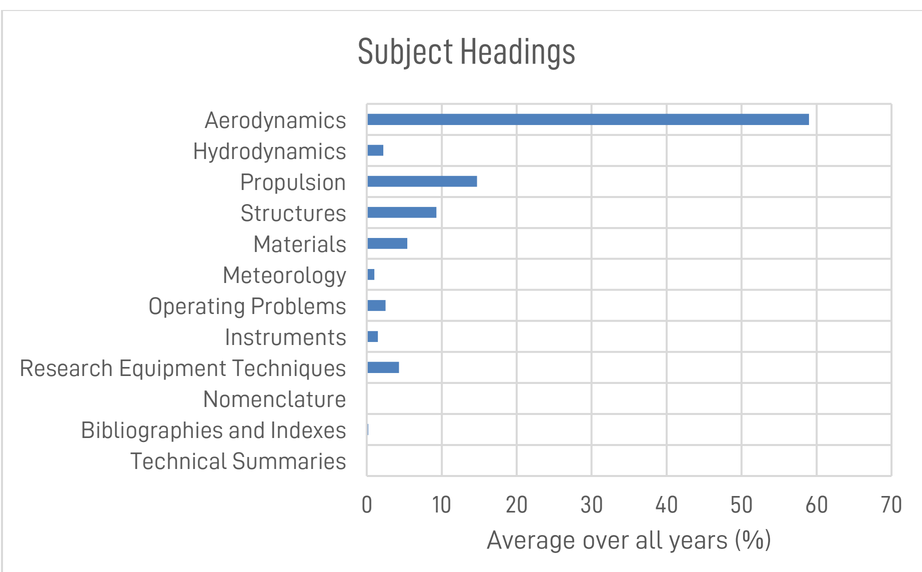


Figure 4: Average Report Subject Distribution (see Table 1)

It was also observed that subjects containing less documents often included more interesting and varying subject matter, hence less prevalent subject headings are worthwhile to have appropriate spatial representation on a chart. The visualization should therefore supply the user with enough focus for these subject headings to be placed into adequate context. Therefore, balance is required between the focus of less numerous specialized fields and the primary central focus of all the Aerodynamic documents. The subject heading order was changed accordingly (see below).

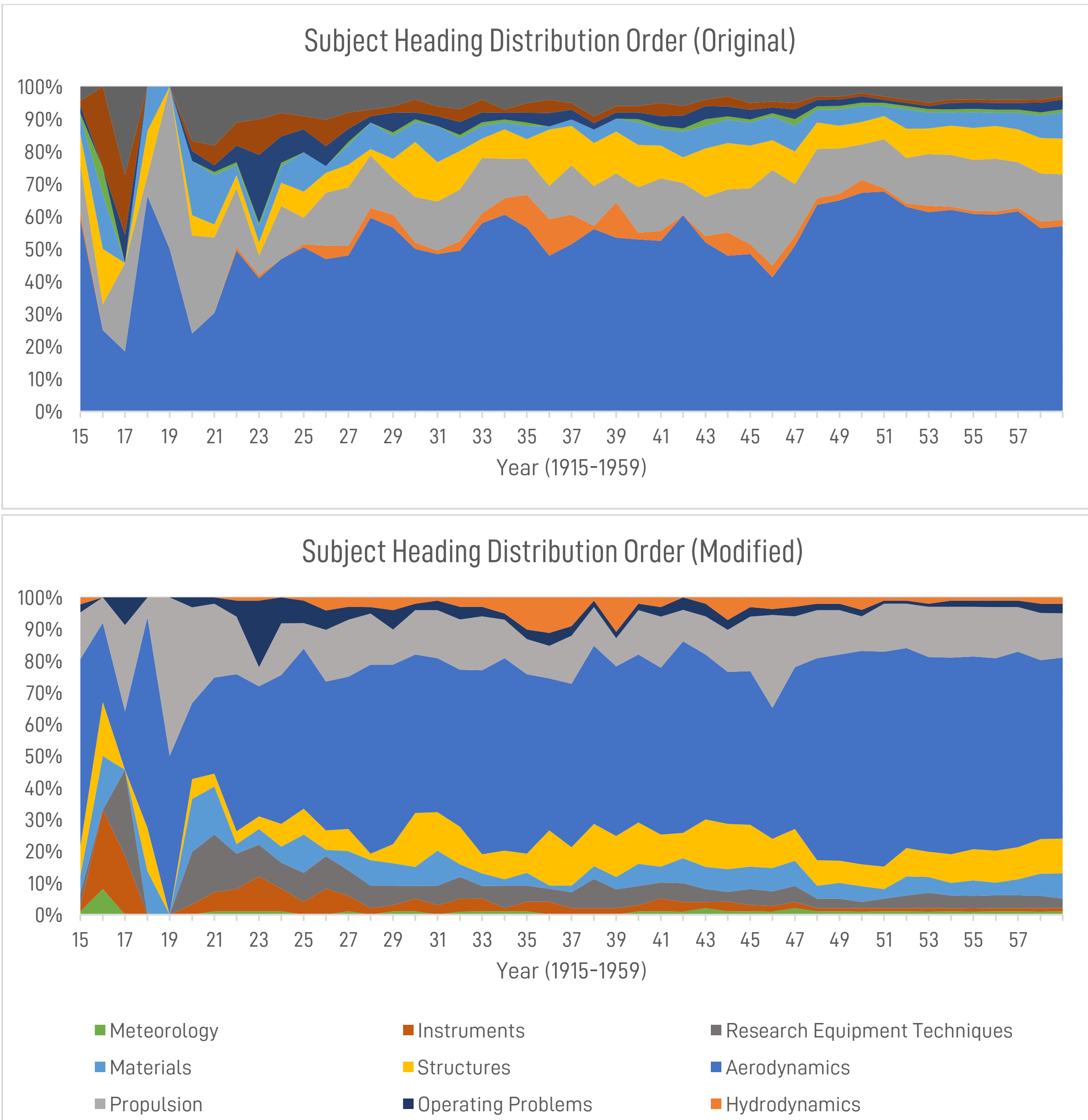


Figure 5: Subject Coverage by Year

To solve the problem of disparate topics competing for plot area, consideration was given to split the visualization into two main different main parts; the first plot area containing Aerodynamics (59%) and the second plot area containing all the other topics (41%). However, such a split could hamper the ability to elicit important and interesting multi-disciplinary connections such as Aeroelasticity (i.e. Aerodynamics and Elastic Structures) or Aero-propulsion.

These multi-disciplinary considerations are important because such connections form the basis of many modern Aerospace research fields. Subsequently it was decided to retain all subject headings and shape the visualization layout to make multi-disciplinary connections possible, without compromising on competing focal points. An attempt was made to minimize the "raggedness" over time i.e. to even out the overall distribution of subject boundaries over years, without interfering with the data itself (The quantities of the Report Distributions are given in Appendix B: NACA Report Distribution). When the re-ordering was conducted, it meant that the original NACA index order changed. However, re-ordering is deemed a small compromise, since subject categories are artificially numbered irrespectively. Re-numbering could be explore in future, but it was decided to just refrain from showing the original NACA index number to avoid unnecessary confusion. This choice might be revisited in a future version.

A subject heading word map was also generated, but is not included in the final online visualization, and is only shown here for additional explanation. This word map was created as part of an assignment for the Typography and Visual design course. The map also shows deeper insight into the underlying content (see Figure 6). The position of the words are mapped approximately to the same position as the final visualization, with the horizontal coordinate corresponding the subject headings and the vertical coordinate mapping corresponding to the year of publication. This mapping is switched from the preceding figures that shows the original and modified subject distributions (Figure 5), but the same as the final version; i.e. the word map and the final visualization will have a vertical timeline, with subject headings spaced from left to right (more on this later).

The word map format is compressed to a landscape aspect ratio, even though the timeline of the final visualization will require an elongated aspect ratio due to space requirements. The number of documents in a specific subject area is mapped to the text size. Note that "Wing Theory" dominates and is surrounded with various other subject types. This word map indicates what type of trends to expect in the final visualization. A full list of all subjects levels are supplied in Appendix C: Full Subject list, but only the third tier of the hierarchy was used in the word map to avoid unnecessary clutter (i.e. 1. to 1.2.3.) with lower 4th and 5th levels (1.2.3.4. and 1.2.3.4.5) aggregated into the 3rd (1.2.3.).

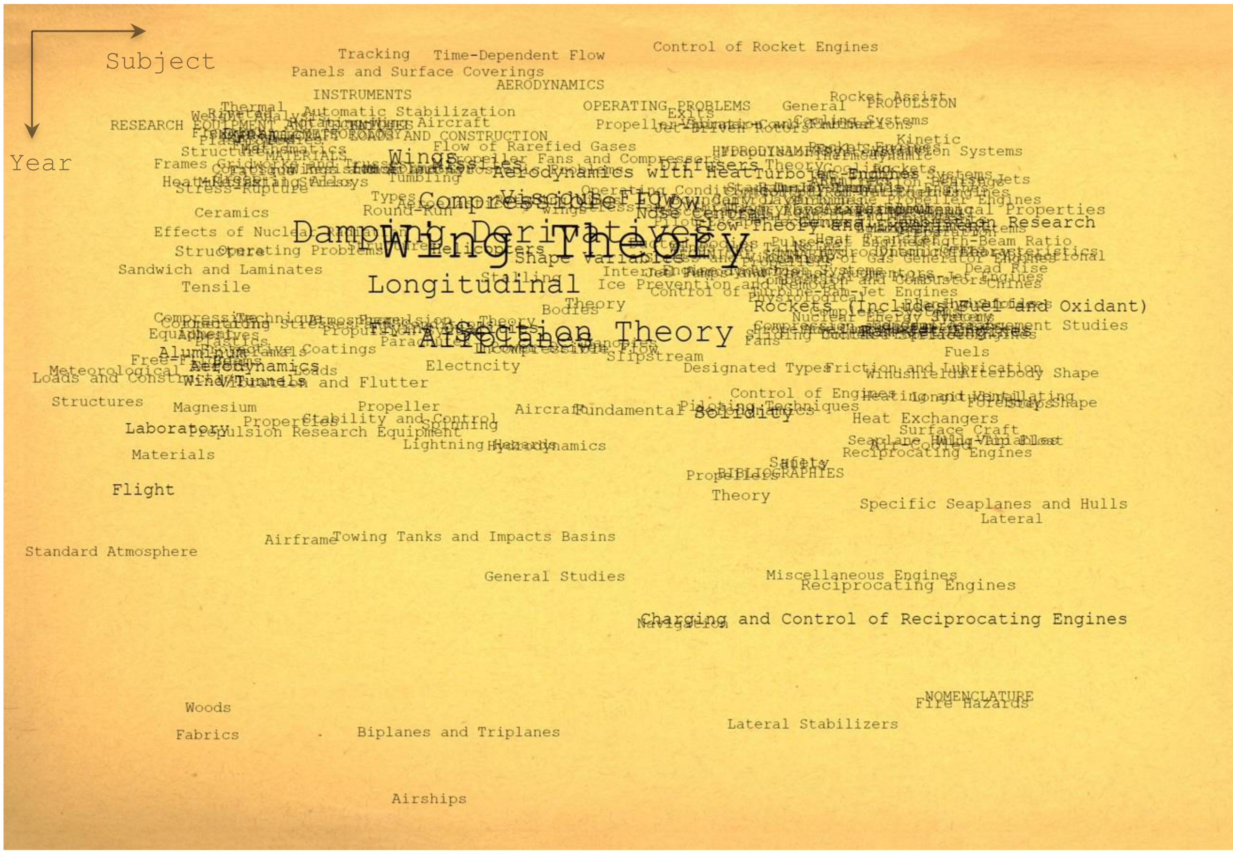


Figure 6: Subject Heading Word Map (horizontal) vs. Time (vertical)

Finally, considering the possible cross-plotting of various multi-disciplinary connections; it seems worthwhile to recall the generalizability of the proposed layout technique toward complex multi-disciplinary subjects (as discussed in previous sections) with regards to trees and networks. Often these multi-disciplinary topics are recognized by dual words like “Aerospace” or “Aeropropulsion”[9]. Although there are obvious fundamental differences in the subject matter of such varying disciplines, structuring complex multi-disciplinary subject matter in standardized data structures seem advantageous. The generalizability of visualization structures for different types of multi-disciplinary topic visualizations is beyond the scope of this thesis, but it is worthwhile mentioning that using the well-documented and categorized Aerospace research collection of NACA could help shape ways of visualizing other multi-disciplinary complexity as well.

---

[9] Considering non-aerospace fields, these and might include other complex subjects such as “Technopolitics” or “Socioeconomics”.

3.3 Metadata Structure

Something that seems to differentiate the NACA research library from many other complex research areas is the meticulous cataloguing and classification that was conducted by the NACA *Division of Research Information* to keep track of main subject headings (NACA, 1949; 1952; 1953; 1954; 1955; 1956; 1957; 1959). A whole section of the original NACA document collection is dedicated to Bibliographies and Indexes (see Table 1) and some of these NACA index catalogues are subsequently used as a primary input to add "top-down" structure hierarchical Table of Contents to the visualization data structure. An extract of the index catalogues is shown in Figure 7, with some of the detail shown in Figure 8. An example of such a paper is shown in Figure 9.

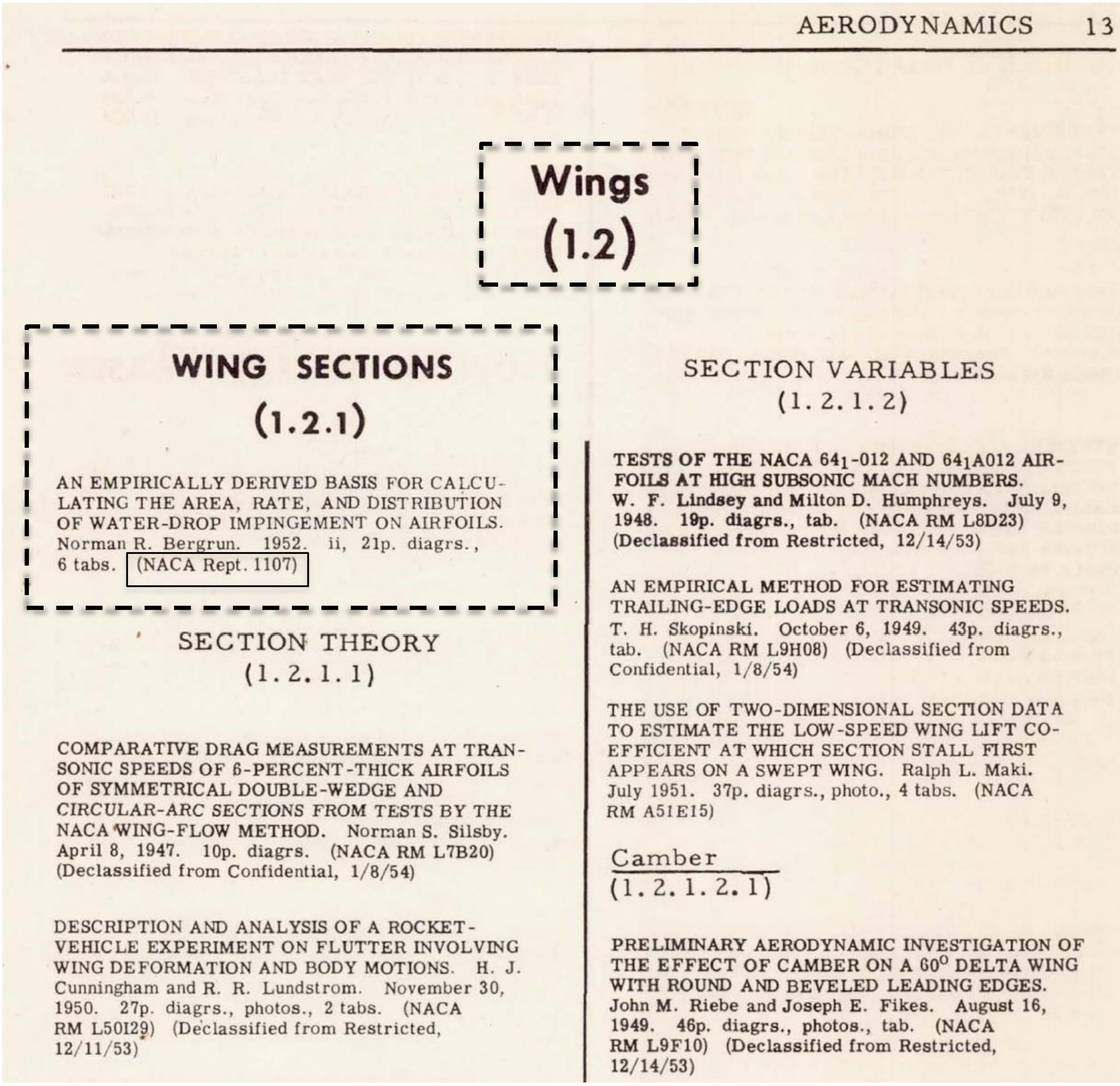
Using these indexes, the document metadata was extracted from the digital Object Character Recognized (OCR) text using Python/Pandas scripts, cleaned, and linked to each other using the report numbers (e.g. NACA Rept. 1107). The OCR text was obtained from online sources using the NASA *Technical Reports Server* (NTRS; ntrs.nasa.gov) as well as the *Waybackmachine (archive.org)*. The digital text was then cleaned using a series of scripts each generating console outputs in the Python Spyder IDE and writing text files to disk in various subsequent steps to ensure data integrity.

Fortunately, the original NACA indexes were mostly high quality with individual document records well-spaced out on each page, with scanned documents mostly being high quality as well. Some OCR artifacts remained, but most challenges were due to excessive whitespace and line spacing which made it more difficult to group individual document records. There were also two different formatting conventions in the index catalogues: the first being lowercase from 1915-1951 and the second being uppercase from 1951-1959 with report numbers in brackets. Two scripts had to be written to handle the different formats. Some minor fields were also found to be obsolete going from 1949 to 1951. These fields were not remapped but should be ideally be remapped in future to be more accurately allocated in the hierarchy.

Without the inclusion of individual report numbers in each Table of Contents entry (or "leaf"), the mapping into the Hierarchy would have been a great challenge. The report numbers were therefore a primary way to classify the reports into the correct subject fields and extra attention was given to make sure the report numbers OCR text are clean and free from artifacts using Python terminal logs. If OCR report numbers were not clean, it would imply that specific subject heading ("copy of a leaf") will not be used to calculate the averaged lateral position of the record on screen. Fortunately the cases for bad OCR report numbers were rare.

Subject Heading Number	Subject Heading Outline	Page
1	AERODYNAMICS	1-181
1.1	Fundamental Aerodynamics	3-26
1.1.1	Incompressible Flow	3-4
1.1.2	Compressible Flow	4-5
1.1.2.1	Subsonic Flow	5-6
1.1.2.2	Mixed Flow	6-7
1.1.2.3	Supersonic Flow	7-12
1.1.3	Viscous Flow	13-14
1.1.3.1	Laminar Flow	14-16
1.1.3.2	Turbulent Flow	16-18
1.1.3.3	Jet Mixing	18-20
1.1.4	Aerodynamics With Heat	20-21
1.1.4.1	Heating	21
1.1.4.2	Heat Transfer	22-24
1.1.4.3	Additions of Heat	25
1.1.5	Flow of Rarefied Gases	25
1.1.5.1	Slip Flow	25
1.1.5.2	Free Molecule Flow	25-26
1.1.6	Time-Dependent Flow	26
1.2	Wings	27-69
1.2.1	Wing Sections	27
1.2.1.1	Section Theory	27
1.2.1.2	Section Variables	27
1.2.1.2.1	Camber	27-28
1.2.1.2.2	Thickness	28
1.2.1.2.3	Thickness Distribution	28-29
*1.2.1.2.4	Inlets and Exits	
1.2.1.2.5	Surface Conditions	29
1.2.1.3	Designated Profiles	29
1.2.1.4	High-Lift Devices	29
1.2.1.4.1	Plain Flaps	30
1.2.1.4.2	Split Flaps	30
1.2.1.4.3	Slotted Flaps	30
1.2.1.4.4	Leading-Edge Flaps	30
1.2.1.4.5	Slots and Slats	30
1.2.1.5	Controls	30
1.2.1.5.1	Flap Type	30-31
1.2.1.5.2	Spoilers	31
1.2.1.6	Boundary Layer	31
1.2.1.6.1	Characteristics	31
1.2.1.6.2	Control	32
1.2.1.7	Reynolds Number Effects	32
1.2.1.8	Mach Number Effects	32-33
1.2.1.9	Wake	33
1.2.2	Complete Wings	33
1.2.2.1	Wing Theory	33-34
1.2.2.2	Wing Variables	34-35
1.2.2.2.1	Profiles	35-37
1.2.2.2.2	Aspect Ratio	37-39
1.2.2.2.3	Sweep	39-44
1.2.2.2.4	Taper and Twist	44-45
1.2.2.2.5	Inlets and Exits	45
1.2.2.2.6	Surface Conditions	45
1.2.2.2.7	Dihedral	45-46
1.2.2.3	High-Lift Devices	46-47
1.2.2.3.1	Trailing-Edge Flaps	47-49
1.2.2.3.2	Slots and Slats	49
1.2.2.3.3	Leading-Edge Flaps	49-50
1.2.2.4	Controls	50-51
1.2.2.4.1	Flap Type	51-53
1.2.2.4.2	Spoilers	53-54
1.2.2.4.3	All-Movable	54-55
1.2.2.5	Reynolds Number Effects	55-57
1.2.2.6	Mach Number Effects	57-67
1.2.2.7	Wake	67-68
1.2.2.8	Boundary Layer	68
1.2.2.8.1	Characteristics	68
1.2.2.8.2	Control	69
1.3	Bodies	70-79
1.3.1	Theory	71
1.3.2	Shape Variables	72-73
1.3.2.1	Fineness Ratio	73-74
1.3.2.2	Cross Section	74
1.3.2.3	Thickness Distribution	74-75
1.3.2.4	Surface Conditions	75
1.3.2.5	Protuberances	75-76
1.3.3	Canopies	76
1.3.4	Ducted Bodies	76
1.3.4.1	Nose Shape	77
1.3.4.2	Tail Shape	77-78
1.3.4.3	Side Inlets	78
1.3.4.4	Side Exits	79
*1.3.5	Hulls	
1.4	Internal Aerodynamics	80-97
1.4.1	Air Inlets	80
1.4.1.1	Nose, Central	80-81
1.4.1.1.1	Propeller-Spinner-Cowl Combinations	81
1.4.1.1.2	Subsonic	81
1.4.1.1.3	Supersonic	81-82
1.4.1.2	Nose, Annular	82-83
1.4.1.3	Wing-Leading-Edge	83
1.4.1.4	Side	83-84
1.4.1.4.1	Scoops	84-85
*1.4.1.4.2	Submerged	
1.4.2	Ducts	85-86
1.4.2.1	Diffusers	86
1.4.2.1.1	Subsonic	86-87
1.4.2.1.2	Supersonic	87-89
1.4.2.2	Nozzles	89-90
1.4.2.3	Pipes	90-91
1.4.2.4	Bends	91
1.4.3	Exits	91-93
1.4.4	Jet Pumps and Thrust Augmenters	93
1.4.5	Cascades	93-94
1.4.5.1	Theory	94
1.4.5.2	Experiment	95
1.4.6	Fans	95-96
1.4.7	Boundary Layer	96
1.4.7.1	Characteristics	96-97
1.4.7.2	Control	97

Figure 7: Extract from NACA Index of Technical Publications (NACA, 1954)



AERODYNAMICS 13

Wings

(1.2)

WING SECTIONS

(1.2.1)

AN EMPIRICALLY DERIVED BASIS FOR CALCULATING THE AREA, RATE, AND DISTRIBUTION OF WATER-DROP IMPINGEMENT ON AIRFOILS. Norman R. Bergrun. 1952. ii, 21p. diagrs., 6 tabs. (NACA Rept. 1107)

SECTION THEORY

(1.2.1.1)

COMPARATIVE DRAG MEASUREMENTS AT TRANSONIC SPEEDS OF 6-PERCENT-THICK AIRFOILS OF SYMMETRICAL DOUBLE-WEDGE AND CIRCULAR-ARC SECTIONS FROM TESTS BY THE NACA WING-FLOW METHOD. Norman S. Silsby. April 8, 1947. 10p. diagrs. (NACA RM L7B20) (Declassified from Confidential, 1/8/54)

DESCRIPTION AND ANALYSIS OF A ROCKET-VEHICLE EXPERIMENT ON FLUTTER INVOLVING WING DEFORMATION AND BODY MOTIONS. H. J. Cunningham and R. R. Lundstrom. November 30, 1950. 27p. diagrs., photos., 2 tabs. (NACA RM L50I29) (Declassified from Restricted, 12/11/53)

SECTION VARIABLES

(1.2.1.2)

TESTS OF THE NACA 64_1-012 AND 64_1A012 AIRFOILS AT HIGH SUBSONIC MACH NUMBERS. W. F. Lindsey and Milton D. Humphreys. July 9, 1948. 19p. diagrs., tab. (NACA RM L8D23) (Declassified from Restricted, 12/14/53)

AN EMPIRICAL METHOD FOR ESTIMATING TRAILING-EDGE LOADS AT TRANSONIC SPEEDS. T. H. Skopinski. October 6, 1949. 43p. diagrs., tab. (NACA RM L9H08) (Declassified from Confidential, 1/8/54)

THE USE OF TWO-DIMENSIONAL SECTION DATA TO ESTIMATE THE LOW-SPEED WING LIFT COEFFICIENT AT WHICH SECTION STALL FIRST APPEARS ON A SWEPT WING. Ralph L. Maki. July 1951. 37p. diagrs., photo., 4 tabs. (NACA RM A51E15)

Camber

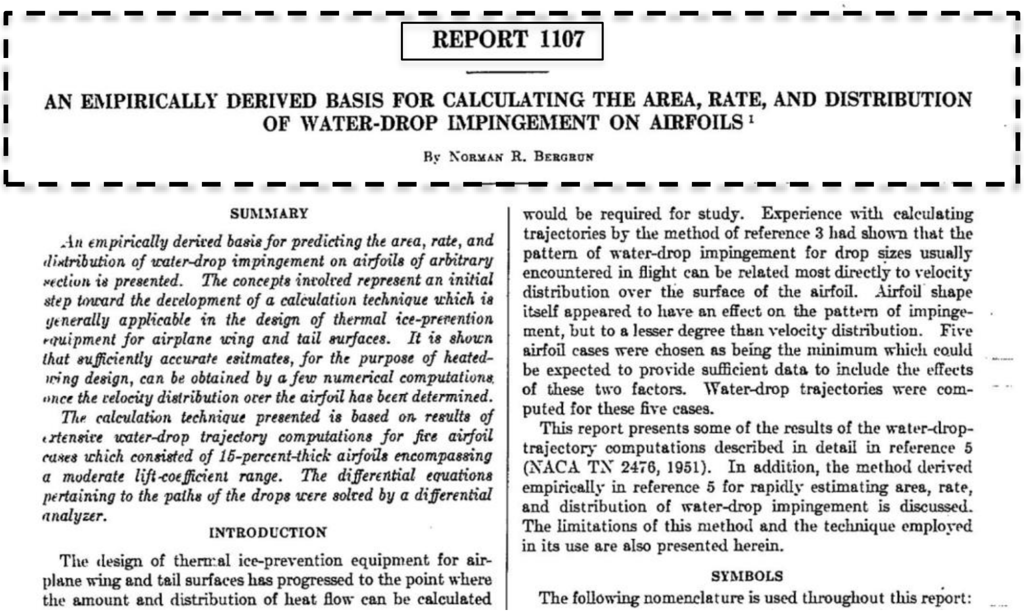
(1.2.1.2.1)

PRELIMINARY AERODYNAMIC INVESTIGATION OF THE EFFECT OF CAMBER ON A 60° DELTA WING WITH ROUND AND BEVELED LEADING EDGES. John M. Riebe and Joseph E. Fikes. August 16, 1949. 46p. diagrs., photos., tab. (NACA RM L9F10) (Declassified from Restricted, 12/14/53)

Figure 8: NACA Index Record

AERODYNAMICS 13 Wings (1.2) WING SECTIONS (1.2.1) AN EMPIRICALLY DERIVED BASIS FOR CALCU- LATING THE AREA, RATE, AND DISTRIBUTION OF WATER-DROP IMPINGEMENT ON AIRFOILS. Norman R. Bergrun. 1952. ii, 21p. diagrs. , 6 tabs. (**NACA Rept. 1107**) SECTION THEORY (1.2.1.1) COMPARATIVE DRAG MEASUREMENTS AT TRAN- SONIC SPEEDS OF 6-PERCENT -THICK AIRFOILS OF SYMMETRICAL DOUBLE -WEDGE AND CIRCULAR-ARC SECTIONS FROM TESTS BY THE NACA WING- FLOW METHOD. Norman S. Silsby. April 8, 1947. lOp. diagrs. (NACA RM L7B20) (Declassified from Confidential, 1/8/54) DESCRIPTION AND ANALYSIS OF A ROCKET- VEHICLE EXPERIMENT ON FLUTTER INVOLVING WING DEFORMATION AND BODY MOTIONS. H. J. Cunningham and R. R. Lundstrom. November 30, 1950. 27p. diagrs., photos., 2 tabs. (NACA RM L50I29) (Declassified from Restricted, 12/11/53)	SECTION VARIABLES (1. 2.1.2) TESTS OF THE NACA 64i-012'AND 64iA012 AIR- FOIU AT mOI SUBSONIC MACH NUMBERS. W. F. LiiKteey and Milton D. Humphreys. July 9, 1948. 19p. diagrs., tab. (NACA RM L8D23) (Declassified from Restricted, 12/14/53) AN EMPIRICAL METHOD FOR ESTIMATING TRAILING-EDGE LOADS AT TRANSONIC SPEEDS. T. H. Skopinskl. October 6, 1949. 43p. diagrs., tab. (NACA RM L9H08) (Declassified from Confidential, 1/8/54) THE USE OF TWO-DIMENSIONAL SECTION DATA TO ESTIMATE THE LOW-SPEED WING UFT CO- EFFICIENT AT WHICH SECTION STALL HRST APPEARS ON A SWEPT WING. Ralph L. Maki. July 1951. 37p. diagrs., photo., 4 tabs. (NACA RM A51E15) Camber (1.2.1.2.1) PRELIMINARY AERODYNAMIC INVESTIGATION OF THE EFFECT OF CAMBER ON A 60° DELTA WING WITH ROUND AND BEVELED LEADING EDGES. John M. Riebe and Joseph E. Flkes. August 16, 1949. 46p. diagrs., photos., tab. (NACA RM L9F10) (Declassified from Restricted, 12/14/53)

Table 2: NACA Index Record OCR Text



REPORT 1107

AN EMPIRICALLY DERIVED BASIS FOR CALCULATING THE AREA, RATE, AND DISTRIBUTION OF WATER-DROP IMPINGEMENT ON AIRFOILS [1]

By NORMAN R. BERGRUN

SUMMARY

An empirically derived basis for predicting the area, rate, and distribution of water-drop impingement on airfoils of arbitrary section is presented. The concepts involved represent an initial step toward the development of a calculation technique which is generally applicable in the design of thermal ice-prevention equipment for airplane wing and tail surfaces. It is shown that sufficiently accurate esitmates, for the purpose of heated-wing design, can be obtained by a few numerical computations once the velocity distribution over the airfoil has been determined.

The calculation technique presented is based on results of extensive water-drop trajectory computations for five airfoil cases which consisted of 15-percent-thick airfoils encompassing a moderate lift-coefficient range. The differential equations pertaining to the paths of the drops were solved by a differential analyzer.

INTRODUCTION

The design of thermal ice-prevention equipment for airplane wing and tail surfaces has progressed to the point where the amount and distribution of heat flow can be calculated would be required for study. Experience with calculating trajectories by the method of reference 3 had shown that the pattern of water-drop impingement for drop sizes usually encountered in flight can be related most directly to velocity distribution over the surface of the airfoil. Airfoil shape itself appeared to have an effect on the pattern of impingement, but to a lesser degree than velocity distribution. Five airfoil cases were chosen as being the minimum which could be expected to provide sufficient data to include the effects of these two factors. Water-drop trajectories were computed for these five cases.

This report presents some of the results of the water-drop-trajectory computations described in detail in reference 5 (NACA TN 2476, 1951). In addition, the method derived empirically in reference 5 for rapidly estimating area, rate, and distribution of water-drop impingement is discussed. The limitations of this method and the technique employed in its use are also presented herein.

SYMBOLS

The following nomenclature is used throughout this report:

Figure 9: Document Example (NACA-TR-1107)

If the whole collection data structure is likened to a hierarchical tree, abovementioned catalogues and links will form the branches and designate the positions of the leaves, which consist of individual papers. However, these leaves could occur on more than one branch at the same time, especially when the topics addressed in each paper is multi-disciplinary, i.e. a document might be classified as Aerodynamics and Structures at the same time (e.g. Aeroelasticity). As an example: For the document NACA Rept. 1107 ("leaf") that was highlighted above; this document is also classified into the subject headings of Meteorology (6.2 Ice Formation) as well as Operating problems (7.3.3 Wings and Tails; see Figure 10 and Figure 11 below).

Ice Formation

(6.2)

AN EMPIRICALLY DERIVED BASIS FOR CALCULATING THE AREA, RATE, AND DISTRIBUTION OF WATER-DROP IMPINGEMENT ON AIRFOILS. Norman R. Bergrun. 1952. ii, 21p. diagrs., 6 tabs. (NACA Rept. 1107)

MAXIMUM EVAPORATION RATES OF WATER DROPLETS APPROACHING OBSTACLES IN THE ATMOSPHERE UNDER ICING CONDITIONS. Herman H. Lowell. October 1953. 56p. diagrs., 3 tabs. (NACA TN 3024)

Figure 10: One document (leaf) in many subject headings (branches)

WINGS AND TAILS

(7.3.3)

AN EMPIRICALLY DERIVED BASIS FOR CALCULATING THE AREA, RATE, AND DISTRIBUTION OF WATER-DROP IMPINGEMENT ON AIRFOILS Norman R. Bergrun. 1952. ii, 21p. diagrs., 6 tabs. (NACA Rept. 1107)

IMPINGEMENT OF WATER DROPLETS ON AN ELLIPSOID WITH FINENESS RATIO 5 IN AXISYMMETRIC FLOW. Robert G. Dorsch, Rinaldo J. Brun and John L. Gregg. March 1954. 50p. diagrs., tab. (NACA TN 3099)

IMPINGEMENT OF WATER DROPLETS ON AN ELLIPSOID WITH FINENESS RATIO 10 IN AXISYMMETRIC FLOW. Rinaldo J. Brun and Robert G. Dorsch. May 1954. 37p. diagrs., tab. (NACA TN 3147)

Figure 11: One document (leaf) in many subject headings (branches)

Abovementioned multi-disciplinary hierarchical layouts could be problematic, because placing this document as a leaf into the tree essentially means that there needs to be copy of the document on both branches it belongs to. Using the catalogue of the NACA subject headings indicated that a large majority of documents were placed in three to four subject headings. If one wanted to visualize these, redundant leaves would need to be added in all these branches and would result more than 40,000 data points (for 13,800 papers). This high number illustrates how meticulous the NACA have categorized the documents into multiple different "branches". Storing and showing all these duplications would add an unnecessary burden to plot and to store. Hence, the average position mapping was used by calculating the mean position of multiple placements such as these.

In addition to the original NACA Indexes and Bibliographies, the University of North Texas NACA archive has metadata related to the typical primary identifiers such as Report number, Title and Author, as well as keywords and descriptions (Abstracts). The UNT API allows access to all this data free of charge and the data is publicly available, with the only condition specified their site being that large queries be done after hours so that an unnecessary burden isn't placed on their server resources.

This XML file does not appear to have any style information associated with it. The document tree is shown below.


```
▼<oai_dc:dc xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
 http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
  ▼<dc:title>
     An empirically derived basis for calculating the area, rate, and distribution of water-drop impingement on airfoils
   </dc:title>
   <dc:title>NACA Technical Reports</dc:title>
   <dc:creator>Bergrun, Norman R.</dc:creator>
   <dc:contributor>Ames Aeronautical Laboratory (U.S.)</dc:contributor>
   <dc:date>1951-05-08</dc:date>
   <dc:language>English</dc:language>
  ▼<dc:description>
     From Summary: "An empirically derived basis for predicting the area, rate, and distribution of water-drop impingement on
     airfoils of arbitrary section is presented. The concepts involved represent an initial step toward the development of a
     calculation technique which is generally applicable to the design of thermal ice-prevention equipment for airplane wing
     and tail surfaces. It is shown that sufficiently accurate estimates, for the purpose of heated-wing design, can be
     obtained by a few numerical computations once the velocity distribution over the airfoil has been determined. The
     calculation technique presented is based on results of extensive water-drop trajectory computations for five airfoil
     cases which consisted of 15-percent-thick airfoils encompassing a moderate lift-coefficient range."
   </dc:description>
   <dc:subject>meteorology</dc:subject>
   <dc:subject>water-drop impingement</dc:subject>
   <dc:rights>Public</dc:rights>
   <dc:rights>Public Domain</dc:rights>
  ▼<dc:rights>
     No Copyright, Unclassified, Unlimited, Publicly available
   </dc:rights>
   <dc:type>Report</dc:type>
   <dc:format>1079 - 1099 p.</dc:format>
   <dc:format>Text</dc:format>
   <dc:identifier>local-cont-no: 93R21433</dc:identifier>
   <dc:identifier>url: http://hdl.handle.net/2060/19930092143</dc:identifier>
   <dc:identifier>rep-no: NACA-TR-1107</dc:identifier>
   <dc:identifier>casi: 19930092143</dc:identifier>
  ▼<dc:identifier>
     https://digital.library.unt.edu/ark:/67531/metadc60466/
   </dc:identifier>
   <dc:identifier>ark: ark:/67531/metadc60466</dc:identifier>
 </oai_dc:dc>
```


Figure 12: UNT API XML OUTPUT

The primary method to extract the data from the API is using XML lists (see Figure 12). JSON outputs are also available, but only for single records at a time. Even though JSON outputs would be preferred above XML, the single record query return is not ideal for downloading the whole library metadata, since 13,801 individual server requests would have to be made and would take hours to complete. Subsequently the XML format was used to obtain the data and was parsed with an online parsing tool to convert into a JSON structure. The structure was not as simple as API to JSON would have been, but individual record metadata were then extracted into a single data file that was reconfigured and restructured to in Pandas/Python Jupyter notebooks for further wrangling.

Even though *keywords* and *abstracts* are common to most types research papers online, what distinguishes this NACA metadata structure is that both a Hierarchical *Table of Contents* metadata (based on OCR extracted indexes), as well as the more common *keyword indexes* (from the UNT library) are available for the entire collection (albeit with additional processing and OCR text data wrangling).

Therefore, the entire research collection now has a "front-of-the book" Table of Contents and the "rear-of-the book" index/keyword collection that connects information both in a tree-like and network-like structure. It seems rare that both these types of data are "easily" available for a full collection of research papers, even though additional research would be required to find similar collections that have both a well-classified hierarchical and network metadata available.

The interesting thing about this duality of human-driven classification into hierarchies and networks emerge from the complexity is that both contribute to the navigation and contextual meaning of the individual interconnected documents. On the one side there is a search for order out of the complexity by grouping documents into subjects, on the other side there is the underlying connections in between documents by the same author(s) or and/or keywords. There might be some overlap between subjects and keywords, but there remains a tension between the two. Recall that the simplest analogy made before is a tree spun with a web (recall Figure 1). A visual summary of the data preparation process is shown below in Figure 13.

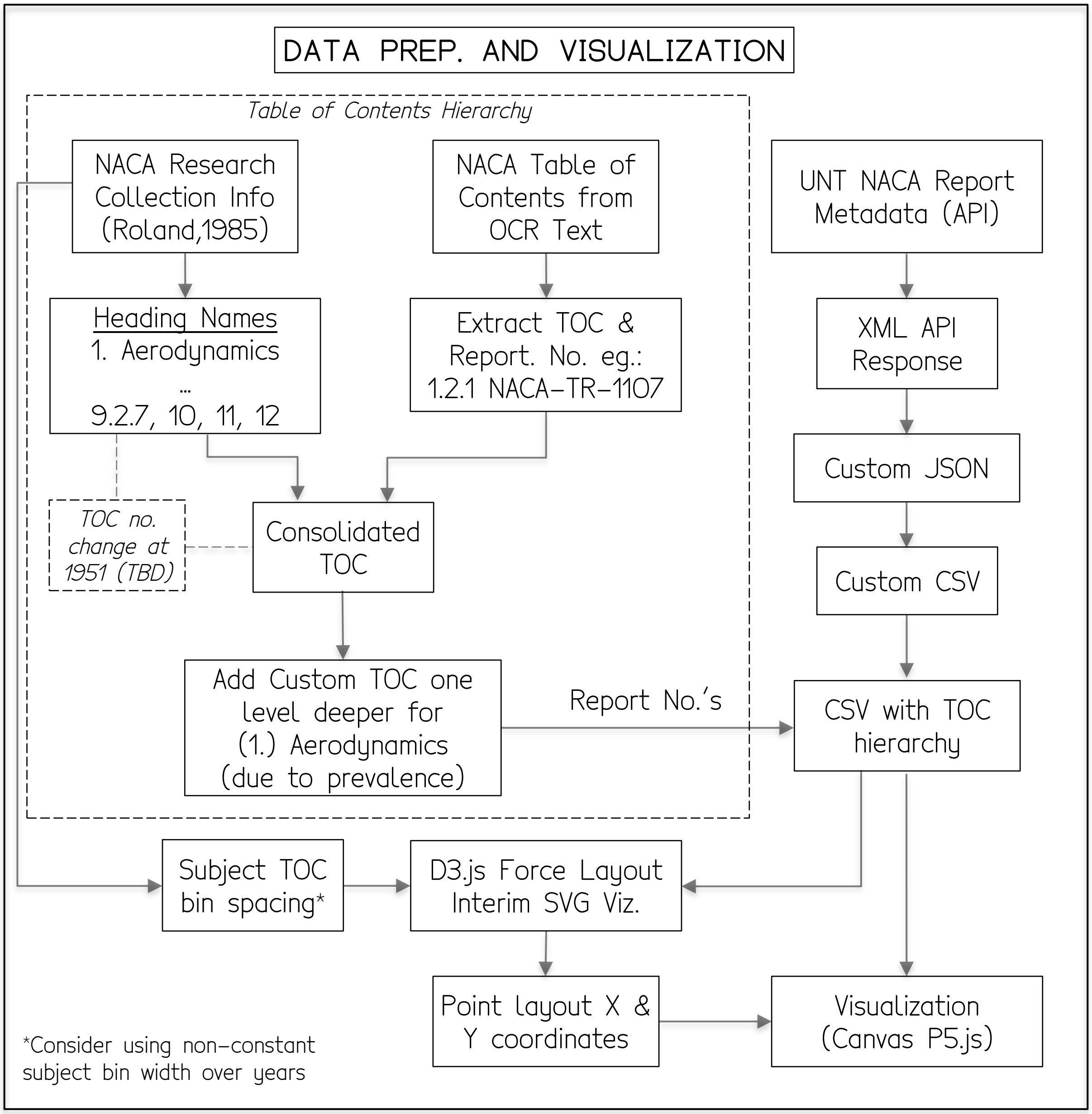


Figure 13: Data Wrangling and Visualization Process Logic

3.4 Visual mapping

As mentioned before, one challenge of visualizing the records in both a tree and a web is that each might be duplicated, with linked connections to other related documents (e.g. via authors and keywords). This might be useful for some aspects, but special encoding could be lost if there aren't enough dimensions to show all the different types of connections, or rather; the relational structure might not be clear.

Some of this relational structure can be solved using network layout tools such as the D3 force layout tool to visually cluster related nodes by specifying the "level of relatedness" with a force attraction value and limiting connected link distances to avoid unnecessary clutter. The only problem is that this strategy easily negates an externally imposed hierarchical structure based on human classification.

When a third aspect of time is added, this makes the visualization even more challenging. Visualizing networks over time while extracting meaning can be especially challenging for a static visualization. With the NACA collection being an historical archive, it would be advantageous to also have the aspect of time as part of a primary dimension of the visualization, because it will likely show some key developments and how these affected subsequent technologies.

Another useful dimension would be the hierarchical classification of the NACA subject heading table of contents. Like the Dewey Decimal system and Library of Congress numbers, these categories have their limitations, but even so, some utility remains in having a structured navigable dimension to aid search functionality. A third dimension of relevance and a fourth dimension of interconnectivity and/or relation would be advantageous. Alas, the number of dimensions is soon running out for standard layouts. Subsequently additional dimensions need to be accommodated by using either collapsed primary dimensions and/or additional retinal variables (according to Bertin).

Collapsed dimensions could be constructed to simulate a higher order dimension with a lower dimension, such as placing a 3D object on a 2D plane and adding a 3D effect (i.e. 2.5D). This approach might be a compromise but could free up a dimension for adding another layer of meaning, if the compromise still allows the dimension to fulfill its intended communicative role.

A question that comes up is the possibility of collapsing a 2D plane to a "1.5D line". If it can be done for a 3D space, surely it can be done for a 2D plane. In fact, this is exactly what is done by listing a table of contents in a book. The hierarchy of the tree is collapsed to lines of text, spaced by the line spacing of the typographic elements. One could argue that the line heights between text lines become encoded with the different categories of classification for individual subject headings. Hence the idea of collapsing a 2D hierarchical structure to 1D lines would be nothing new. The manifestation of this strategy in a visualization for the NACA document collection would result in a categorical axis, where documents are binned into a specific subject area (as was done in the word map that was shown before in Figure 6 on page 23).

For a scatter plot these primary dimensions manifests itself as a scatter plot timeline, with the vertical axis denoting time (year of document publishing) and the horizontal axis denoting a categorical classification (i.e. the NACA subject heading table of contents). Inspiration is drawn from Nadieh Bremer's data sketch "Royal Constellations", which uses year of birth/death as vertical axis and royal family/country as horizontal axis which employs this type of mapping (Bremer,2016).

A third dimension would be challenging to show on the 2D screen without using a 3D visualization. Whereas this might be possible, it was decided to rather employ the Bertin retinal variable of lightness/opacity to more easily distinguish points that are more relevant or less relevant by bringing lighter values to the forefront with a filter connected to a primary navigation slider. The rods in the human eye (retina) are most sensitive to lightness and ideally plotting varying lightness values need to be spaced close to each other to allow a comparative effect and exploit the light sensitivity of the retina, hence lightness is ideal for distinguishing many dots on a screen to show grouping. Colour/hue might also be used to create a distinguishing effect; however, care should be taken that the use of many colors still communicate a technical/research oriented visual language. Subsequently, lightness is used as the third dimension (depth).

A fourth dimension of interconnectivity is a primary feature of a network graph and is subsequently used in the NACA document visualization for papers with common authors. Various parameters could be used to create connections, but only if these are distinguishable from each other. Paper Authorship is considered a primary connecting factor between different individual documents, since it would be highly beneficial for searching related content. Keyword connections would be another connection of interest, although keywords might have some overlap with classification categories, which might draw lines to records that are close to each other, hence keywords are at first deemed a secondary connective network parameter.

To differentiate the various lines that connect different records, lightness and/or line weight could be used, especially because a backdrop of points would imply the lines won't clash as much on the backdrop of a scatter plot. Line texture (solid or dashed) is usually effective in Engineering and Architectural drawings and could create a pleasing visual aesthetic also associate with the visual language for an engineering discipline, however experimentation is required to make sure a dashed line does not clash with the texture of the backdrop of many scattered points. Experimentation is therefore needed to isolate which retinal variable: line lightness, weight, and/or texture/dash, is most appropriate for connecting lines.

Lastly an additional layer of meaning can be communicated with regards to the record/point type using size or shape of the point itself, as long as this retinal variable doesn't clash with the spatial spread of points (depending on the density of point spacing), nor should it compete visually with the retinal variable of point lightness (which simulates relevance through depth in/out of the page).

3.5 Implementation

To implement the visualization, the most notable consideration that were made were the available display surface area and graphics performance. Display area is a primary concern, because eliciting context from the full dataset requires records to be visible all at the same time, even if some are shown "inactive" in the background. For 13,801 display points this puts a limitation on the number of pixels on screen that is available to present data. A point is the smallest shape that is visible, allowing for the most records that can be displayed, hence a scatterplot of points is used. To distinguish individual points, these need to be spaced with some buffer in between them. Furthermore, surface area for different types of data representations need to be distinguishable, hence points and lines need to be combined with appropriate size/weight, thickness and layering that do not visually compete.

The D3.js force layout is ideal for clustering since it will both spread out the points into categories (albeit a flattened table of contents hierarchy), and also do individual coordinate spacing required to distinguish individual points using collision detection with an invisible perimeter around each point. The approximate position of the point is applied using a force position along the x-axis and year along the y-axis. This will ensure that all records are positioned in their approximate subject heading region and year of publishing.

To cluster multi-disciplinary records that are binned into multiple subjects (i.e. leaves that are "attached" to multiple branches at once), the average positions of record parent categories are used. Hence, if a record is "mostly aerodynamic" it will end up in the x-axis region of aerodynamics and if it is “mostly structural” it will end up in the x-axis region of “structural”. Points widely varying x-categories will end up in the average subject position they are binned into. This fuzzy clustering will result in a lateral offset when filtered points and hopefully highlight multi-disciplinary records.

The lateral parent category spacing (i.e. the force layout x-position) is determined directly from the NACA Table of Contents, with category bin size based on the approximate amount of records in each binned category, as determined by the order that was determined before (refer to the report coverage by year from Figure 5, p.21). Through experimentation the packing density was optimized by setting the D3.js force collision boundaries slightly larger than the point size to get an acceptable spread and distinguishability of individual records (see Figure 14 and Figure 15 below). Note the shape of the clustering and “trickling” over time; a pattern generated by major subject heading categories.

The full force simulation gave an effective area distribution of points, but performance was sub-optimal for the dynamic layout just after the page loads. This was partly expected due to the large amount of points. According to online forums the maximum number of circles that should be used in SVG is approx. 3000-5000. Hence 13,800 are too many.

Subsequently it was decided to use Canvas/P5 by exporting the converged nodal coordinates from the D3.js force layout to avoid the run-time calculation of point coordinates and increase performance. Using the D3.js layout html page, a custom button was incorporated in the force layout to generate a .csv file containing the x- and y-coordinates stored in a file that is downloaded directly from the browser. The .csv file was then imported into the Python/Pandas scripts that did the final data preparation for use in the final visualization.

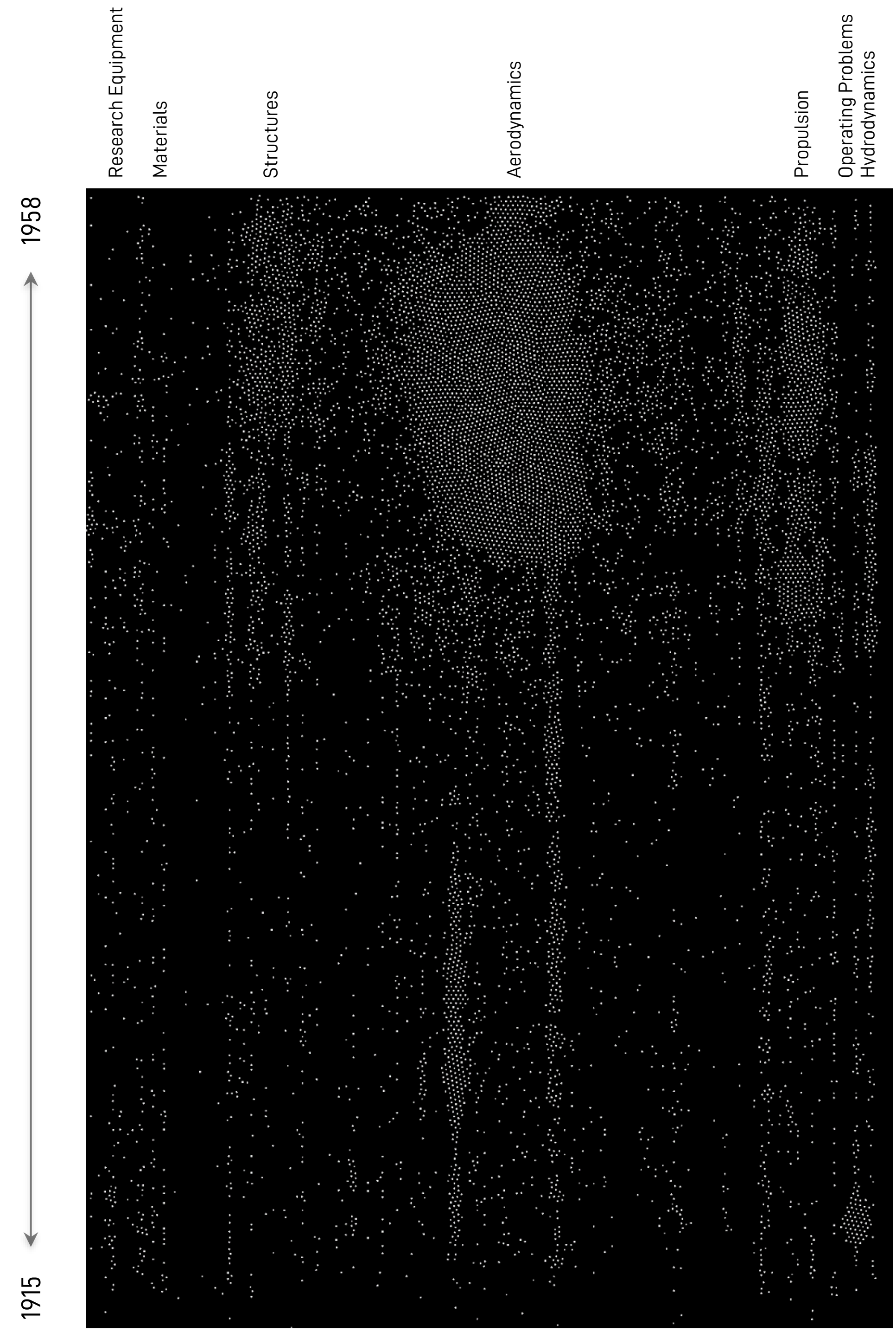


Figure 14: D3 Force Layout of all records as points

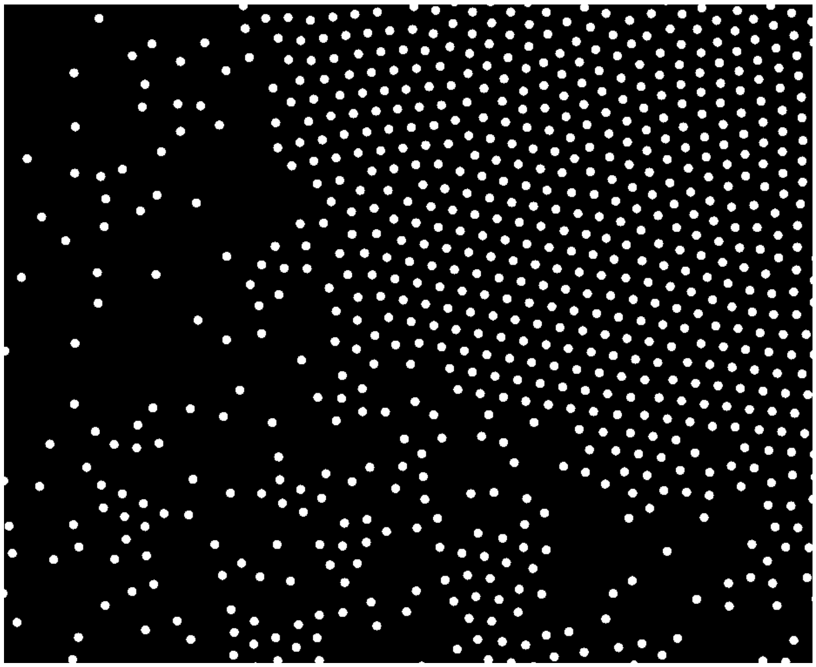


Figure 15: D3 Force Layout Closeup showing point cluster spacing

The final visualization was implemented in p5.js to take advantage of both the ease of synchronous programming, as well as the mouse functions and other supporting methods. Initial loading time for a github hosted site was found to be relatively slow, even though the data files are only 2.5mb, but after the data is loaded the visualization is extremely responsive; showing the advantage of using the Canvas instead of SVG for such a large amount of data points. Streamlining the initial data loading time and initial plotting should be considered in future. One consideration is to exclude the tooltip data from the original data file and only loading relevant point coordinate and category information to do the main plot. Another "hack" would be to create a static image of the scattered point backdrop. Furthermore, a request can be made to the UNT API with the JSON structure returned for each individual point as needed (when clicked upon). Optimizing the response further is beyond the scope of this current project, but should be prioritized in the future.

3.6 Visual Style

The visual style that was employed for the final visualization was that of a blueprint design using the visual language of 20th century architectural and engineering drawings. This style fits well into the era that NACA existed.

The visualization background was graph paper photographed and processed to resemble a blueprint by first inverting the color of the image and then shifting its color temperature to cool colors, while also amending the brightness and contrast. The graph paper has the subtle effect of supplying lines to trace individual points toward axes. Hand drawn icon symbols in the blueprint style were also added to communicate that the x-axis is a categorical axis, as compared to the y-axis which is a continuous axis for time (years). The icons only resemble a general field of the approximate subject heading where it is located. These Icons were taken from assignments done as part of the Typography and Visual design course.

Text fonts were chosen from technical engineering drawing standards such as Tecnico (from European standard IRAM 4503) and the American Society for Mechanical Engineers (ASME) Y14.5M. Another Font that is used is **Courier New** to show a typewriter style. The main logo was taken from the original **NASA** font, but the "C" was custom made to create a visual connection between both the organizations. It was decided not to incorporate the NASA logo in the final page, since it is not the focus of the project, but NACA is. Furthermore, there might be some concerns of legal liability.

Figure 16: The NASA Logo (Note the font)

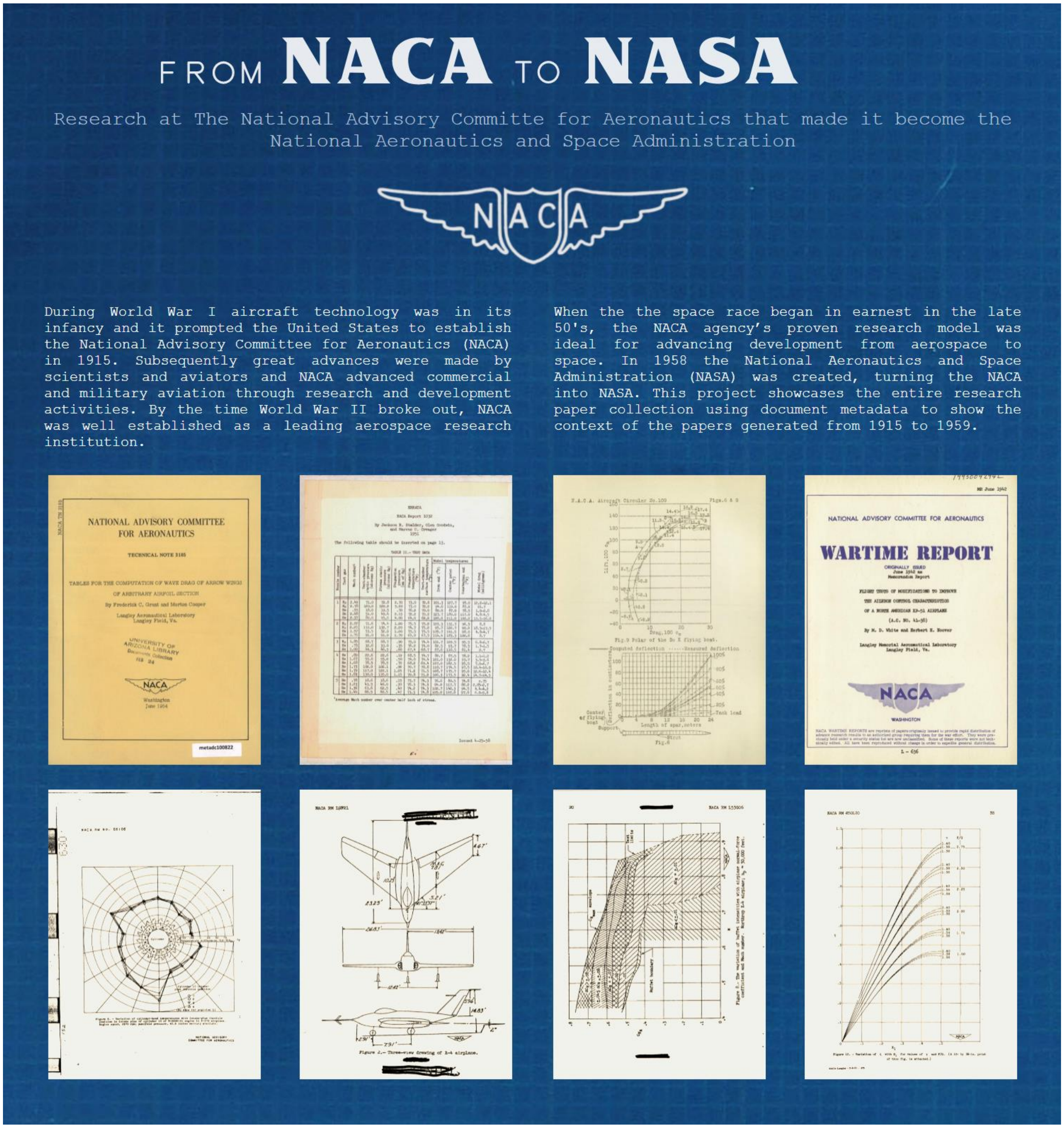


Figure 17: Web Site Opening Section

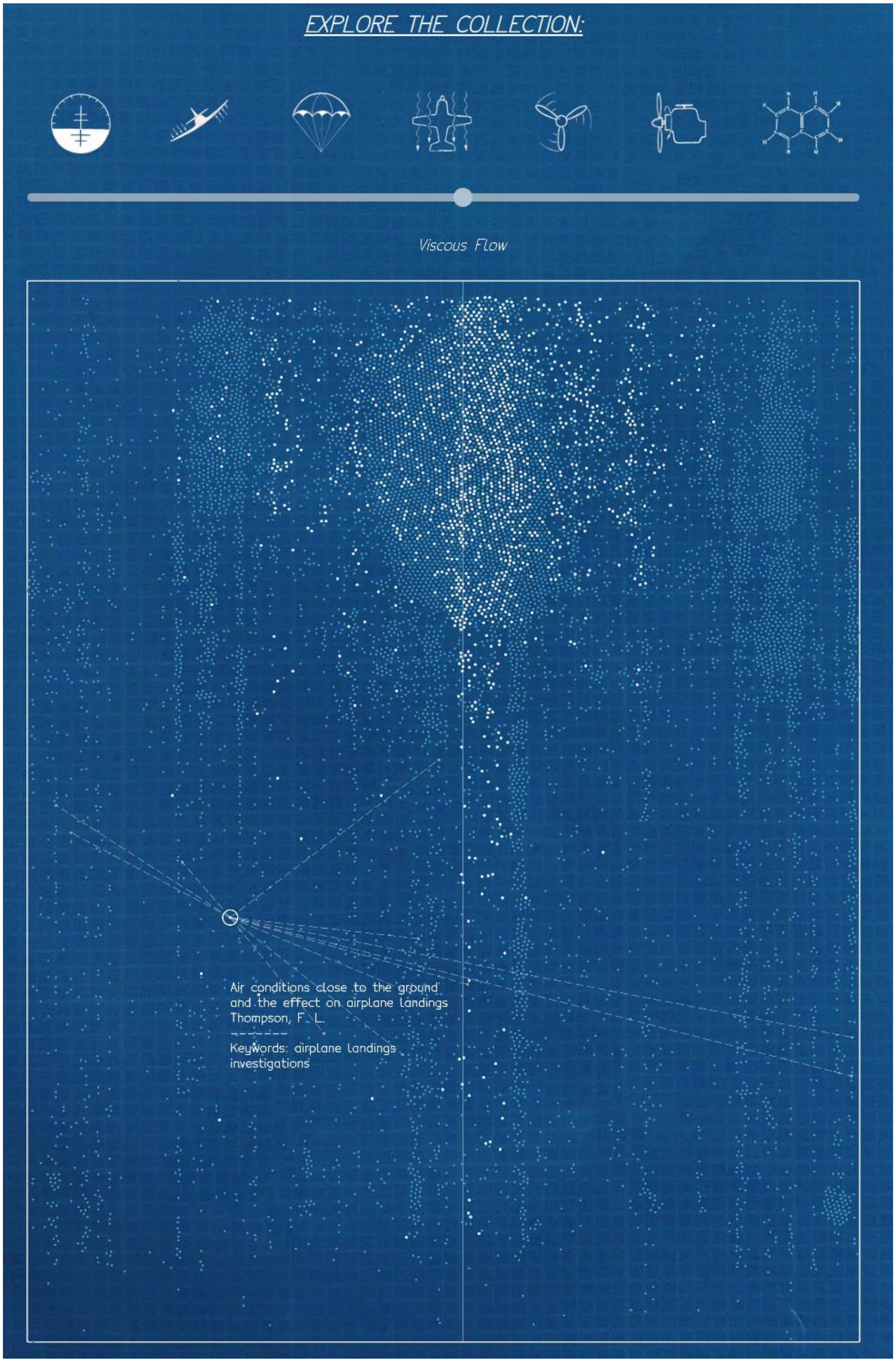


Figure 18: Example State of Main Visualization

4 Conclusion

This thesis presents a data visualization project that shows the entire National Advisory Committee for Aeronautics (NACA) research paper collection located at the University of North Texas digital libraries. The NACA research model historically had great success in the field of Aeronautics and it led to the establishment of NASA. The NACA's research document metadata is used to place Aerospace engineering into context and integrates it into one comprehensive knowledge system. This makes it possible to gain insight into the collection as well as the field of aerospace engineering in general.

Knowledge management was explored in context of the deluge of information that many modern knowledge workers are experiencing. Trees and networks were conceptually explored to present complex information in a structured way. It was found that these underlying data structures have existed for a long time in the form of Table of Contents and Indexes in books. The Table of Contents would typically consist of a top-down classification system of categories driven by splitting and lumping information into respective bins. The bins occur as a parent/child relationship such as cataloguing systems in a library. Indexes and networks typically relate peer-to-peer sibling and bottom-up structuring of information, where the bigger picture could easily be lost without some type of map. When both the top-down and bottom-up structures are combined in a framework likened to a tree that is engulfed in a web, the context of individual records has the potential to become much more clear.

A brief history of developments in Aerospace engineering was then described and shown to include various types of multi-disciplinary subjects. It was found that the NACA collection has a "Table of Contents" index record that was used to extract research paper categories as "leaves on a tree" and provided a way to map the research papers. These links were extracted using OCR text and Python scripts. Single papers ("leaves") were found under an average of 3-4 subject headings (or tree branches).

To visualize individual papers inside this structure without unnecessary duplicating the "leaves", the tree was collapsed to a single dimension with points/documents located on the average position along an axis of the flattened subject headings. A slider selector/filter was incorporated so that the user can highlight applicable categories. This is analogous to running a finger along a list of headings through a Table of Contents in a book (even though the slider/axis is implemented horizontally). The slider is then used to scroll through categories and highlight individual documents (points) in the scatter plot of document records. The user can then select individual filtered points/records.

The date of publication was used to map the documents (points) vertically on the chart. The overall clustering of points illustrates the collection's subject focus over time. Where subjects are scattered in the lateral direction, it indicates multi-disciplinary research and where subjects/points are clustered along vertical lines, it shows single-disciplinary research papers (i.e. papers found in a single subject heading, or in headings that are closely related to each other).

Lastly, each paper's author was added as links using basic "spider network" lines. These lines would spider out from an author that was involved in many types of research across time and subject headings. If an author only published a single paper, there would be no connecting lines. The network-like connections of authors are just one of many types of connections that can be used to interlink the different papers. Other metadata fields include keywords and links that could be added from other sources such as machine learning clusters, however such connections are beyond the scope of this project. The depth of network links could be another future consideration; such as co-authors or co-authors of co-authors.

Finally it is recommended to implement a connection tool that allows a user to navigate the plot as one would do for a "normal" text search online, but to do so visually to "connect the dots". This would allow a user to see the navigation history as a route along the map over time and subject (like a start chart). When such routes are stored and recalled later, it could be a powerful teaching tool to allow a mentor or teacher to show a mentee or student a particular "route" along an assigned map to master a particular topic. The implementation of such a "route map" functionality is envisioned to add a significant contribution as a pedagogical tool to communicate insight into complex multi-disciplinary fields.

5 References:

Bair, Sheila. “Toward a Code of Ethics for Cataloging.” Technical Services Quarterly 23, no. 1 (2005): 13–26. https://doi.org/10.1300/j124v23n01_02.

Besser, Howard (1987a). Digital Images for Museums, Museum Studies Journal 3 (1), Fall/Winter, pages 74-81

Besser, Howard (1987b). The Changing Museum, in Ching-chih Chen (ed), Information: The Transformation of Society (Proceedings of the 50th Annual Meeting of the American Society for Information Science), Medford, NJ: Learned Information, Inc, (pages 14-19)

Bilstein, Roger E. Orders of Magnitude: A History of the NACA and NASA, 1915-1990. The NASA History Series. 1989. https://history.nasa.gov/SP-4406/contents.html

Bremer N., 2016, Royal constellations,http://www.datasketch.es/october/code/nadieh/

FAA,2008: DOT/FAA/AR-08/39 History of Aviation Safety Oversight in the United States

Fitzpatrick, L. Kelly. Shared Practices in Museum Open Collections Data. Feb 22, 2017 https://medium.com/berkman-klein-center/shared-practices-in-museum-open-collections-data-72e924c4849a. Accessed online 03/26/2020.

Gardner, Howard E. Multiple Intelligences: New Horizons in Theory and Practice. New York: Basic Books, 2006.

Geisler, Eliezer. Knowledge and Knowledge Systems Learning from the Wonders of the Mind. Hershey, PA: IGI Publications, 2008

Gitelman, Lisa. "Raw Data" Is an Oxymoron. Cambridge, MA: The MIT Press, 2013.

Kington Miles, 2003, UK Independent, 28 March 2003

Knight, G. Norman (1979) Indexing, the Art of: A Guide to the Indexing of Books and Periodicals (HarperCollins), pp. 17–18

Lima, Manuel. The Book of Circles: Visualizing Spheres of Knowledge. New York: Princeton Architectural Press, 2017.

Lima, Manuel. The Book of Trees: Visualizing Branches of Knowledge. New York: Princeton Architectural Press, 2014.

Lima, Manuel. Visual Complexity: Mapping Patterns of Information. New York: Princeton Architectural Press, 2013.

Markman, Jon. Elon Musk's Neuralink Is Sci-Fi Made Real. Forbes. Aug 30, 2019 https://www.forbes.com/sites/jonmarkman/2019/08/30/elon-musks-neuralink-is-sci-fi-made-real/#18c7c60b4e2f (accessed 03/26/2020)

NACA, 1949, Index of NACA Technical Publications, 1915 – 1949, Division of Research information, Washington,
(Accessed online: https://ntrs.nasa.gov/search.jsp?R=20160003570 on 2020-04-09)

NACA, 1952, Index of NACA Technical Publications, 1949 – May, 1951, Division of Research information, Washington, NASA-TM-80991,
(Accessed online: https://ntrs.nasa.gov/search.jsp?R=19800073356 on 2020-04-09)

NACA, 1954, Index of NACA Technical Publications, June 1953 – May 1954, Division of Research information, Washington,
(Accessed online: https://ntrs.nasa.gov/search.jsp?R=20160003567 on 2020-04-09)

NACA, 1955, Index of NACA Technical Publications, June 1954 – May 1955, Division of Research information, Washington,
(Accessed online: https://ntrs.nasa.gov/search.jsp?R=20150021831 on 2020-04-09)

NACA, 1956, Index of NACA Technical Publications, June 1955 – June 1956, Division of Research information, Washington,
(Accessed online: https://ntrs.nasa.gov/search.jsp?R=20150021830 on 2020-04-09)

NACA, 1957, Index of NACA Technical Publications, July 1956 – June 1957, Division of Research information, Washington,
(Accessed online: https://ntrs.nasa.gov/search.jsp?R=20150021829 on 2020-04-09)

NACA, 1959, Index of NACA Technical Publications, July 1957 – Sept. 1958, Division of Research information, Washington,
(Accessed online: https://ntrs.nasa.gov/search.jsp?R=20150021828 on 2020-04-09)

NACA,1953, Index of NACA Technical Publications, June 1951 – May, 1953, Division of Research information, Washington, NASA-TM-80989,
(Accessed online: https://ntrs.nasa.gov/search.jsp?R=20150021906 on 2020-04-09)

Nelson, Ted: “A File Structure for The Complex, The Changing and the Indeterminate”. ACM 20th National Conference (1965)

Powell, Walter W., and Kaisa Snellman. “The Knowledge Economy.” Annual Review of Sociology 30, no. 1 (2004): 199–220. https://doi.org/10.1146/annurev.soc.29.010202.100037.

Roberts, J. (2015). A very short, fairly interesting and reasonably cheap book about knowledge management 55 City Road, London: SAGE Publications Ltd doi: 10.4135/9781473919709

Roland, Alex. Model Research: The National Advisory Committee for Aeronautics, 1915-1958. Management History; The NACA. 1985. SP-4103. https://history.nasa.gov/SP-4103/

Russell L. Ackoff (1989), “From Data to Wisdom,” Journal of Applied Systems Analysis 16: 3-9

Russell L. Ackoff (1989), Ackoff’s Best. New York: John Wiley & Sons, pp 170 – 172

Smith, Ruth S., 1994, NACA Collections: A directory of significant collections of the documents of the National Advisory Committee for Aeronautics, NASA-CR-195686

UNT, 2020, National Advisory Committee for Aeronautics (NACA) in UNT Digital Library. University of North Texas Libraries. (Accessed online: https://digital.library.unt.edu/explore/collections/NACA/ March 16, 2020).

Victor, Bret. Kill math. http://worrydream.com/#!/KillMath Accessed March 2020.

Watkins R., 2010, Image, Trees cocooned in spider webs, an unexpected side effect of the flooding in Sindh, Pakistan. UK Department for International Development, (Accessed online: https://www.flickr.com/photos/dfid/5331062034, March 28, 2020)

Weinberger, David. Everything Is Miscellaneous: the Power of the New Digital Disorder. New York, NY: Holt, 2008

Appendix A: Web Page Abstract

During World War I aircraft technology was in its infancy; this prompted the United States to establish the National Advisory Committee for Aeronautics (NACA) in 1915. Subsequently, significant research, testing, and development was undertaken by scientists and aviators. The NACA greatly advanced commercial and military aviation through these activities. When World War II occurred NACA was already well established as a leading aerospace research institution.

When the the space race began in earnest in the late 50's, the NACA agency's proven research model was ideal for advancing development from aerospace to space. In 1958 the National Aeronautics and Space Administration (NASA) was created, this was when the NACA became NASA. This project showcases the entire research paper collection using document metadata to show the context of the papers generated from 1915 to 1959.

Appendix B: NACA Report Distribution

Subject Heading by year (1915-1935)	Total	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
Aerodynamics	59	25	17	63	50	23	30	49	41	46	50	46	48	59	56	50	48	50	58	60	56	47
Hydrodynamics	2.2	-	-	-	-	-	-	1	1	-	1	4	3	3	4	2	1	3	3	5	10	11
Propulsion	14.7	8	25	6	50	29	23	18	6	16	8	16	18	16	11	14	15	16	17	12	11	10
Structures	9.3	17	-	13	-	6	4	4	4	7	8	6	7	2	6	17	12	12	6	9	6	17
Materials	5.4	17	-	13	-	16	15	3	5	5	12	2	6	8	7	6	11	4	4	2	4	1
Meteorology	1	8	-	-	-	-	1	1	1	1	-	-	1	-	1	1	-	1	1	1	1	-
Operating Problems	2.5	-	8	-	-	3	2	5	21	8	7	6	4	2	6	2	3	4	3	2	3	4
Instruments	1.5	25	17	-	-	3	6	7	11	7	4	8	5	2	2	4	3	4	4	1	3	4
Research Equipment Techniques	4.3	-	25	-	-	16	18	11	10	8	9	10	8	7	6	4	6	7	4	7	5	4
Nomenclature	0	-																				
Bibliographies and Indexes	0.2	-																				
Technical Summaries	0	-																				

Subject heading by year (1936-1958)	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58
Aerodynamics	51	55	54	53	52	61	52	47	48	45	51	63	65	68	67	63	62	62	62	60	61	57	57
Hydrodynamics	9	1	11	2	3	-	2	7	3	4	3	2	2	4	1	1	2	1	1	1	1	2	2
Propulsion	15	12	9	14	16	10	12	13	17	32	16	15	14	11	15	14	16	16	16	16	14	15	14
Structures	12	13	13	13	10	8	15	14	13	10	10	8	7	7	7	9	8	9	10	10	10	11	11
Materials	2	4	4	7	5	8	7	7	7	8	8	4	5	5	3	6	5	4	5	4	5	7	8
Meteorology	-	-	-	1	1	1	2	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
Operating problems	3	2	2	2	3	4	4	3	3	2	3	2	2	2	1	1	1	2	2	2	2	3	3
Instruments	2	2	2	2	4	3	2	3	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1
Research equipment techniques	5	9	6	6	5	6	4	3	5	5	5	3	3	2	3	4	5	4	4	4	4	4	3
Nomenclature																							
Bibliographies and indexes																							
Technical summaries																							

Appendix C: Full Subject list

Number	Subject-Heading
1	AERODYNAMICS
1.1	Fundamental Aerodynamics
1.1.1	Incompressible Flow
1.1.2	Compressible Flow
1.1.3	Viscous Flow
1.1.4	Aerodynamics with Heat
1.1.5	Flow of Rarefied Gases
1.1.6	Time-Dependent Flow
1.2	wings
1.2.1	Wing Sections
1.2.2	Complete Wings
1.3	Bodies
1.3.1	Theory
1.3.2	Shape Variables
1.3.3	Canopies
1.3.4	Ducted Bodies
1.3.5	Hulls
1.4	Internal Aerodynamics
1.4.1	Air Inlets
1.4.2	Ducts
1.4.3	Exits
1.4.4	Jet Pumps and Thrust Augmentors
1.4.5	Cascades
1.4.6	Fans
1.4.7	Boundary Layer
1.5	Propellers
1.5.1	Theory
1.5.2	Design Variables
1.5.3	Designated Types
1.5.4	Slipstream
1.5.5	Selection Charts
1.5.6	Operating Conditions
1.5.7	Propeller-Spinner-Cowl Combinations
1.6	Rotating Wings
1.6.1	Theory
1.6.2	Experimental Studies
1.7	Aircraft
1.7.1	Airplanes
1.7.2	Missiles
1.7.3	Rotating-Wing Aircraft
1.7.4	Seaplanes
1.7.5	Airships
1.7.6	Biplanes and Triplanes
1.8	Stability and Control
1.8.1	Stability
1.8.2	Control
1.8.3	Spinning
1.8.4	Stalling
1.8.5	Flying Qualities
1.8.6	Mass and Gyroscopic Problems
1.8.7	Tumbling
1.8.8	Automatic Stabilization

1.8.9	Tracking
1.9	Aeroelasticity
1.10	Parachutes
2	HYDRODYNAMICS
2.1	Theory
2.2	General Arrangement Studies
2.3	Seaplane Hull Variables
2.3.1	Length-Beam Ratio
2.3.2	Dead Rise
2.3.3	Steps
2.3.4	Afterbody Shape
2.3.5	Forebody Shape
2.3.6	Chines
2.4	Specific Seaplanes and Hulls
2.5	Lateral Stabilizers
2.5.1	Wing-Tip Float
2.6	Planing Surfaces
2.7	Hydrofoils
2.8	Surface Craft
2.9	Ditching Characteristics
2.1	Stability and Control
2.10.1	Longitudinal
2.10.2	Lateral
2.10.3	Directional
3	PROPULSION
3.1	Complete Systems
3.1.1	Reciprocating Engines
3.1.2	Reciprocating Engines Turbines
3.1.3	Turbojet Engines
3.1.4	Turbo-Propeller Engines
3.1.5	Ducted Propeller Engines
3.1.6	Pulse-Jet Engines
3.1.7	Ram-Jet Engines
3.1.8	Rocket Engines
3.1.9	Jet-Driven Rotors
3.1.10	Nuclear Energy Systems
3.1.11	Miscellaneous Engines
3.1.12	Comparison of Engine Types
3.2	Control of Engines
3.2.1	Charging and Control of Reciprocating Engines
3.2.2	Control of Turbojet Engines
3.2.3	Control of Turbine-Ram-Jet Engines
3.2.4	Control of Turbine Propeller Engines
3.2.5	Control of Pulse-Jet Engines
3.2.6	Control of Ram-Jet Engines
3.2.7	Control of Rocket Engines
3.2.8	Control of Gas Generator Engines
3.3	Auxiliary Booster Systems
3.3.1	Reciprocating Engines
3.3.2	Gas Turbines
3.3.3	Rocket Assist
3.4	Fuels

3.4.1	Preparation
3.4.2	Physical and Chemical Properties
3.4.3	Relation to Engine Performance
3.5	Combustion and Combustors
3.5.1	General Combustion Research
3.5.2	Effect of Engine Operating Conditions and Combustion Chamber Geometry
3.6	Compression and Compressors
3.6.1	Flow Theory and Experiment
3.6.2	Stress and Vibration
3.6.3	Matching
3.7	Turbines
3.7.1	Flow Theory and Experiment
3.7.2	Cooling
3.7.3	Stress and Vibration
3.7.4	Matching
3.8	Friction and Lubrication
3.8.1	Theory and Experiment
3.8.2	Sliding Contact Surfaces
3.8.3	Rolling Contact Surfaces
3.8.4	Sliding and Rolling Contact Surfaces
3.8.5	Lubricants
3.9	Heat Transfer
3.9.1	Theory and Experiment
3.9.2	Heat Exchangers
3.1	Cooling of Engines
3.10.1	Reciprocating Engines
3.10.2	Gas-Turbine Systems
3.10.3	Ram jets
3.10.4	Pulse Jets
3.10.5	Rockets
3.11	Properties of Gases
3.11.1	Kinetic
3.11.2	Thermodynamic
3.12	Accessories and Accessory Functions
3.12.1	Fuel Systems
3.12.2	Ignition Systems
3.12.3	Starting Systems
3.12.4	Lubrication Systems
3.12.5	Cooling Systems
3.13	Vibration and Flutter
4	AIRCRAFT LOADS AND CONSTRUCTION
4.1	Loads
4.1.1	Aerodynamic
4.1.2	Landing
4.2	Vibration and Flutter
4.2.1	Wings and Ailerons
4.2.2	Tails
4.2.3	Bodies
4.2.4	Propeller, Fans, and Compressors
4.2.5	Rotating-Wing Aircraft
4.2.6	Panels and Surface Coverings
4.3	Structures
4.3.1	Columns
4.3.2	Frames, Gridworks, and Trusses
4.3.3	Plates

4.3.4	Beams
4.3.5	Shells
4.3.6	Connections
4.3.7	Loads and Stresses
4.3.8	Weight Analysis
5	MATERIALS
5.1	Types
5.1.1	Aluminum
5.1.2	Magnesium
5.1.3	Steels
5.1.4	Heat-Resisting Alloys
5.1.5	Ceramics
5.1.6	Plastics
5.1.7	Woods
5.1.8	Adhesives
5.1.9	Protective Coatings
5.1.10	Fabrics
5.1.11	Sandwich and Laminates
5.1.12	Ceramals
5.1.13	Titanium
5.2	Properties
5.2.1	Tensile
5.2.2	Compressive
5.2.3	Creep
5.2.4	Stress-Rupture
5.2.5	Fatigue
5.2.6	Shear
5.2.7	Flexural
5.2.8	Corrosion Resistance
5.2.9	Structure
5.2.10	Effects of Nuclear Radiation
5.2.11	Thermal
5.2.12	Multiaxial Stress
5.2.13	Plasticity
5.3	Operating Stresses and Conditions
5.3.1	Airframe
5.3.2	Propulsion System
6	METEOROLOGY
6.1	Atmosphere
6.1.1	Standard Atmosphere
6.1.2	Gusts
6.1.3	Electncity
6.2	Ice Formation
7	OPERATING PROBLEMS
7.1	Safety
7.1.1	Pilot-Escape Techniques
7.2	Navigation
7.3	Ice Prevention and Removal
7.3.1	Engine Induction Systems
7.3.2	Propellers
7.3.3	Wings and Tails
7.3.4	Windshields
7.3.5	Miscellaneous Accessories
7.3.6	Propulsion Systems
7.4	Noise
7.5	Heating and Ventilating

7.6	Lightning Hazards
7.7	Piloting Techniques
7.8	Physiological
7.9	Fire Hazards
7.10	General
8	INSTRUMENTS
8.1	Flight
8.2	Laboratory
8.3	Meteorological
9	RESEARCH EQUIPMENT AND TECHNIQUES
9.1	Equipment
9.1.1	Wind/Tunnels
9.1.2	Free-Flight
9.1.3	Towing Tanks and Impacts Basins

9.1.4	Propulsion Research Equipment
9.1.5	Propeller
9.1.6	Materials
9.1.7	Structures
9.2	Technique
9.2.1	Corrections
9.2.2	Aerodynamics
9.2.3	Hydrodynamics
9.2.4	Loads and Construction
9.2.5	Propulsion
9.2.6	Operating Problems
9.2.7	Mathematics
10	NOMENCLATURE
11	BIBLIOGRAPHIES
12	TECHNICAL SUMMARIES