The tyranny of formatting

Maryann E. Martone, Ph. D.*
Dept of Neuroscience
University of California, San Diego

San Francisco Veterans Administration Hospital

*Founder and equity interest in SciCrunch Inc

It is 4 am in the morning and the grant is due the next day. You have 12 pages to tell a story that will determine whether or not you can pay your salary and support your lab next year. You delete a sentence on page 5. The document swells to 13 pages. Oh no! The figures have all jumped around! The next paragraph is now in Courier 14 bold! You hit your head against the screen and spend the next hour trying to get everything back in place and the sentence deleted. The next day, you're tired, irritable and have a broken computer screen.

-formatting articles for different publishers
-formatting references
-difficulty of text mining

Isn't it time that we as scholars throw our copy of Microsoft Word out the window and say "I'm mad as hell and I'm not going to take it anymore". Has anyone tried to calculate the unproductive hours spent by scholars on formatting? Has anyone else been reduced to tears by jumping figures at 4 am? Has anyone then asked "Why?"

-From "The tyranny of formatting", blog by Maryann Martone first published at FORCE11 in 2013

https://zenodo.org/record/7851203

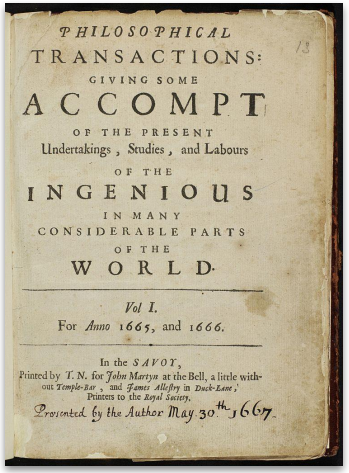
My journey so far…

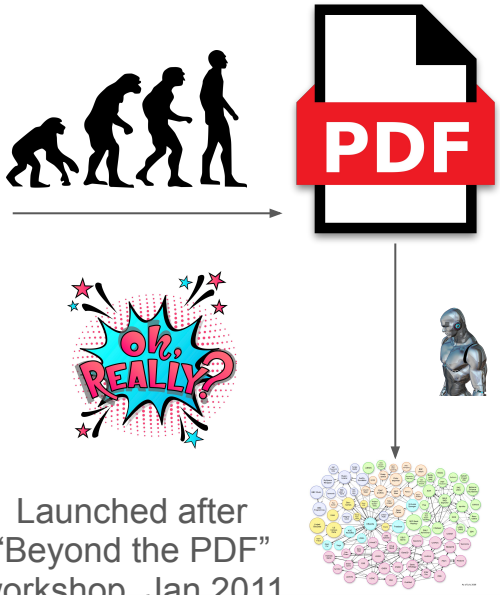
- I am a neuroscientist by with a specialty in neuroanatomy and neuroinformatics
- My career spans the introduction of personal computers and the internet
- In the 1990's, neuroscience like many biomedical domains began to adopt computers
- Many of us saw the potential of the internet and computers to transform the way we do science
- That movement came together as FORCE11 in 2011

https://force11.org/

FORCE11

- **F**uture of **R**esearch **C**ommunications and **E**-Scholarship
  - A grassroots effort to accelerate the pace and change the nature of scholarly communications and e-scholarship through technology, education and community
- Brought together groups that typically don't interact as peers: researchers, scholars, publishers, librarians, technologists
- FORCE11 Manifesto: Outline for harnessing the new medium of the web for nextgen scholarly communications

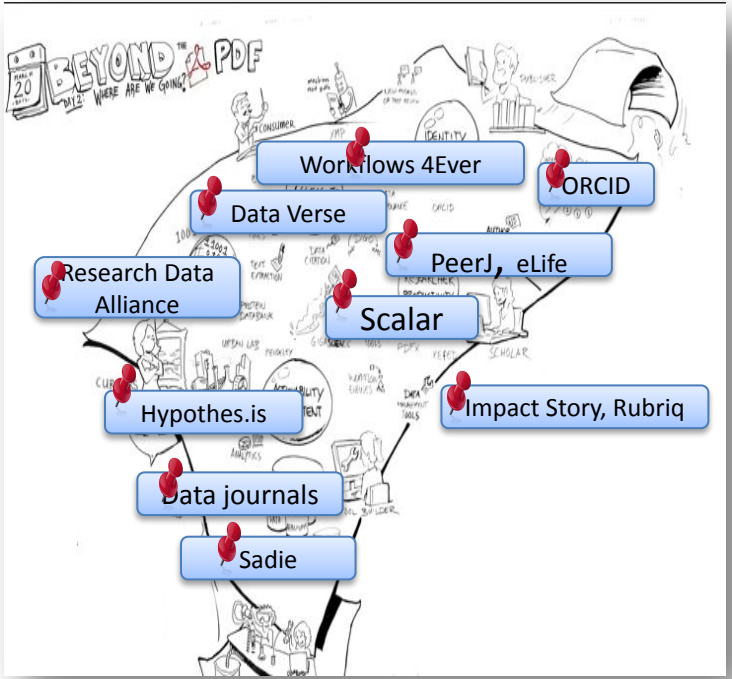

PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.
Vol I.
For Anno 1665, and 1666.
In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little without Temple-Bar, and James Allestry in Duck-Lane,
Printers to the Royal Society.
Presented by the Author May 30th 1667.



Launched after "Beyond the PDF" workshop, Jan 2011 UCSD

Those early conferences were eye opening…

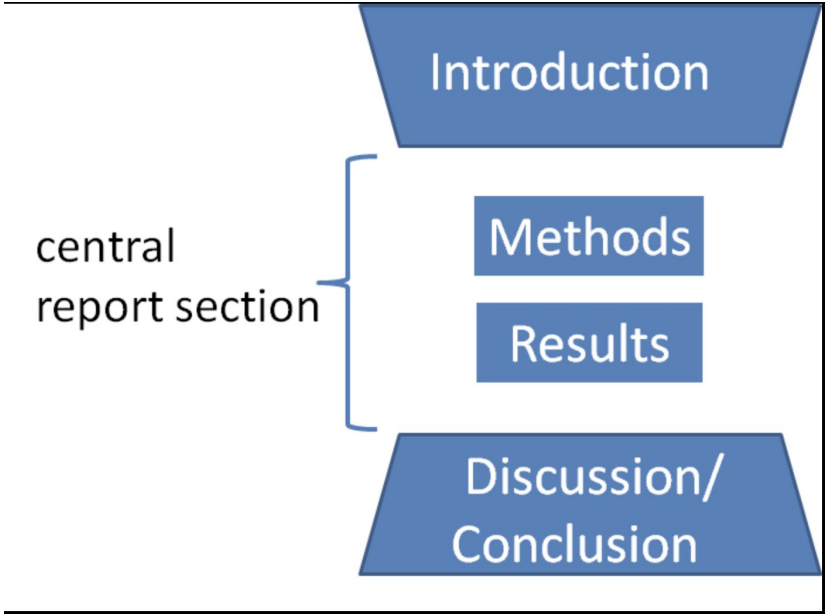
- The atmosphere was electric in those first conferences
- New tools and technology dominated, but great discussions as well
- Memorable quotes:
  - *"You know, other fields don't do authorship like that."* - Dan O'Donnell
  - *"Whether you like it or not, publishers got to the internet first"* - unknown publisher
  - *"Formatting references is completely unnecessary"* - *unnamed publisher in conversation*
  - *"Articles should be published in XML and nothing else"* - Kavegh Bazargan



“Whether you like it or not, publishers got to the internet first”

- Current paper-based practices were enshrined in software and therefore difficult and expensive to change:
  - Practice in biomedicine of only first and last author receiving credit
  - Putting the burden of formatting references on authors
  - Forcing authors to conform to multiple variants of article styles
  - Charging authors for typesetting
- The scientific article is just 6000 variations on a theme

Structure is so stereotyped it has a name: IMRaD

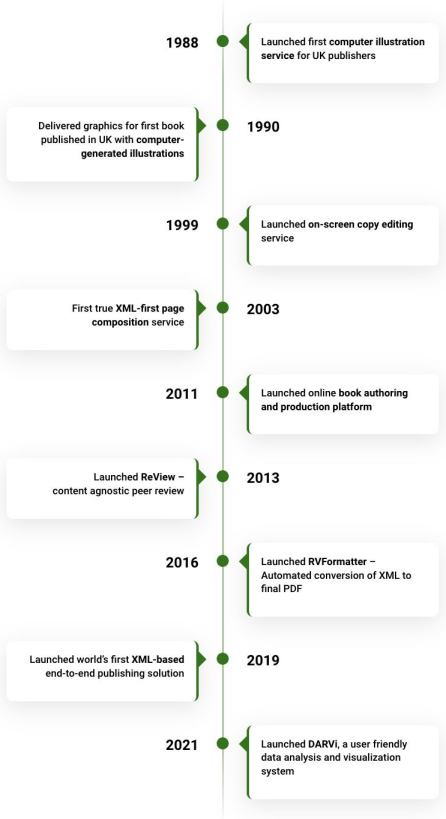


https://en.wikipedia.org/wiki/IMRAD

Tom Toyosaki

“Articles should be published in XML and nothing else”

- Kavegh Bazargum gave a talk at vision session at the 2013 Beyond the PDF conference in Amsterdam
- Head of River Valley Technologies, a scholarly publishing platform
- The pdf is a very inflexible and computationally intractable form
- Html is a little better for linking on the web but it is for formatting
- Advocated for publishing all articles in XML



https://rivervalley.io/

Too many practices are still relics of the print age

- The US National Institutes of Health requires grant applications to conform to specific page limits, usually 12 pages
- The switch to computers led to researchers taking advantage of flexibility in fonts and spacing to cram more information in
- *“Adherence to font size, type density, line spacing, and text color requirements is necessary to ensure readability and fairness.”*
- As a researcher, I don’t like page limits, but as a reviewer, I very much appreciate them
- But relying on page limits and font sizes imposes a real cost on the research enterprise

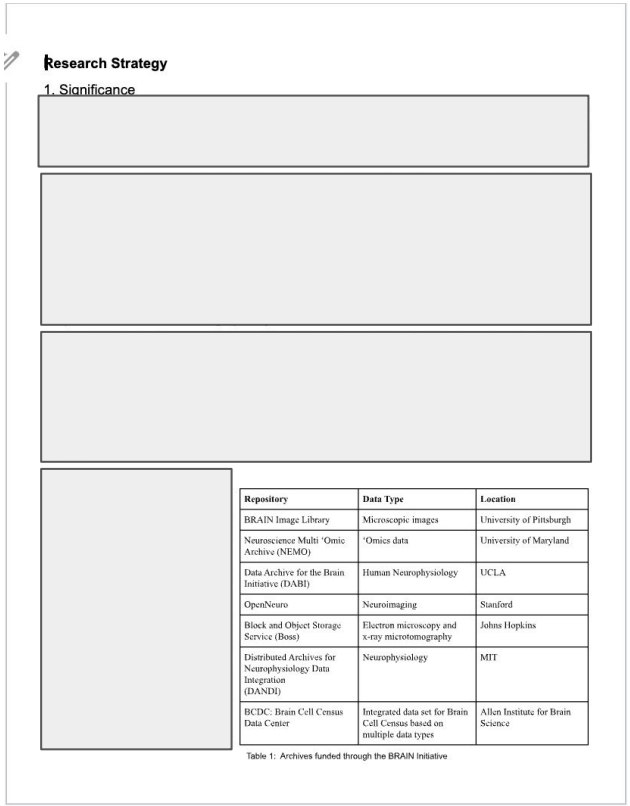
https://grants.nih.gov/grants/how-to-apply-application-guide/format-and-write/format-attachments.htm#

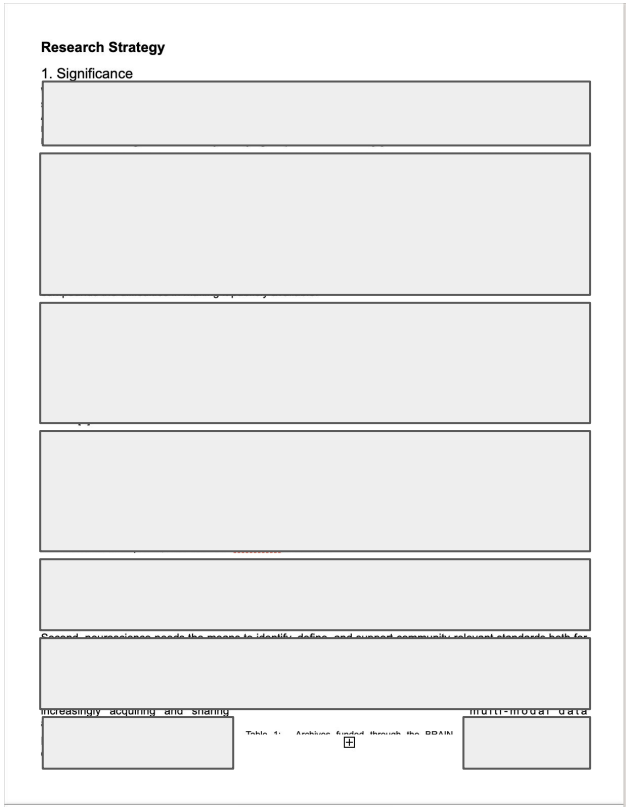
How are things now: Technology

- Google Docs has made great strides since 2013.
- Figures and tables work much better
- You can use Google Docs with some reference managers now
- Google has realized we don't need to constrain our work by the size of a page if we don't want to
- There are other platforms for collaborative editing, but Google is by far and away my favorite
- PubMed Central open access is available for text mining in XML and JSON
- But, NIH hasn't changed their policy and that 4 am scenario still holds
- Why…

About those figures and tables…

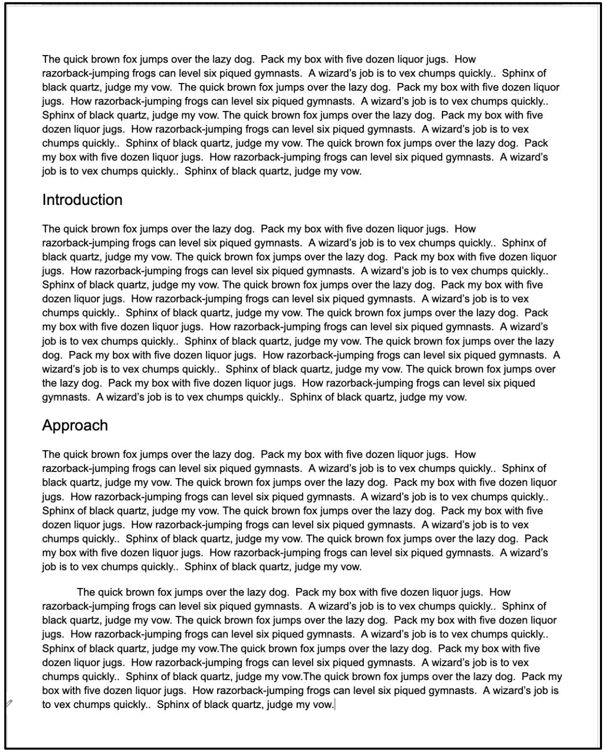
- Final formatting is still usually done in Word
- Downloading from Google Docs is always a crap shoot
- Still a source of considerable frustration





Google Docs → Pages

Interoperability across word processing programs still buggy



The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.

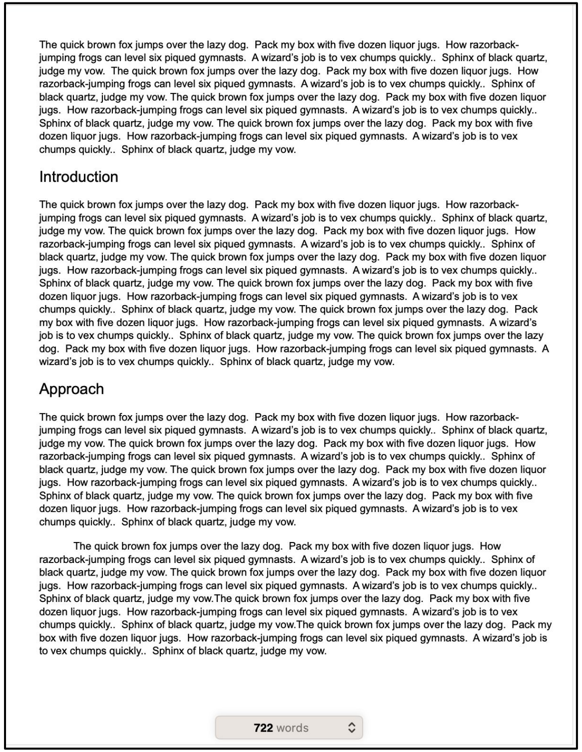
Introduction

The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.

Approach

The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.

The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.



The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.

Introduction

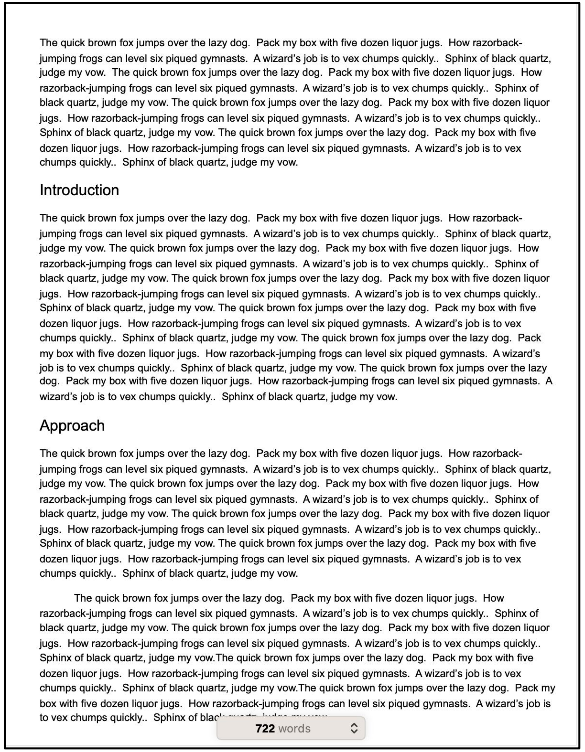
The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.

Approach

The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.

The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.

722 words



The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.

Introduction

The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.

Approach

The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.

The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow. The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of black quartz, judge my vow.The quick brown fox jumps over the lazy dog. Pack my box with five dozen liquor jugs. How razorback-jumping frogs can level six piqued gymnasts. A wizard's job is to vex chumps quickly.. Sphinx of blac

722 words

Created in Google Docs; Saved as Word doc → Opened in Pages; saved as Word → Opened again in Pages

How are things now: policies

- Most publishers impose word and figure limits for different article types
- Conferences impose word limits with a figure counting for a given number of words
- Some publishers are loosening requirements for formatting upon submission:
  - Cascading reviews
  - Integration with preprint services
  - Formatting after a review is done

Formatting is not FAIR: Text as data

- Articles are no longer just human artifacts but data to consume
- Text mining the entire biomedical corpus is still difficult
  - Licenses: Open access subset of PubMed Central that can be text mined is roughly 50% of the articles available
  - Tables are still very difficult to extract
- Publishers may have gotten to the internet first, but their obligations have changed

“Tyranny of formatting” generated by Google Doc’s resident AI artist

Difficulty in extracting tables from articles

- A lot of data and information is contained in tables
- Recognizing and extracting tables in text still a big challenge

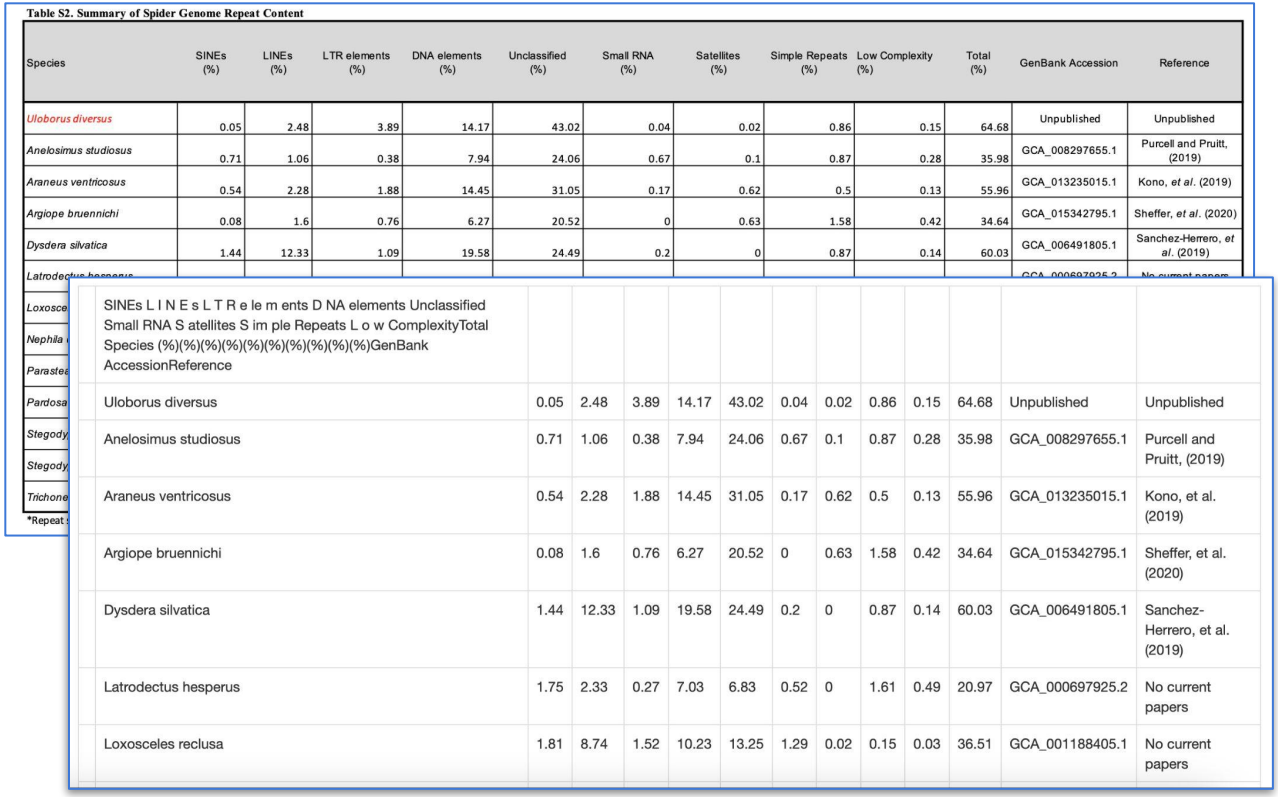


Table S2. Summary of Spider Genome Repeat Content

Species	SINEs (%)	LINEs (%)	LTR elements (%)	DNA elements (%)	Unclassified (%)	Small RNA (%)	Satellites (%)	Simple Repeats (%)	Low Complexity (%)	Total (%)	GenBank Accession	Reference
Uloborus diversus	0.05	2.48	3.89	14.17	43.02	0.04	0.02	0.86	0.15	64.68	Unpublished	Unpublished
Anelosimus studiosus	0.71	1.06	0.38	7.94	24.06	0.67	0.1	0.87	0.28	35.98	GCA_008297655.1	Purcell and Pruitt, (2019)
Araneus ventricosus	0.54	2.28	1.88	14.45	31.05	0.17	0.62	0.5	0.13	55.96	GCA_013235015.1	Kono, *et al.* (2019)
Argiope bruennichi	0.08	1.6	0.76	6.27	20.52	0	0.63	1.58	0.42	34.64	GCA_015342795.1	Sheffer, *et al.* (2020)
Dysdera silvatica	1.44	12.33	1.09	19.58	24.49	0.2	0	0.87	0.14	60.03	GCA_006491805.1	Sanchez-Herrero, *et al.* (2019)
Latrodectus hesperus												
Loxosce												
Nephila												
Parastea												
Pardosa												
Stegody												
Stegody												
Trichone												

*Repeat

SINEs L I N E s L T R e le m ents D NA elements Unclassified Small RNA S atellites S im ple Repeats L o w ComplexityTotal Species (%)(%)(%)(%)(%)(%)(%)(%)(%)(%)GenBank AccessionReference

Uloborus diversus	0.05	2.48	3.89	14.17	43.02	0.04	0.02	0.86	0.15	64.68	Unpublished	Unpublished
Anelosimus studiosus	0.71	1.06	0.38	7.94	24.06	0.67	0.1	0.87	0.28	35.98	GCA_008297655.1	Purcell and Pruitt, (2019)
Araneus ventricosus	0.54	2.28	1.88	14.45	31.05	0.17	0.62	0.5	0.13	55.96	GCA_013235015.1	Kono, et al. (2019)
Argiope bruennichi	0.08	1.6	0.76	6.27	20.52	0	0.63	1.58	0.42	34.64	GCA_015342795.1	Sheffer, et al. (2020)
Dysdera silvatica	1.44	12.33	1.09	19.58	24.49	0.2	0	0.87	0.14	60.03	GCA_006491805.1	Sanchez-Herrero, et al. (2019)
Latrodectus hesperus	1.75	2.33	0.27	7.03	6.83	0.52	0	1.61	0.49	20.97	GCA_000697925.2	No current papers
Loxosceles reclusa	1.81	8.74	1.52	10.23	13.25	1.29	0.02	0.15	0.03	36.51	GCA_001188405.1	No current papers

Example courtesy of Peter Eckmann and Dr. Anita Bandrowski, UCSD

https://www.biorxiv.org/content/10.1101/2022.06.14.495972v1

Imagining a world without formatting…

Glia takes my XML and makes it look very pretty. They advertise it on their website and sell the pretty version back to institutions and individuals so they can read it. Some scholars have accused the publishers of existing only to sell us formatting. That's OK. I pay for formatting all the time. I pay extra for the floral pattern on the Kleenex box because it goes with my color scheme. But my mother, who has no institutional subscription, can go to Pub Med Central and get my vanilla-formatted but still functional document. Sometimes, I buy generic.

Imagining a world without formatting…

Meanwhile, my text miner friends have a subscription to the Global ArXiv via their institution or a personal one, if they are not affiliated with an institution. They agree that they will not try to recreate an individual article for resale and then mine away. Perhaps a new target is discovered for my child's disease or maybe there isn't enough information available. But at least we have uniform access to the entire corpus in a form suitable for mining.

Imagining a world without formatting…

*Now, with all my extra time saved by no formatting, I am largely done writing my grant proposal at 11 pm instead of 4 am. It needs to be 10,000 words with each figure counting for 500 words. At 11 pm, I am 10 words over. I delete a sentence and hit "Done". At 11:15, I'm brushing my teeth. Grant reviewers get access to a nifty program that allows them to view/print/read in a variety of formats depending on their medium. One reviewer prefers to print things out on paper. Each figure is a full page and the font is 18 point so the grant is 40 pages long. But a younger reviewer with better vision who still likes paper chooses 12 point font and a half page for figures and her version is 12 pages. Another reviewer has an iPAD and likes to swipe and zoom but he still has my same 8, 000 words and 4 figures. Same content, same length-**no extra burden on the reviewers**-just different formatting.*

But we are not quite there yet

“Freedom from formatting” generated by
Google Doc’s resident AI artist