

## WHY PDF?

**Key Value Prop and its implications for the future of PDF**



**Catherine Andersz  
CEO, PDFTron Systems Inc.**

Catherine Andersz, CEO  
PDFTron Systems Inc.



# Agenda

---

- **A bit about PDF**
- **Why PDF? - Different perspectives**
- **Understanding the core value prop of PDF**
- **What does the future hold for PDF?**
- **Questions**



## Background: PDFTron

**We provide Complete, Cross-platform Document Processing Technology for PDF, Office and 30+ other file formats.**

**We helped 5,000+ innovative start-ups, government, and Fortune 500 businesses create document experiences loved by millions of users.**

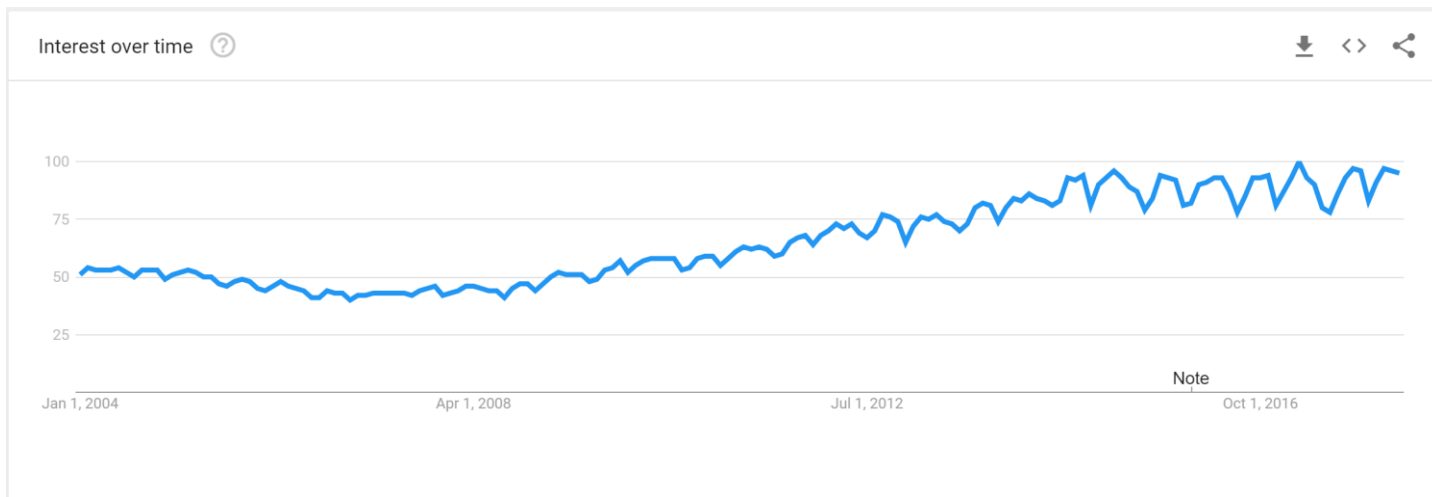


Catherine Andersz, CEO  
PDFTron Systems Inc.



## A bit about PDF

PDF is one of the **most common** file formats on the planet  
...And interest seems to be growing steadily over time.



**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.

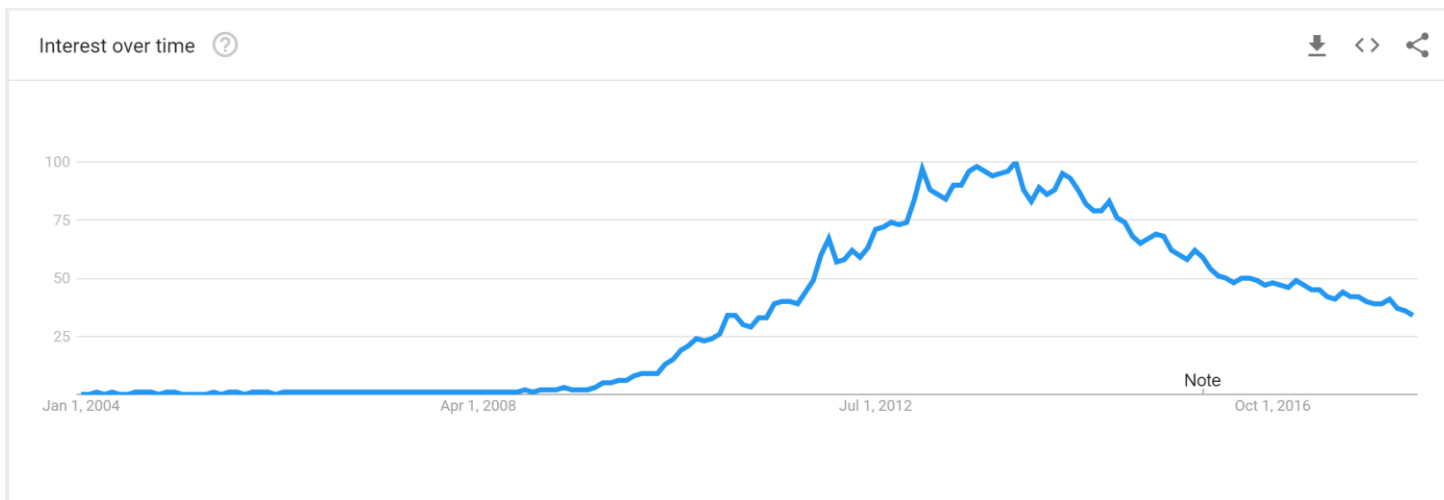
2018-05-14

(Ref. Google Trends)



## A bit about PDF

### In comparison, ePUB interest ...



**PDFTRON**

(Ref. Google Trends)

Catherine Andersz, CEO  
PDFTron Systems Inc.



## A bit about PDF

### PDF is Everywhere:

- **Used in airplane cockpits, clinical trials, courtrooms, government websites...**
- **We run into PDF in libraries, online research, emails, signing documents...**
- **More and more PDFs occupy our digital storage**
- **Used even when we don't know it, when...**
  - Printing/faxing
  - Sending physical mail
  - Collaborating with others online

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



# Why PDF?

---

- **Why is PDF so pervasive?**
- **What is the key to its success?**
- **Why is it so important?**

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Why PDF? - Different Perspectives

**John E. Warnock, Cofounder of Adobe, about launching PDF:**

**“... the world didn’t get it. They didn’t understand how important sending documents around electronically was going to be.”**

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.





# Why PDF – Different Perspectives

---

**25 years later...**

**Very few ❤️ PDF, but we all **need** it ... WHY?**

**Why are we even talking about a file format in the increasingly format-less world?**

- **music/movies → streaming services**
- **Gdoc “file” in your Google Drive?**

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Why PDF – Different Perspectives

**Jakob Nielsen, usability guru, declared PDF as “Unfit for Human Consumption”**



# Why PDF – Different Perspectives

Science publishing researchers once described the format as:



# Why PDF

**Evolution of format over last 25 years, added layers of complexity**



**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Why PDF

**PDF continued to adapt. Flexibility to accommodate to an ever-increasing range of capabilities & use cases.**

- **Text Search**
- **Copy/Paste**
- **Bookmarks**
- **Links**
- **Forms**
- **Geo-tags**
- **Metadata**
- **Structure**
- **JavaScript**
- **3D**
- **Movies**
- **Sounds**
- **Annotations**
- ... And much, much more!

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Let's Peel Away the Layers



# Value Prop

## PDF is “electronic paper”



**PDFTRON**

Catherine Andersz, CEO  
PDFtron Systems Inc.



## Value Prop

**PDF is “electronic paper”**

**Like paper, the essential purpose of PDF is to provide ...**

**... a cheap, simple, efficient and reliable  
communication of visual information chunks.**

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.





## Value Prop

Like paper, PDF is relatively **cheap to produce** and **view**:

- **Cheaply produced:** via print drivers
- **Cheaply viewed:** universal viewing across devices and operating systems. (Built-in support on all major OS-s, and many free advanced viewers on every platform)

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Value Prop

Like paper, PDF is a **globally accepted standard** used by **billions of people**.



**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Value Prop

Like paper, PDF is **self-contained** and **simple to share**.

Letter **P** in PDF stands for “**PORTABLE**”



**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Value Prop

Like paper, PDF is **lightweight** and **doesn't take much space**.

- **Faster to transmit and store** with support for scalable vector graphics, advanced compression schemes.
- **Much smaller memory footprint** (compared to SVG or HTML).
- **Content is paginated** (can be rapidly accessed in **constant time**).



## Value Prop

Like paper, PDF is **durable**.

- **Even more durable with PDF/A** - does not suffer from wear & tear.
- **PDF is a sharp contrast to the ephemeral nature of web and application formats.**

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Value Prop

Like paper, most people feel **safe** reading PDF.

- no viruses
- malware
- expected viewing behaviors (no pop-ups, unusual interfaces, etc.)
- reliable & offline

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Value Prop

Like paper, PDF makes it **simple** for people to 'get on the **same page**' with a reliable and consistent frame of reference.

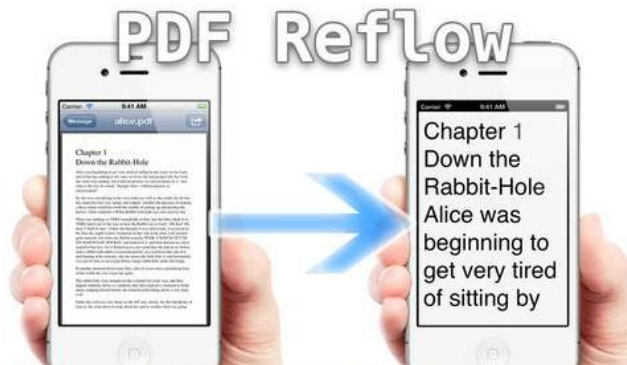
- page numbers
- spatial positioning
- accurate rendering



## Value Prop

Like paper, PDF is **difficult to stretch** to:

- reflow content
- accommodate pages with different aspect ratios



**PDFTRON**

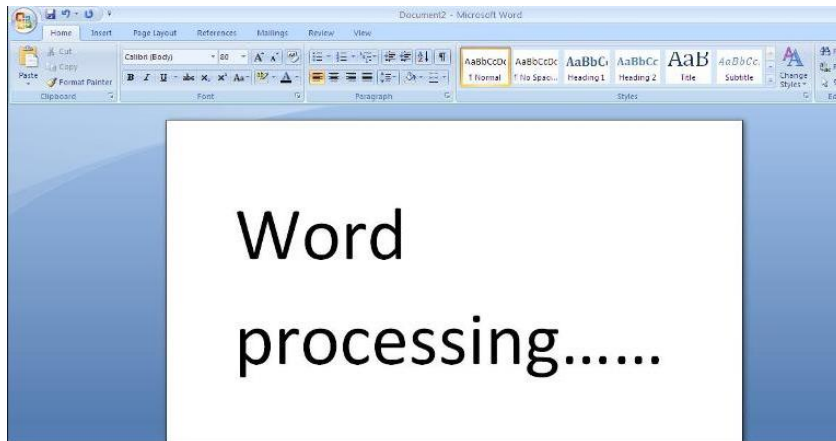
Catherine Andersz, CEO  
PDFTron Systems Inc.





## Value Prop

Like paper, PDF is **difficult to repurpose** and **edit**.



## Value Prop

Like paper, PDF is our **ground truth** — the objective reality.



**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Value Prop

In comparison, semantics is **highly subjective** and **context dependent**:

- **Most education goes into interpreting document semantic structures** (children flipbooks, tax forms, music scores, encyclopedias, novels).
- **No universally agreed-on semantics, just endless variations.**



## Value Prop

Semantics is anchored in culture/language, which (like the web platform) is **constantly changing**:

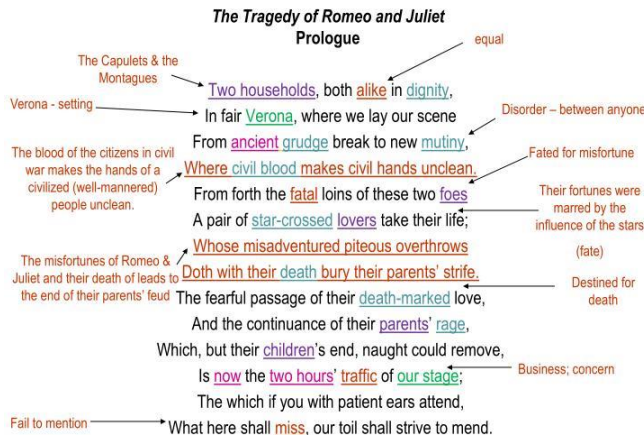
- Documents of ancient cultures (missing semantic info and knowledge to interpret).



# Value Prop

Semantics is anchored in culture/language, which (like the web platform) is **constantly changing**:

- Shakespeare (requires many annotations to decipher/understand).



## Value Prop

Like paper, **semantic information** is **not part** of PDF essence.

- Yes, some semantics info may be present:
- Many **PDFs are searchable** these days
- It is possible to **tag** and embed metadata and semantic structure.
- In near future, we may be able to embed HTML, data, and even JavaScript in PDF as part of **next-generation PDF**.

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Value Prop

... but on a big scale of things, it did not happen:

1) we'll never have “Unicode” of document structures.



**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



## Value Prop

### 2) tension with other value propositions

- **Simple, cheap creation**
- **Durability**
- **Efficiency**
- **Security**

**Requiring the structure to be present (e.g. PDF/UA, PDF/A, or next-generation PDF) goes against the grain of paper and its electronic counterparts.**





## Value Prop

**PDF's loose semantic requirements are its:**

**greatest strength  
and  
greatest weakness**

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



# Value Prop

## Core Value Prop

= Key to Understand, Market and Evolve PDF  
= Answer to "Why PDF?"

Understanding Core Value Prop, helps us not only to:



**Better Market the technology**



**Decide on future of the technology**

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



# What does the future hold for PDF?

**Given the key value prop of PDF, are we there yet?**

**Not quite...**

**Ideally, we want to increase its strengths and decrease its weaknesses.**

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



# What does the future hold for PDF?

## For example:

- **If PDF's greatest strength is that it is reliable (our ground truth), let's make it truly reliable.**
  - **include some of the PDF/A-B provisions in the core ISO32000**
- **Can we make PDF more safe & predictable?**



# What does the future hold for PDF?

## What about PDF's greatest weakness (semantic structure)?

- Continue effort of **adding explicit structure** in PDF... part of the effort behind '**next-generation PDF**'.
- With emergence of new **AI** techniques and advancements in **machine learning**, can we:
  - **transcend PDF limitations** (reflow/repurpose PDF, data recognition/extraction)
  - **build on PDF's key value prop**
  - **without necessarily the need to "fix PDF"**?



# What does the future hold for PDF?

**We need to work as a community to ensure that:**



- Don't undermine the key value prop of PDF, but instead build on its strengths.
- Meet the broadest future market needs and associated use cases & workflows required over the next 25 years or more.

**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.



# Questions



**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.

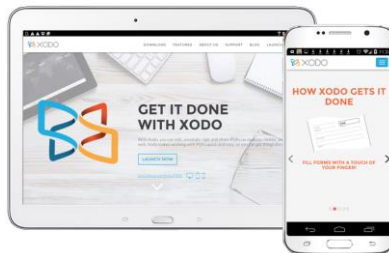


# Thank you!

Get in Touch: [info@pdftron.com](mailto:info@pdftron.com)

PDFTron Website: [www.pdftron.com](http://www.pdftron.com)

Xodo Showcase: [www.xodo.com](http://www.xodo.com)



**PDFTRON**

Catherine Andersz, CEO  
PDFTron Systems Inc.

