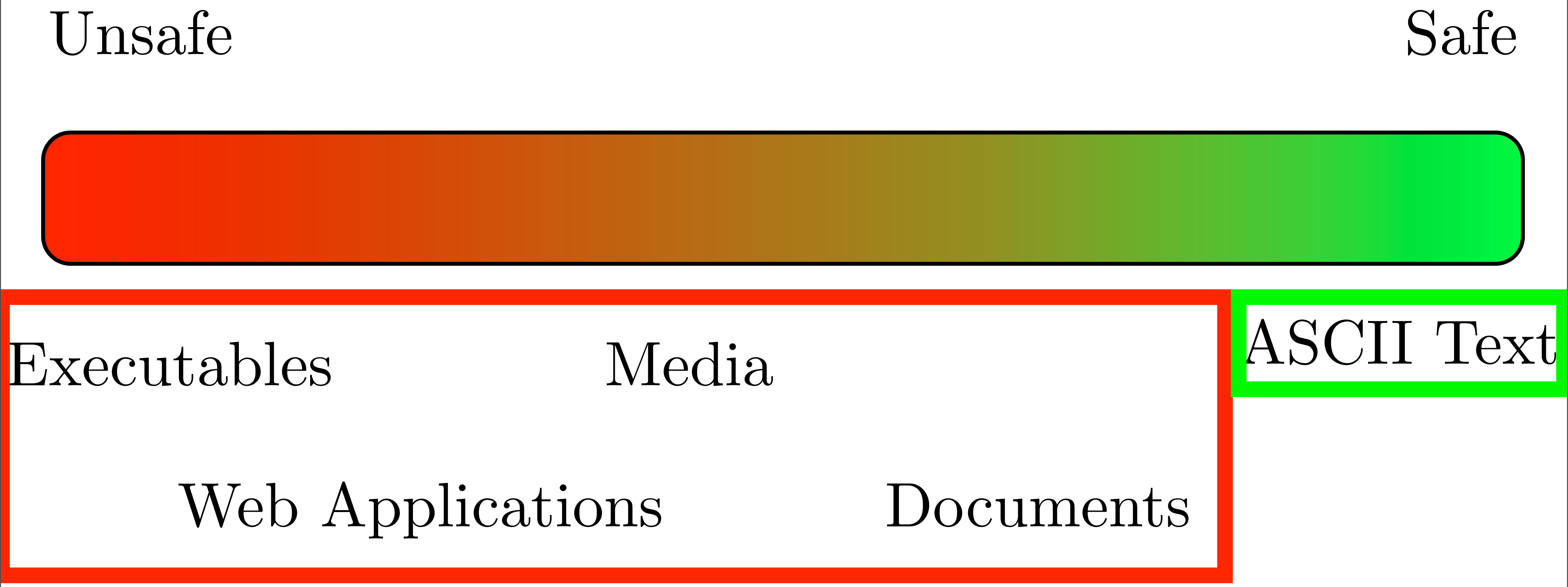
Are Text-Only Data Formats Safe?

Stephen Checkoway, Hovav Shacham, Eric Rescorla

Intuitive data-safety scale



Unsafe

Safe

Executables

Media

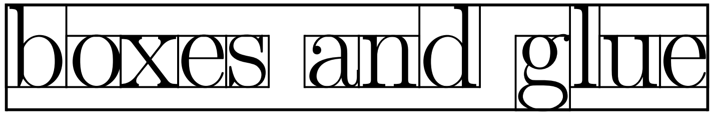
Web Applications

Documents

ASCII Text

TeX

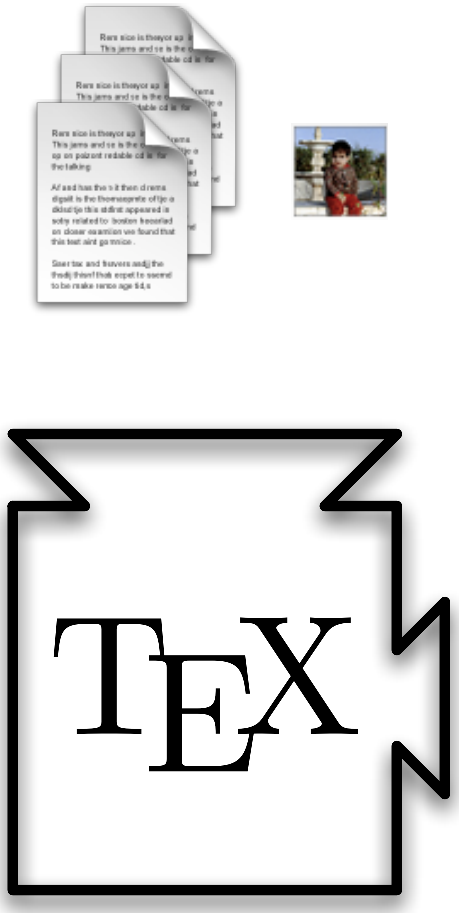
- Document preparation language
- 7-bit ASCII text
- Understands boxes and glue
- Makes pretty equations



boxes and glue

$$D(H\|R) = \sum_{x,y\in\mathcal{X}} H(x,y)\;\log\frac{H(x,y)}{R(x,y)}$$

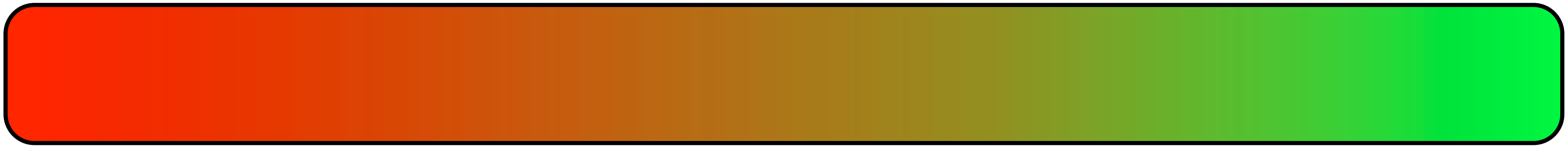
How we use $\TeX$



Intuitive data-safety scale

Unsafe

Safe



Executables

Media

ASCII Text

Web Applications

Documents

TeX

More TeX

- Turing-complete, macro language: `\def`
- Read/write files: `\read`, `\write`
- Extremely malleable syntax: `\catcode`

Taking control with TeX

Distribution	Operating System	How
MiKTeX		Write to `Startup`
TeX Live		Write to web directory

LaTeX virus lifecycle

- Compile `sploit.tex`
- C:\DOCUME~1\ADMINI~1\STARTM~1 \PROGRAMS\STARTUP\sploit.js
- Restart computer
- `sploit.js` finds `.tex` files; inserts the virus

Data exfiltration

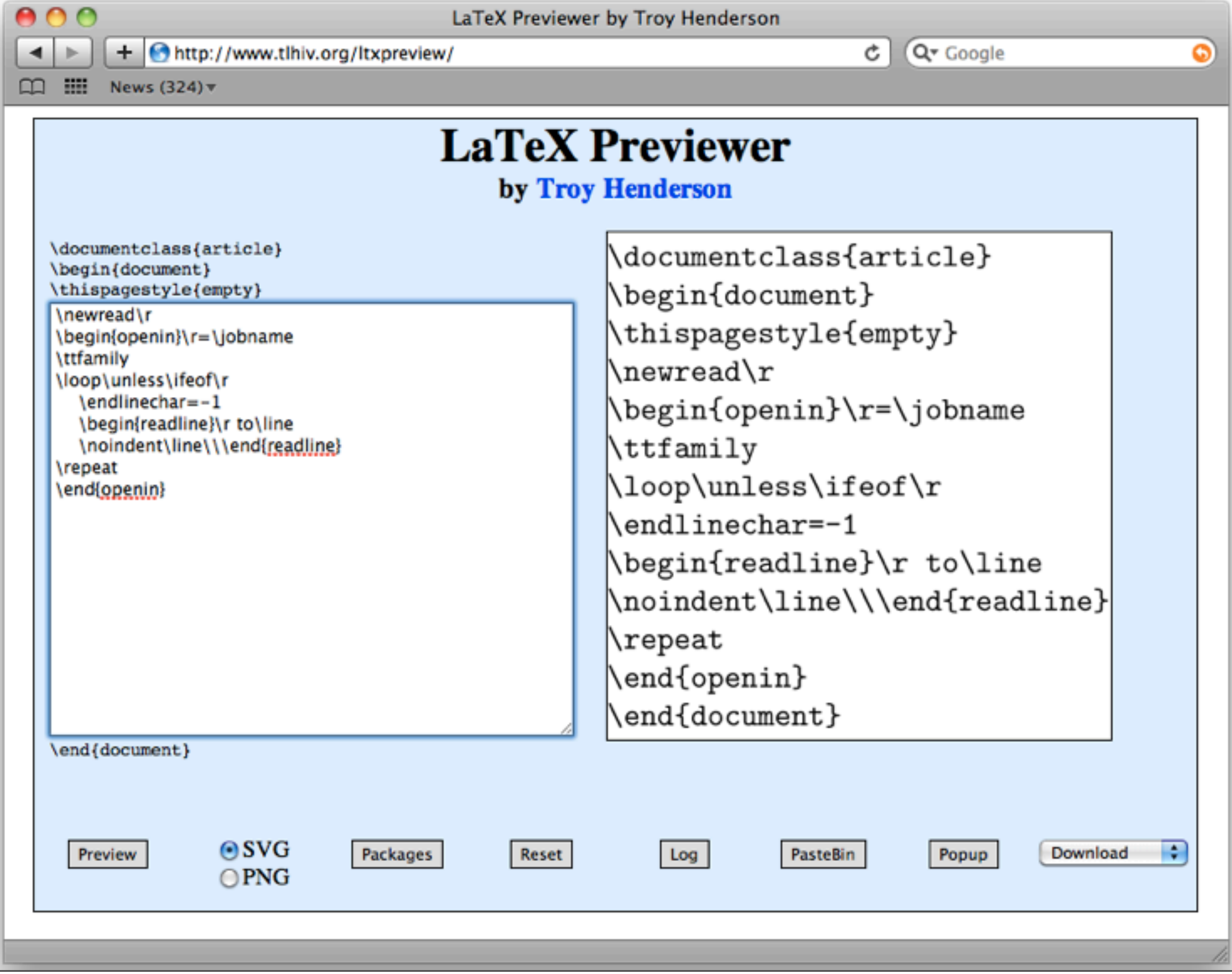
- Read sensitive files
  - `\input`, `\include`
  - `\read`, `\readline`
- Typeset data in output PDF

Input filtering

- Filter out dangerous control sequences
- Math mode

TEXniques to bypass filters

- Macros like `\input`
  - `\@input`, `\@iinput`, `\@input@`, `\@@input`
  - `\lstinputlisting`, `\verbatiminput`
- Bypass filters
  - `\csname`, `\begin`, `^^xy`, `\catcode`
- Escape math mode
  - `\end{eqnarray}`, `\end{align}`


LaTeX Previewer by Troy Henderson
http://www.tlhiv.org/ltxpreview/
Google
News (324)
Preview
Packages
Reset
Log
PasteBin
Popup
Download

LaTeX Previewer

by Troy Henderson

```
\documentclass{article}
\begin{document}
\thispagestyle{empty}
\newread\r
\begin{openin}\r=\jobname
\ttfamily
\loop\unless\ifeof\r
   \endlinechar=-1
   \begin{readline}\r to\line
   \noindent\line\\\end{readline}
\repeat
\end{openin}
\end{document}
```

```
\documentclass{article}
\begin{document}
\thispagestyle{empty}
\newread\r
\begin{openin}\r=\jobname
\ttfamily
\loop\unless\ifeof\r
\endlinechar=-1
\begin{readline}\r to\line
\noindent\line\\\end{readline}
\repeat
\end{openin}
\end{document}
```

● SVG
○ PNG

TEX's malleability

- Category codes control functionality
- Can be changed by `\catcode`

```
    \catcode`Z=0 ZTeX
```

An example: xii.tex

By David Carlisle

```
\let~\catcode~`76~`A13~`F1~`j00~`P2jdefA71F~`7113jdefPALLF
PA''FwPA;;FPAZZFLaLPA//71F71iPAHHFLPAzzFenPASSFthP;A$$FevP
A@@FfPARR717273F737271P;ADDFRgniPAWW71FPATTFvePA**FstRsamP
AGGFRruoPAqq71.72.F717271PAYY7172F727171PA??Fi*LmPA&&71jfi
Fjfi71PAVVFjbigskipRPWGAUU71727374 75,76Fjpar71727375Djifx
:76jelse&U76jfiPLAKK7172F71l7271PAXX71FVLnOSeL71SLRyadR@oL
RrhC?yLRurtKFeLPFovPgaTLtReRomL;PABB71 72,73:Fjif.73.jelse
B73:jfiXF71PU71 72,73:PWs;AMM71F71diPAJJFRdriPAQQFRsreLPAI
I71Fo71dPA!!FRgiePBt'el@ lTLqdrYmu.Q.,Ke;vz vzLqpip.Q.,tz;
;Lql.IrsZ.eap,qn.i. i.eLlMaesLdRcna,;!;h htLqm.MRasZ.ilk,%
s$;z zLqs'.ansZ.Ymi,/sx ;LYegseZRyal,@i;@ TLRlogdLrDsW,@;G
LcYlaDLbJsW,SWXJW ree @rzchLhzsW,;WERcesInW qt.'oL.Rtrul;e
doTsW,Wk;Rri@stW aHAHHFndZPpqar.tridgeLinZpe.LtYer.W,:jbye
```

Conclusions

- Binary/text distinction not a good classifier
- Arbitrary code execution
- Exfiltrate sensitive data

Questions?

Owning people through a typesetting language;
it seems unsporting, somehow. – Keaton Mowery