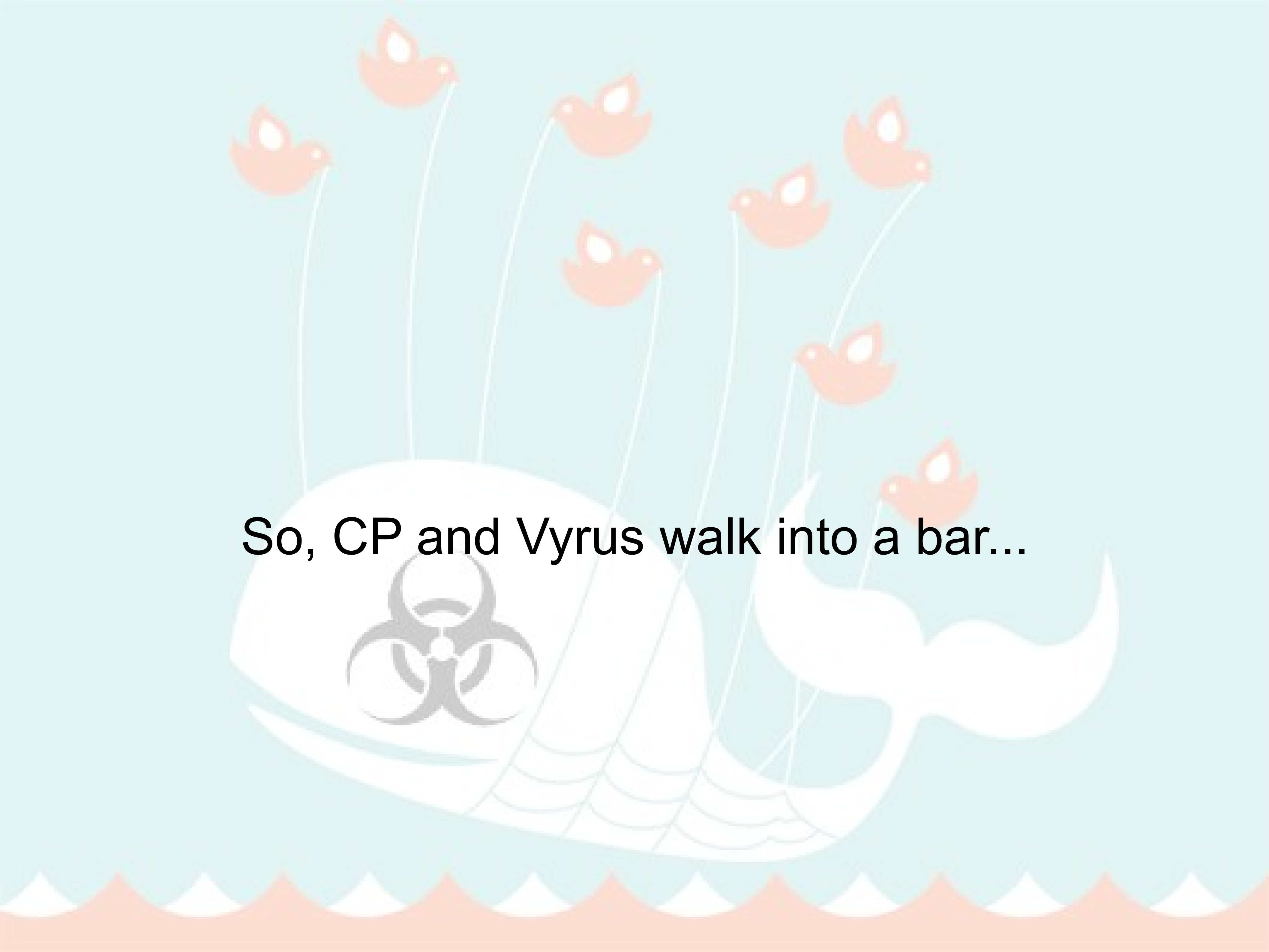
TwatFS

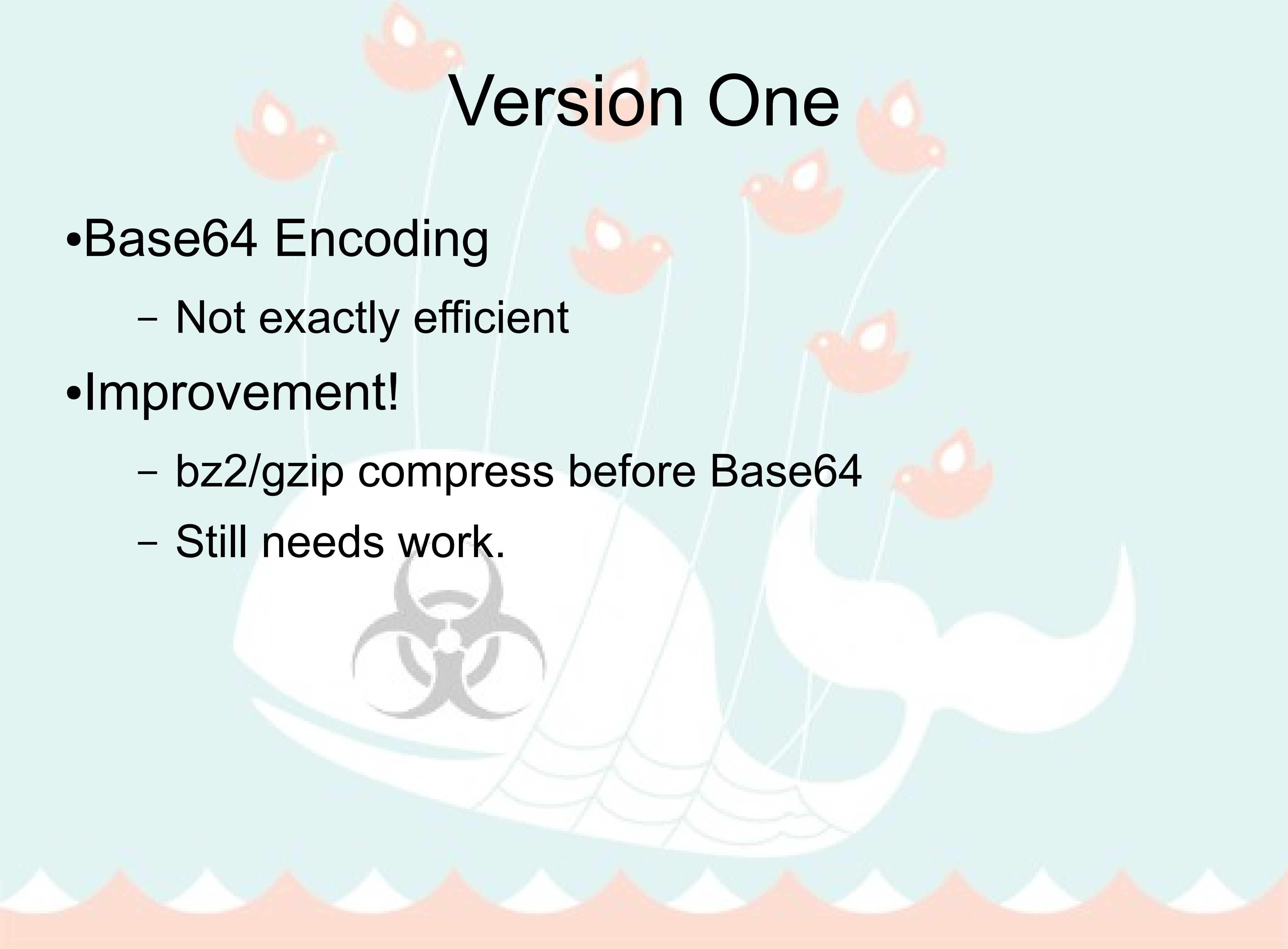
Surly Abuse of
Social Networking Bandwidth



So, CP and Vyrus walk into a bar...

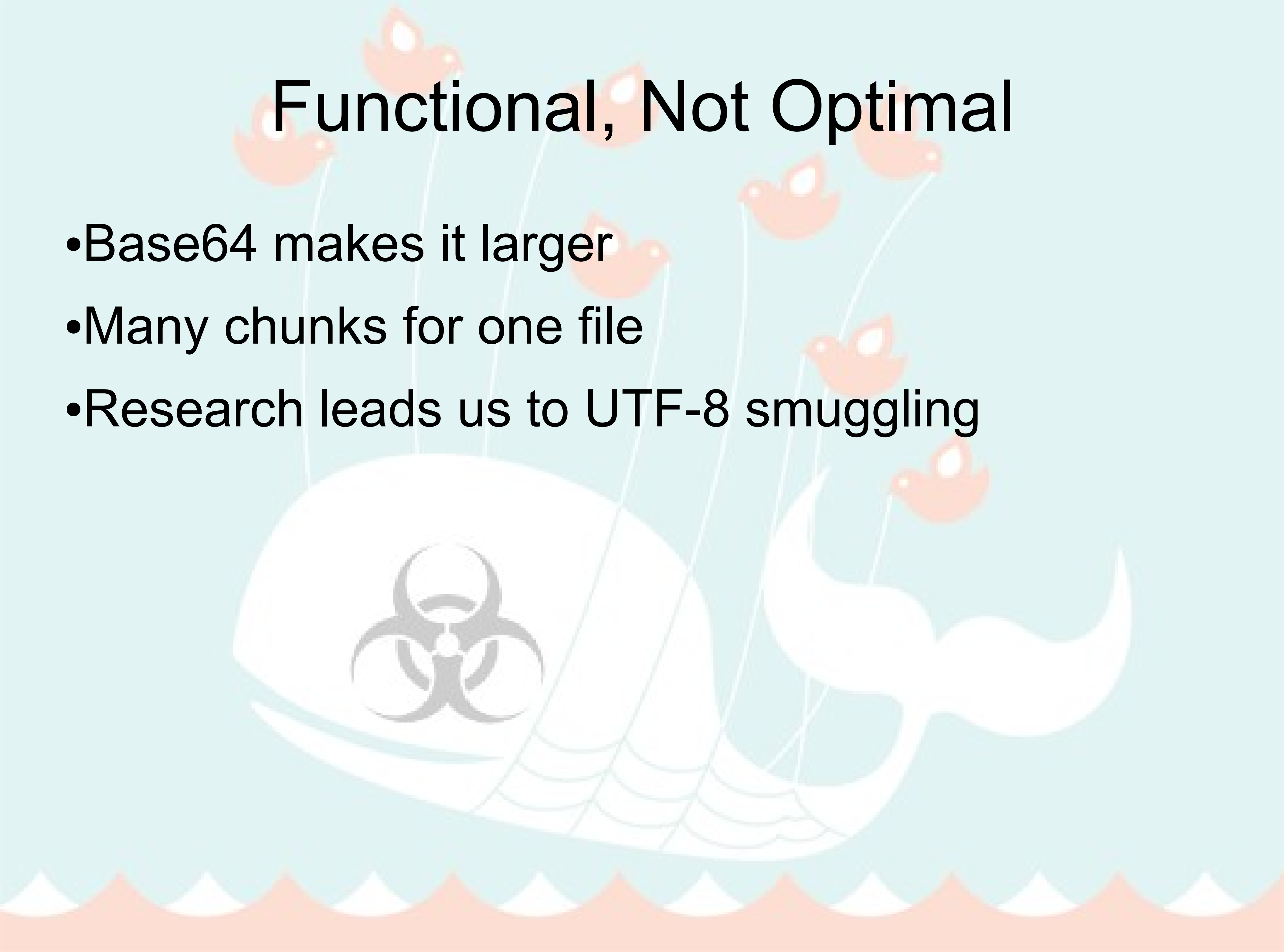
Version One

- Base64 Encoding
  - Not exactly efficient
- Improvement!
  - bz2/gzip compress before Base64
  - Still needs work.



Functional, Not Optimal

- Base64 makes it larger
- Many chunks for one file
- Research leads us to UTF-8 smuggling

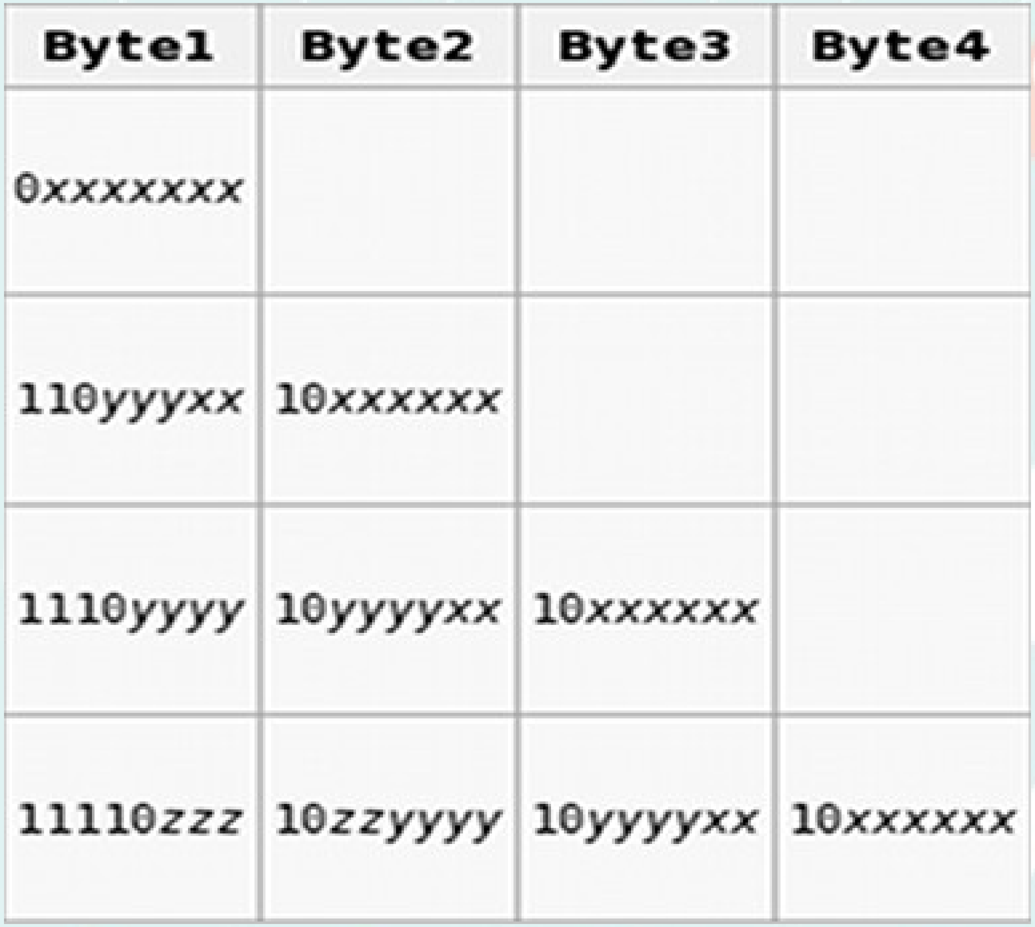


What Is UTF-8?

- Implementation of Unicode
- Variable number of bytes
- Length indicated by bit patterns
- Can encode from U+0000 through U+10FFFF
- For more info, see RFC 3629

Encoding Scheme

- Encode in 20 bit blocks



Byte1	Byte2	Byte3	Byte4
0xxxxxxx			
110yyyxx	10xxxxxx		
1110yyyy	10yyyyxx	10xxxxxx	
11110zzz	10zzyyyy	10yyyyxx	10xxxxxx

7 bits

11 bits

16 bits

21* bits

Version 2: Naïve UTF-8 Encoding

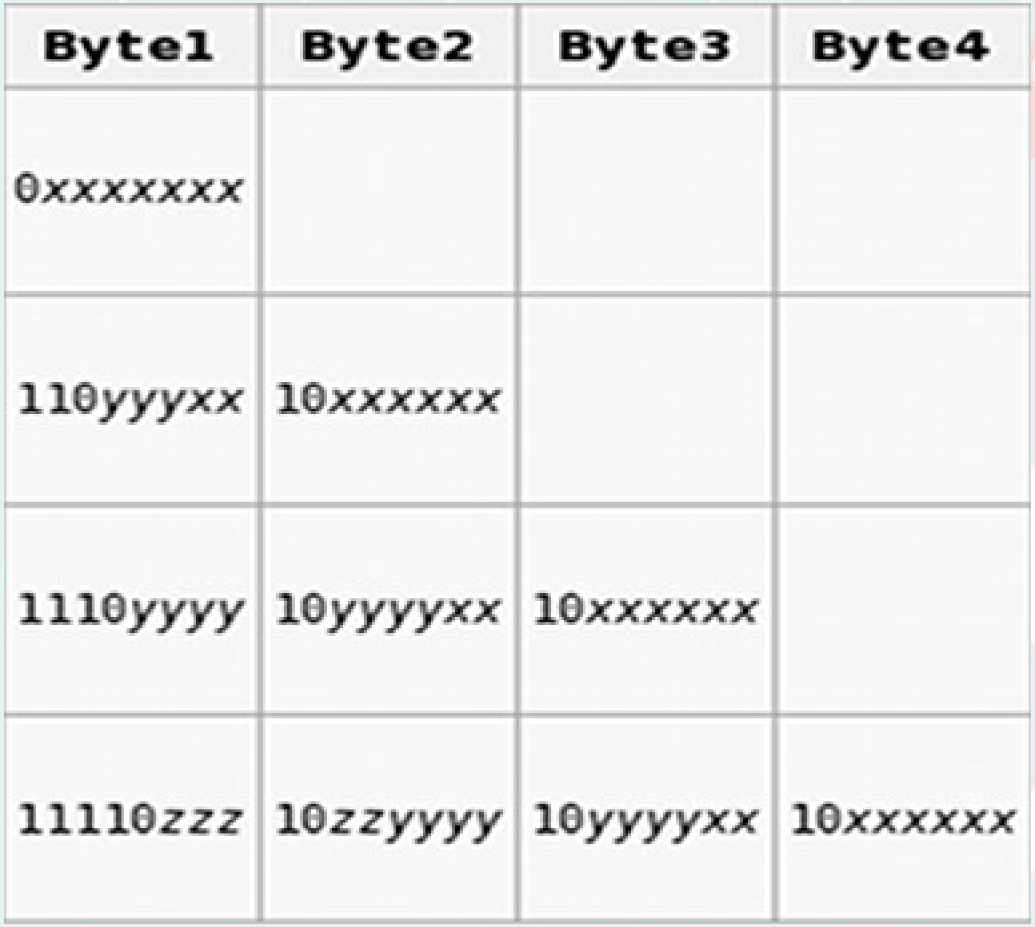
- Compressed and Base64 Encoded Data
- Shifted values to fit within 6 bits
- Encoded 3 “data chars” into each UTF-8 char
- UTF8 Chars = (Compressed_bytes * 4/3) / 3

Version 3: Proper UTF-8

- Base64 increases size by 33%
- Direct encoding of binary
- 20 bits of real data per UTF-8 character
  - Read bytes in 20-bit chunks
  - Convert to integer value
  - Insert into UTF-8 character

Encoding Scheme

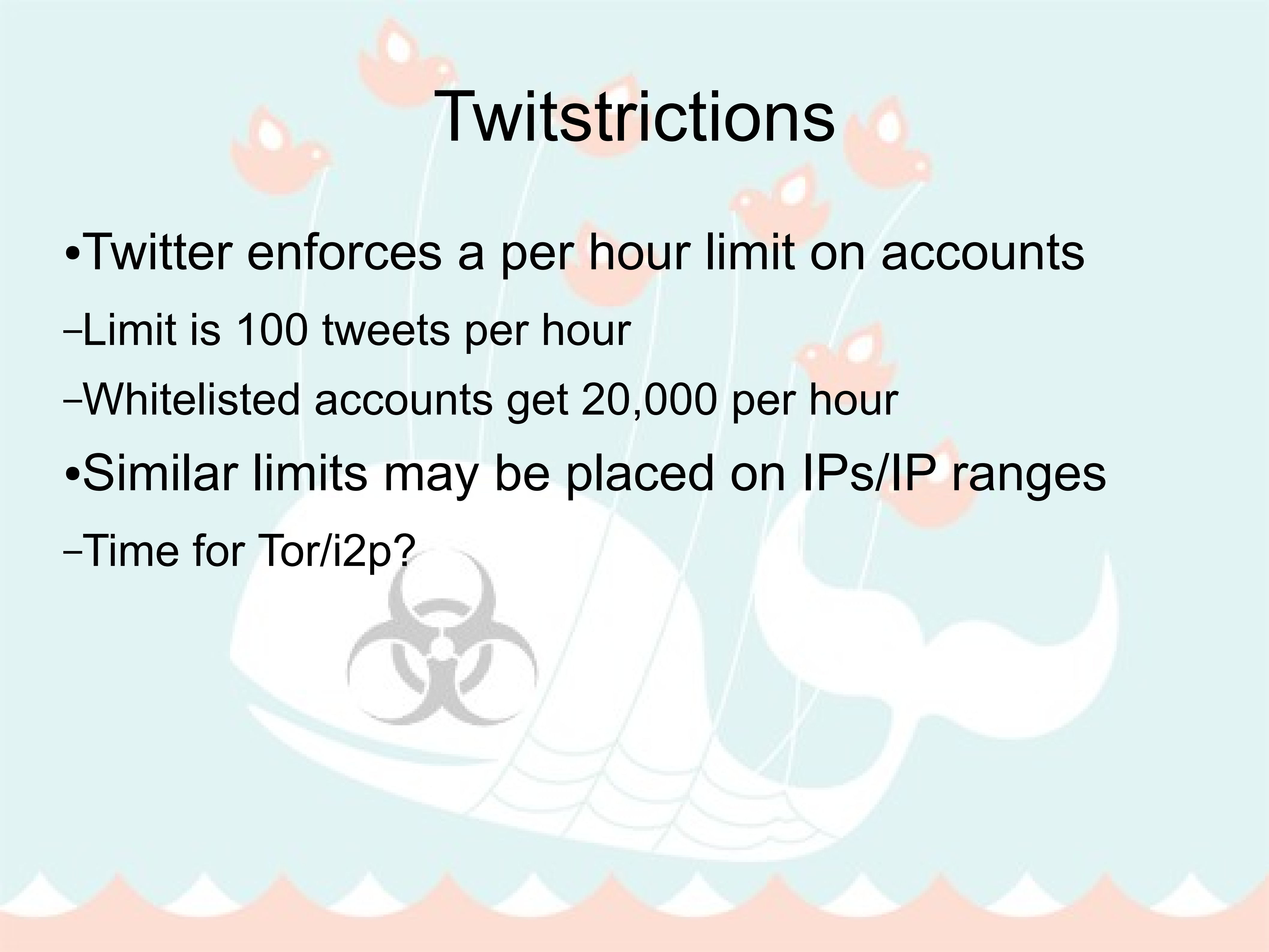
- Encode in 20 bit blocks



Byte1	Byte2	Byte3	Byte4	
0xxxxxxx				0 - 127
110yyyxx	10xxxxxx			128 - 2047
1110yyyy	10yyyyxx	10xxxxxx		2048 - 65535
11110zzz	10zzyyyy	10yyyyxx	10xxxxxx	65536 - 1048576

Data Comparison

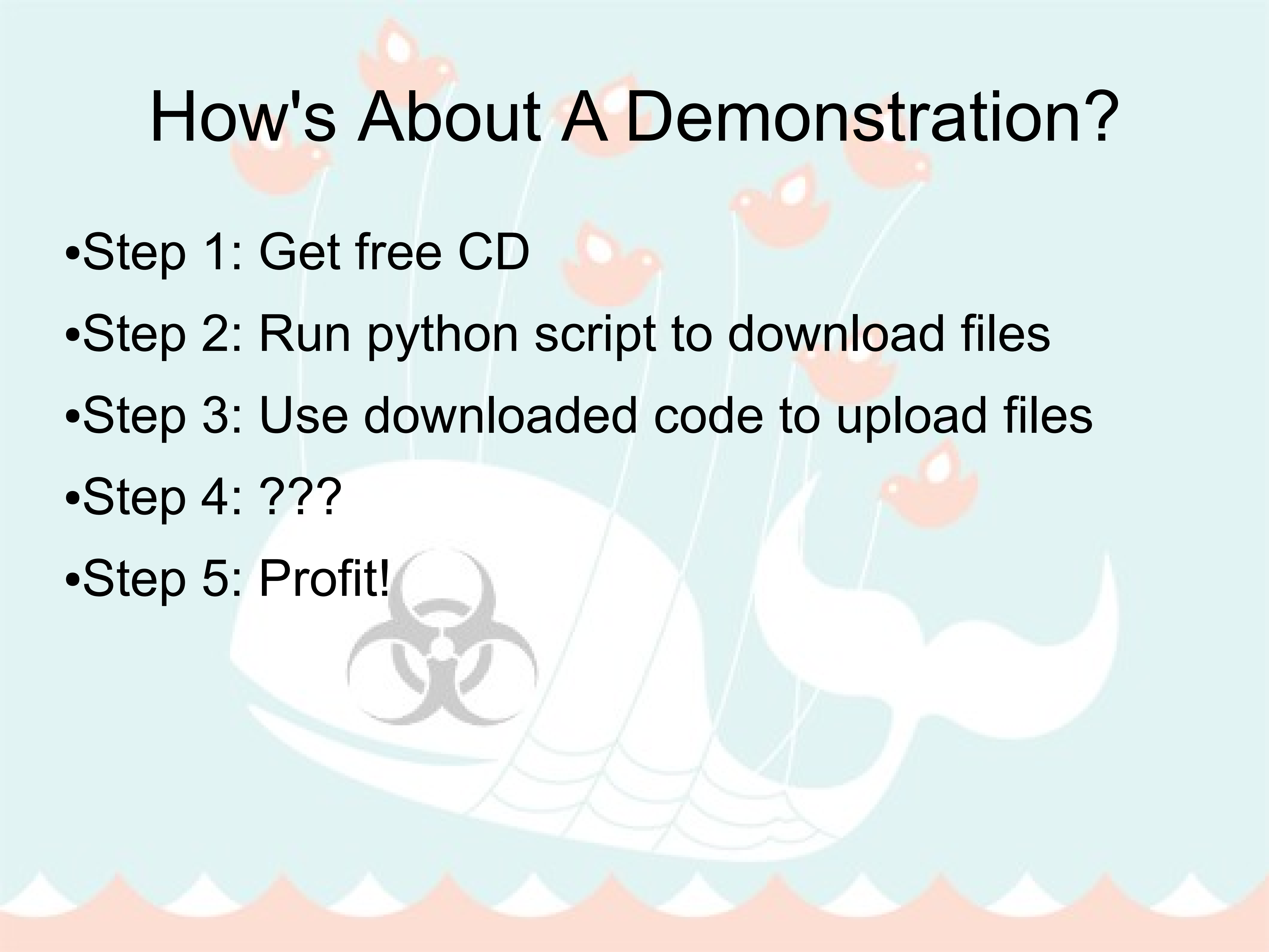
Transformation	Compressed	B64 Chars	UTF-8 Chars
b64	2048	2731	2731
b64+UTF-8	2048	2731	911
Pure UTF-8 – BCS	2048	N/A	820
Pure UTF-8 – WCS	2048	N/A	861



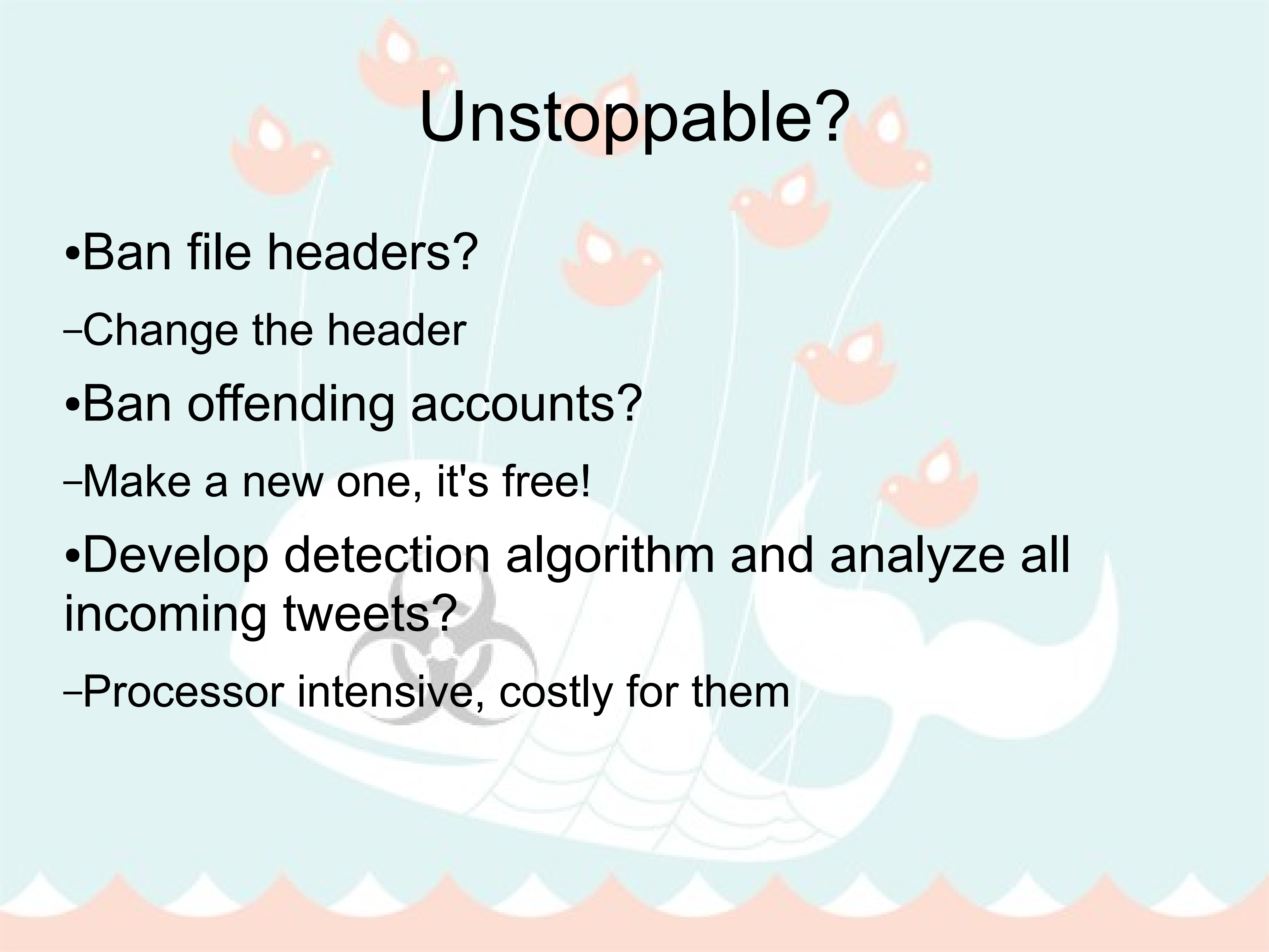
Twitstrictions

- Twitter enforces a per hour limit on accounts
  - Limit is 100 tweets per hour
  - Whitelisted accounts get 20,000 per hour
- Similar limits may be placed on IPs/IP ranges
  - Time for Tor/i2p?

How's About A Demonstration?



- Step 1: Get free CD
- Step 2: Run python script to download files
- Step 3: Use downloaded code to upload files
- Step 4: ???
- Step 5: Profit!

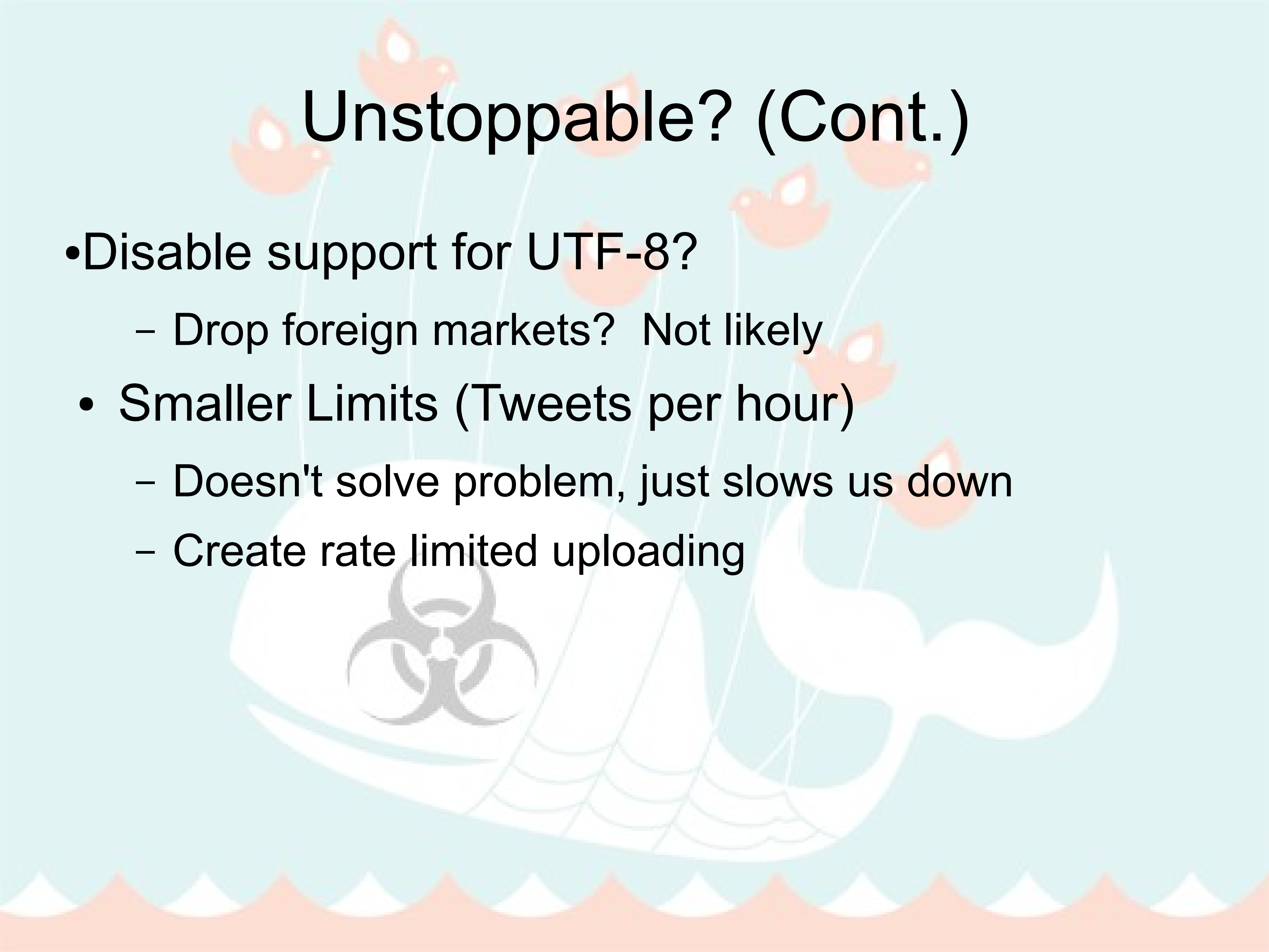


Unstoppable?

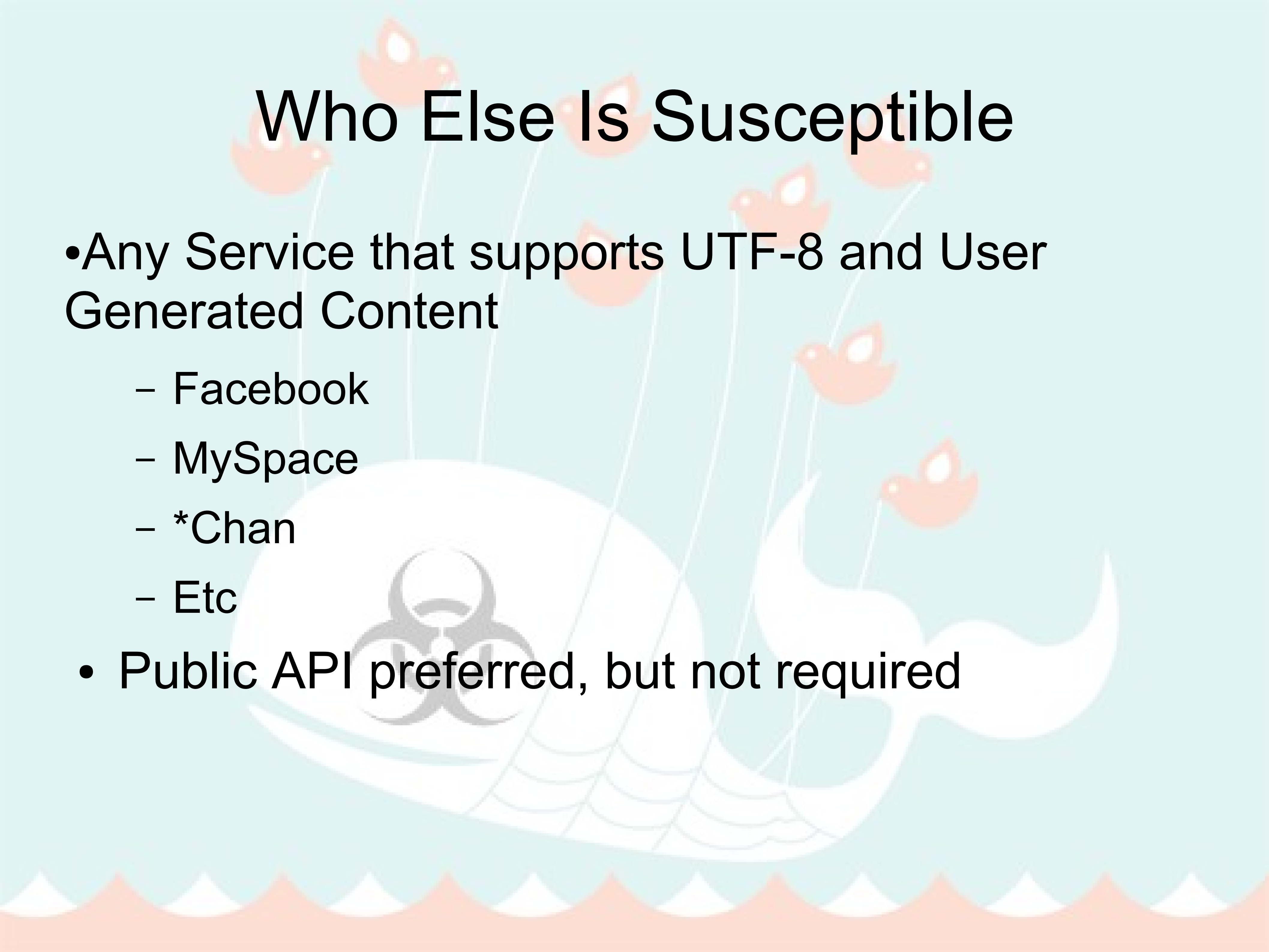
- Ban file headers?
  - –Change the header
- Ban offending accounts?
  - –Make a new one, it's free!
- Develop detection algorithm and analyze all incoming tweets?
  - –Processor intensive, costly for them

Unstoppable? (Cont.)

- Disable support for UTF-8?
  - Drop foreign markets? Not likely
- Smaller Limits (Tweets per hour)
  - Doesn't solve problem, just slows us down
  - Create rate limited uploading

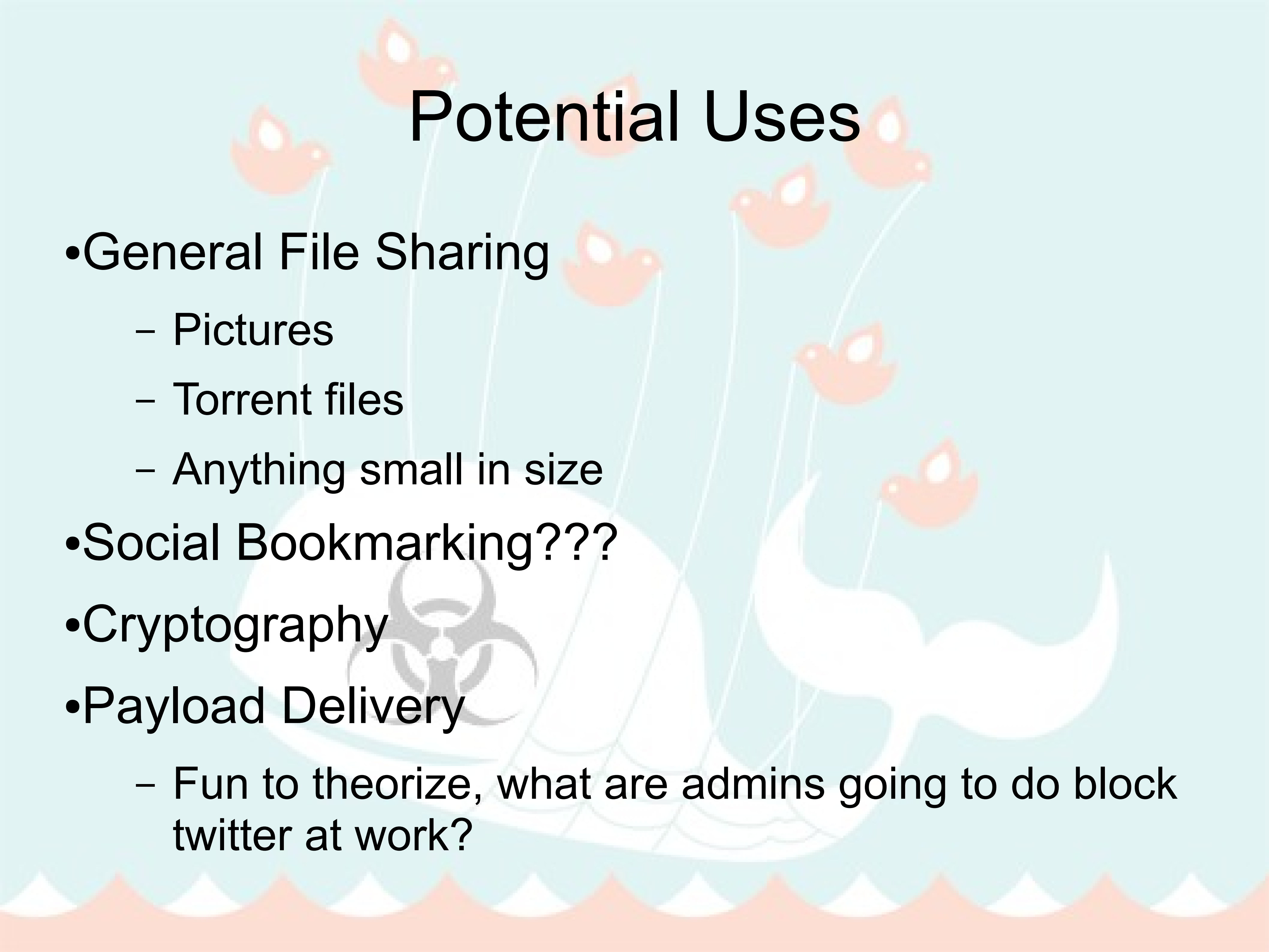


Who Else Is Susceptible



- Any Service that supports UTF-8 and User Generated Content
  - Facebook
  - MySpace
  - *Chan
  - Etc
- Public API preferred, but not required

Potential Uses



- General File Sharing
  - Pictures
  - Torrent files
  - Anything small in size
- Social Bookmarking???
- Cryptography
- Payload Delivery
  - Fun to theorize, what are admins going to do block twitter at work?

TwatFS=end;file=presentation.odp;comp=bz2;chunks=949;enc=utf8;

- Email
  - TwatFS@dc949.org
- Twittersesz
  - cps_rants
  - franksquared
  - Vyrus001