

株式会社 ITS MORE

2020年4月始動

2020年4月23日 投稿者: YSATO@DELEGATE.ORG

data schemeは健在だった

Macに触れたのがきっかけで、つい昔の事を思い出してしまう。URLと言う表現形式の成長期に提案された dataと言うスキームもそのひとつ。

URLは http や ftp と言ったプロトコル名（つまりアクセス手段）と、//ホスト名/パス名、という「リソースのロケーション」つまりアクセス先のアドレスを表す表記法（scheme）だ。その中のひとつとして、リソースのデータ自体をURLに表記する方法として提案されたのが「data」だった。似たものに、抽象的にメールアドレスを表記する mailto（具体的な転送手段を規定していない）があり、現在広く使われている。

CPUの命令アドレスフィールドに即値（immediate value）を持たせられるのは当たり前だし、その他の類想からも、コンピュータ分野で育てば当然考えられるべき事なのだが、当時はなるほど（膝をポン）と思った記憶がある。一般的にも「名は体を表す」と言うが、体自体を名前にしてしまうジョークや商品名は昔からある。data scheme はまさにそれに相当する。

以下に具体例を示す。上がデータ（gifイメージ）の生表現で、下がそれをブラウザが解釈して表現した結果だ（data scheme 提案者のLarry Mainster 氏のポートレート）。

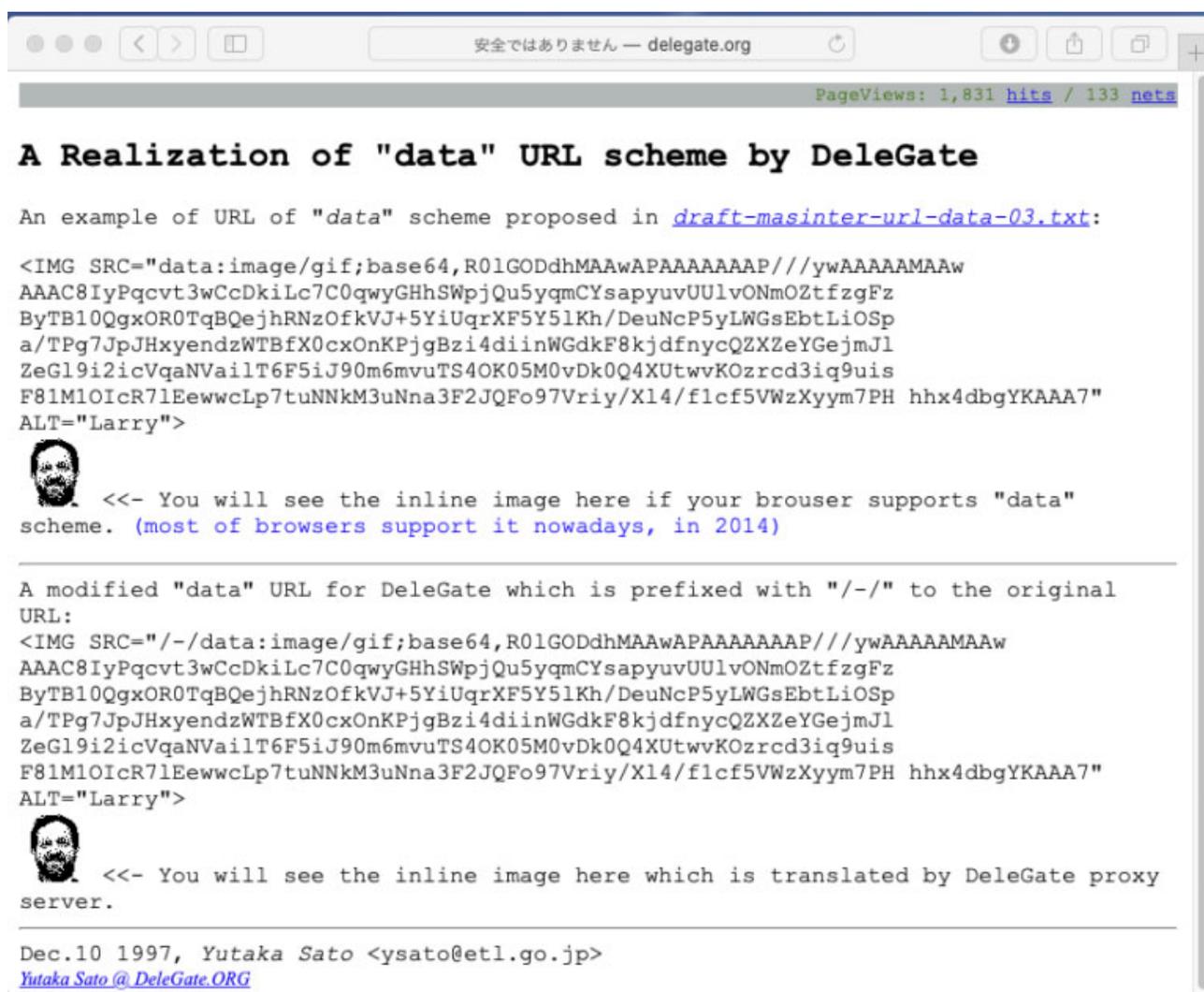
```
<IMG SRC="data:image/gif;base64,R0lGODdhMAAwAPAAAAAAP///ywAAAAAMAAwAAAC8IyPqcvt3wCcDkiLc7C0qwyGHhSWpjQu5yqmCYsapyuvUUVONmOZtfzgFzByTB10QgxOR0TqBQejhRNzOfkVJ+5YiUqrXF5Y5lKh/DeuNcP5yLWGsEbtLiOSp a/TPg7JpJHxyendzWTBfX0cxOnKPjgBzi4diinWGdkF8kjdfnycQZXZeYGejmJlZeGI9i2icVqaNVailT6F5iJ90m6mvuTS40K05M0vDk0Q4XUtwvKOzrcd3iq9uisF81M10Icr7IEewwLp7tuNNkM3uNna3F2JQFo97Vriy/Xl4/f1cf5VWzXyym7PH hxx4dbgYKAAA7" ALT="Larry">
```



ここにはきっちりと、当時確立していたMIMEのデータ型名と、base64エンコーディングが、やはりWWWの初期に確立していたHTMLのSRCタグの中に表現されている。だから、それら極めて基本的な標準仕様に準拠し、gifを表示できるブラウザなら、必ず表示できる表記だった（はずだったが、多くのブラウザの実装ではそうではなかった…）。

data scheme には実際、特に当時の未成熟なブラウザの状況もあり、色々な応用が期待された。data: が提案されたのは1997（翌1998年にRFC2397となった）。私はこれを使って、ブラウザによってサポートされていない文字や画像を埋め込む事を考えた。当時書いたデモ用のページが今も残っている（というか、動態保存している 笑）

URL: <http://www.delegate.org/delegate/sample/data-url.shtml> ↓

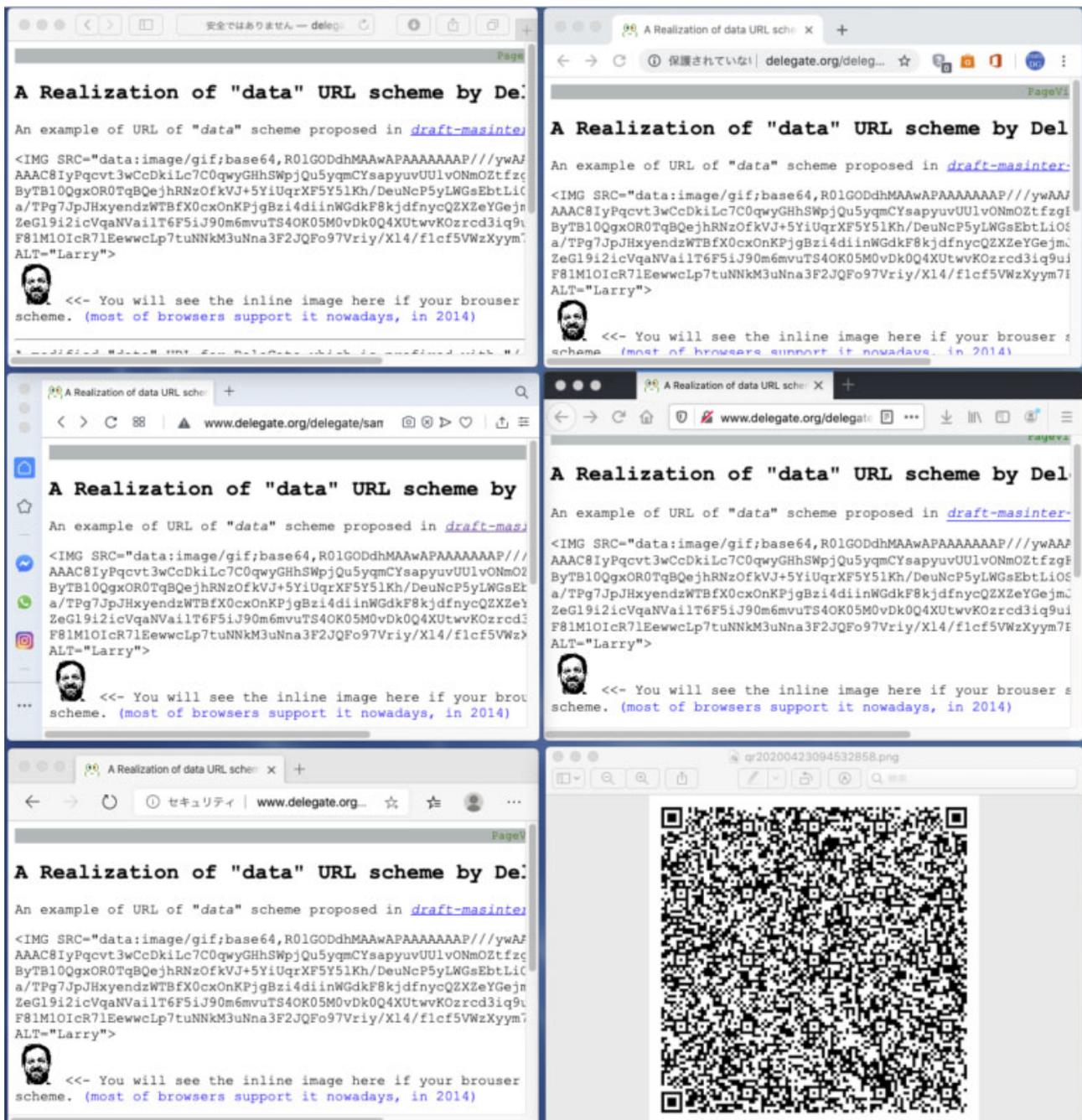


これは、ブラウザローカルに解釈できる事がメリットのはずの data: を実装していないブラウザのために、リモートの http: サーバでそれを補ってやろうという、分かってはいた

が本末転倒の試みであった（実際、ウケなかった）。そして、data がきちんと実装されるより前に、ブラウザの多言語対応や各種イメージ型の表示能力が達成され、当時考えた応用の目的は失われた。

その後の2014年の追記にあるように、やがてほとんどのブラウザがこれを実装した。ただ、私の偏見から来る記憶違いかもしれないが、MSのIEの対応は遅かったように思う。

以下は2020年4月現在の、5大メジャーブラウザでの data scheme の表示。いずれもちゃんと表示できている。



右下は昨今流行りのQRコードでこの data を表現したもの。残念ながら、iPhoneのカメラアプリは、これを認識してくれなかった（笑）

ここまで長く、これを書いたのは、data scheme という歴史に埋もれ忘れ去られた表記を、今も全ての主要ブラウザが解釈できるという事実に基づく、新しい応用があると考えているからだ。応用するのは、良い人かも知れないし、悪い人かも知れないが…

たとえば、文字にしてもUnicodeでは足りないケースも多いはずだ。そういうニッチは今後も完全に無くなる事は無い。ニッチをグローバルなインフラからサポートするのはコストに合わない。

インターネットは100倍早くなったかも知れないが、一方でユーザは100倍短気になった。それに、いくらネットワークの片道スループットが上がったとしても、往復レスポンスは光の速さを超えられない。途中の停車駅だってそうは減らせないだろう。実際、ギガビットネットのRTTは当時の10倍程度しか高速していない、ミリ秒の単位のような。

ウェブには1ページを表示するのに数十回のHTTPリクエストが発生するのはざらにある。細々としたイメージデータを多数取得するためである場合が多い。それらは、転送時に data: にして埋め込んでやれば、一発転送で終わる。（いや、それをDeleGateにやらしてもらふ必要は無いけど 笑）。同じデータを繰り返し送るとい無駄は発生するが、データの素性を知っていれば自動的・機械的なトレードオフはできる。それは長期的に保存するにしても、埋め込み型が有用なケースも多いだろう。

data scheme を IMG SRC 以外でも使えたとしたら、何が（何か）起こるか。応用は、全てをひとつのHTMLファイルに押し込む手段としてだけでは無い。MIME multipart 型と同様、古びた骨董の中に新しい可能性が埋もれている可能性はある。普通に普及しているアーカイブ形式をファイルシステムとして解釈する手法との優劣かなと思う。形式的には格好良くは無い。XMLが、ASN.1 があると言う人もいるだろう。しかし実現上の強みは、すでに data scheme は（長い時間をかけて）実装普及が完了していて準備OKであり、トランスペアレントな移行が可能な事だ。

…と、つい昔のエキサイティングな時代を思い出してエキサイトしてしまった…

インターネットに破れて死んだOSIが電子証明書の中に残して行った形見のような X.509 を思いつつ。

2020-0423 sato@izmoh

