InkML Activity

Stephen M. Watt U. Western Ontario
Sriganesh Madhvanath HP
on behalf of the W3C MMI Pen Input Group

ICDAR 2005, Seoul Korea, August 31

- With particular thanks for contributions to
  - Yi-Min Chee (IBM)
  - Jose-Antonio Magaña (HP)
  - Werner Krandick (Drexel)

  for materials.

Outline

- ~14:00 Overview and Current Status of InkML
- ~14:30 InkML Applications
- ~14:50 Discussion of Community Needs
- ~15:20 Summary and Conclusions

Overview and Current Status of InkML

Overview of InkML Activity

- **World Wide Web Consortium** W3C
- **Multimodal Interaction Working Group** MMI
  - Standards for web pages you can speak to, gesture at …
  - Chartered Feb 2002
  - Extensible Multimodal Annotation (EMMA) ML
  - MMI Architecture
  - MMI Interface
  - DOM
  - …
  - **Pen Input Modality**

Pen Input Subgroup

- **Contributors to Current Draft**

  Yi-Min Chee, IBM
  Jose-Antonio Magaña, HP
  Katrin Franke, Fraunhofer Gesellschaft
  Max Froumentin, W3C
  Gregory Russell, IBM
  Sriganesh Madhvanath, HP
  Giovanni Seni, Motorola
  Christopher Tremblay, Corel
  Stephen Watt, University of Western Ontario
  Larry Yaeger, Apple

InkML

- An XML markup for digital ink data
  - Ink data captured by pen-enabled system
    e.g. Digitizing pad, Tablet PC, PDA, Anoto Pen, …
  - Information about pen movement
    e.g. x,y,z coordinates, angles, force, …
  - Device information
    e.g. sample rate, resolution, cross coupling, …

InkML Goals

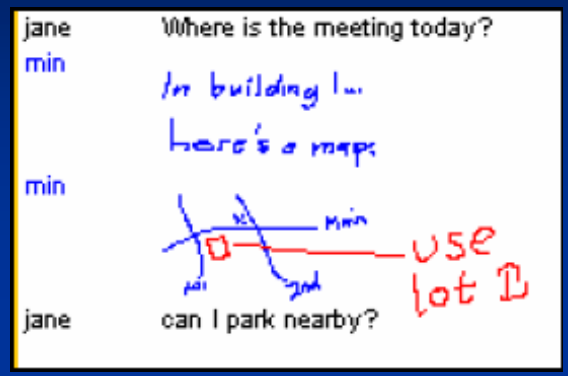
- Design is open
  - Developed in public process
  - Extensible by inclusion in, or annotation with, other XML or textual formats
- General purpose
  - Streaming ink from device
  - Archival ink
  - Structured, annotated results from analysis

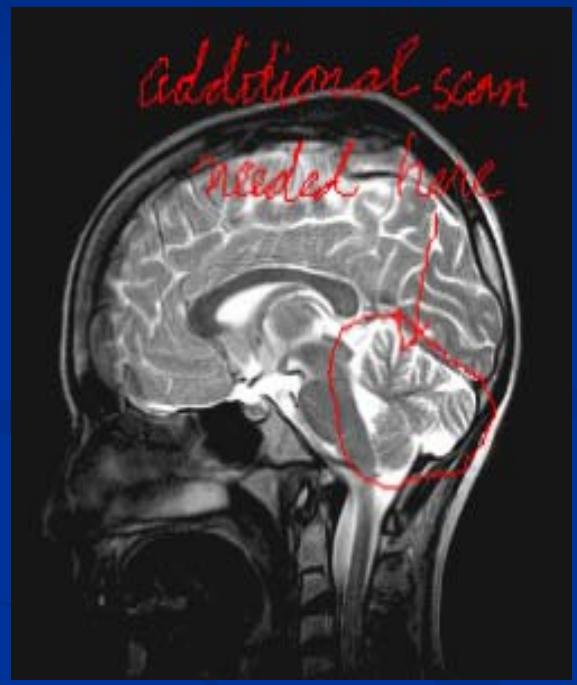
InkML Motivation

- Growth in pen-enabled devices
- Multimodality
- *versus* special purpose, proprietary formats

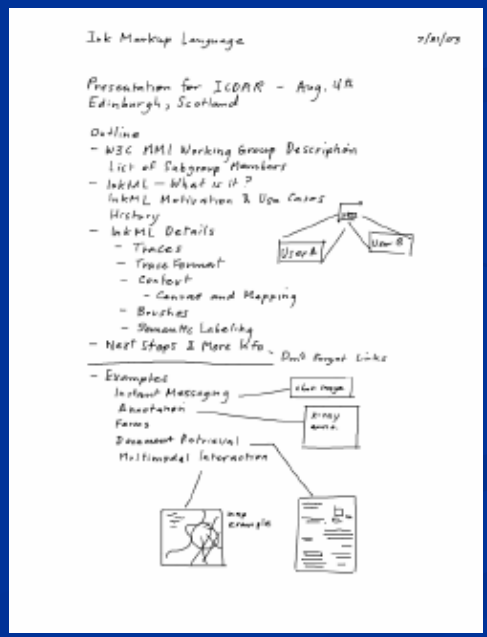
Simple Use Cases

- Ink Messaging





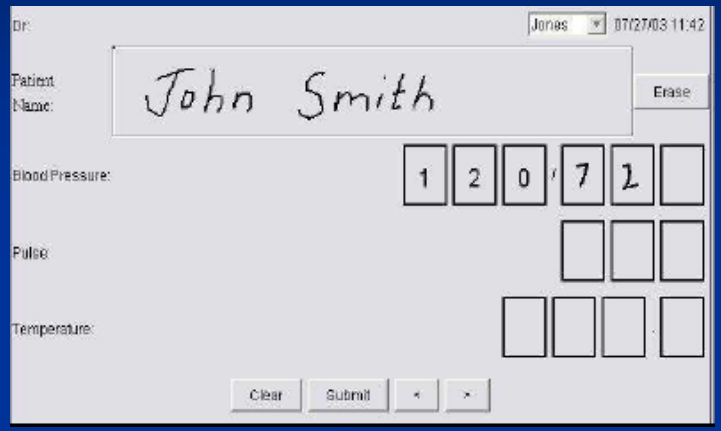
- Annotation



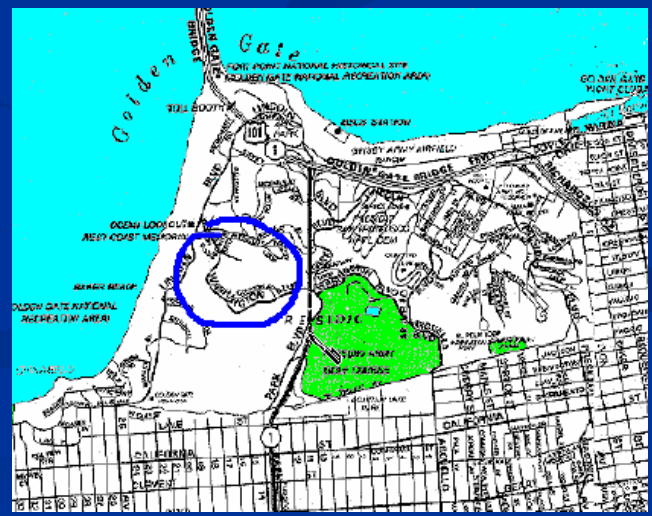
- Archival

Simple Use Cases

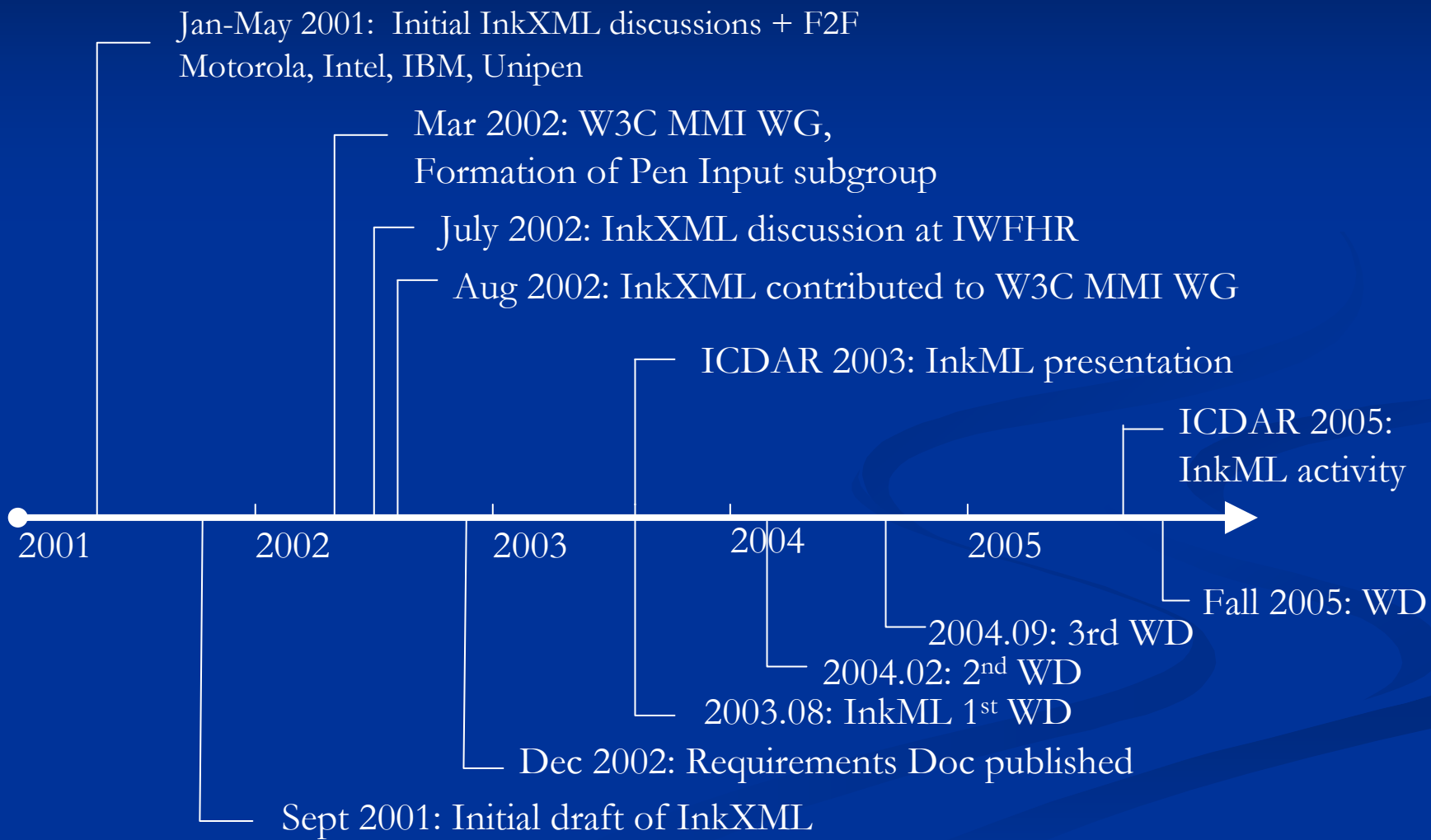
- Form Filling



- Multimodal systems
  (e.g. pen synchronized with voice)

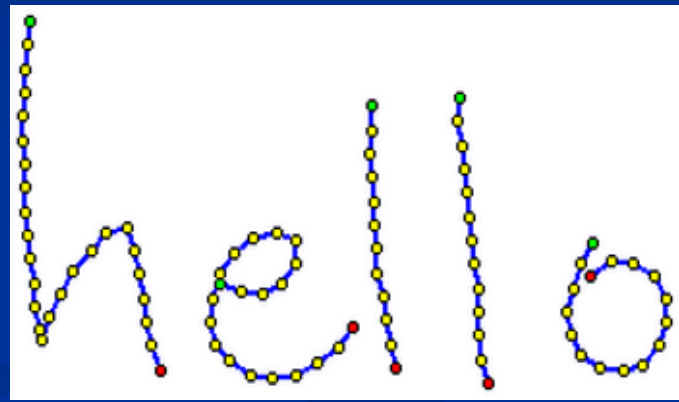


A Brief History



“Hello World” in InkML

```
<ink>
   <trace> 10 0 9 14 8 28 7 42 6 56 6 70 8 84 8 98
      8 112 9 126 10 140 13 154 14 168 17 182 18 188
      23 174 30 160 38 147 49 135 58 124  72 121 77
      135 80 149 82 163 84 177 87 191 93 205
   </trace>
   <trace> 130 155 144 159 158 160 170 154 179 143
      179 129 166 125 152 128 140 136 131 149 126 163
      124 177 128 190 137 200 150 208 163 210 178 208
      192 201 205 192 214 180
   </trace>
   <trace> 227 50 226 64 225 78 227 92 228 106 228 120
      229 134 230 148 234 162 235 176 238 190 241 204
   </trace>
   <trace> 282 45 281 59 284 73 285 87 287 101 288 115 290
      129 291 143 294 157 294 171 294 185 296 199 300 213
   </trace>
   <trace> 366 130 359 143 354 157 349 171 352 185 359
      197 371 204 385 205 398 202 408 191 413 177 413 163
      405 150 392 143 378 141 365 150
   </trace>
</ink>
```



Traces

- Represented by **<trace>** element in InkML
- Describes a sequence of sequential points
  state penUp, penDown, indeterminate, continuation
- Choice of coordinates may be specified
  regular or intermittent
- Given as values, differences or 2nd differences
- May be grouped and structured in various ways

Writing Context

- **<context>** Information to interpret traces
  - **<brush>** named brushes
  - **<traceFormat>** format of data within traces
  - **<captureDevice>** characteristics of the device
  - **<canvasTransform>** trasforrmation to canvas

<brush>

- Application-specific attributes, e.g.
  - “highlighter” “eraser” “spraypaint” “toothpaste”
  - “red” “green” “blue”
  - “Althea” “Ian” “Isaac” “Lori”
- Associated with trace at time of capture

<traceFormat>

- Specifies channels and their order in `<trace>`
  default x y: x1 y1 x2 y2 x3 y3 …
- Regular vs intermittent
- Predefined vs user-defined
- Can give mapping from digitizer to data units

Predefined Channel Names

Name	Interpretation
X	X coordinate (horizontal pen position)
Y	Y coordinate (vertical pen position)
Z	Z coordinate (height of pen above writing surface)
F	pen tip force
S	tip switch state (touching/not touching the writing surface)
$B_1 \ldots B_n$	button states
Tx	tilt along the x-axis
Ty	tilt along the y-axis
A	azimuth angle of the pen (yaw)
E	elevation angle of the pen (pitch)
R	rotation about pen axis (roll)
T	time

Example

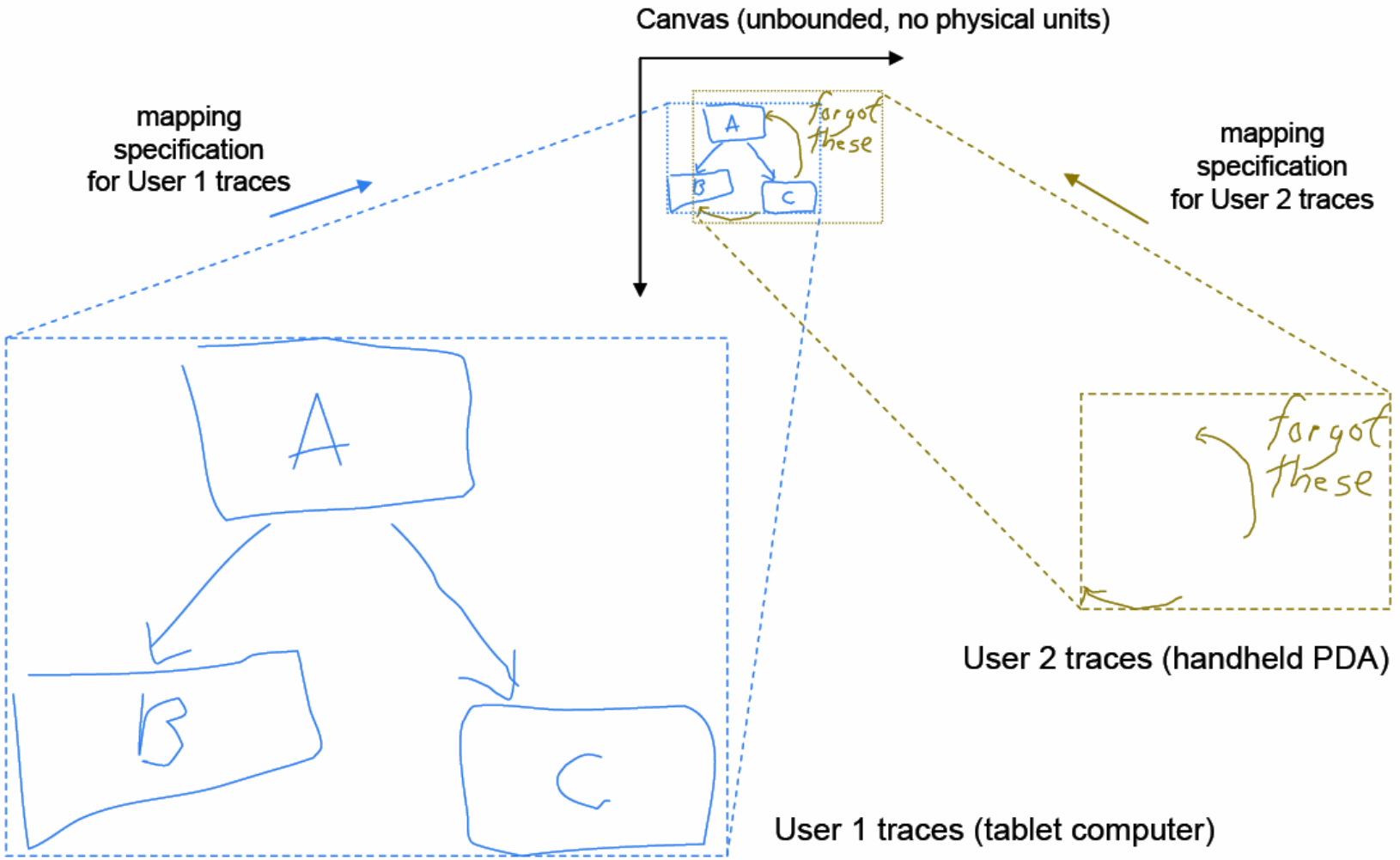
```
<traceFormat>
   <regularChannels>
     <channel name="X" type="decimal"/>
     <channel name="Y" type="decimal"/>
   </regularChannels>
   <intermittentChannels>
     <channel name="B1" type="boolean" default="F"/>
     <channel name="B2" type="boolean" default="F"/>
   </intermittentChannels>
 </traceFormat>

 <trace id="id4525abc">
   1125 18432'23'43"7"-8 3-5 7 -3 6 2.5 6.1 8 3 6:T;
   2 4:*T; 3 6 3-6:FF;
 </trace>
```

<mapping>

- Store raw digitizer coordinates
  fast recording, high fidelity, *or*…
- Store virtual coordinates
  ink from multiple sources

- Can be used with any channel (e.g. force, time)
- Contextual mappings
  channel cross-coupling
  shared virtual canvas

Canvas Transformations



<captureDevice>

- Describes characteristics of the ink digitizer
  - Manufacturer, Model, …
  - Sample rate, uniformity, …
  - Channels: name, type, characteristics
    - Resolution, quantization, noise, cross-coupling, skew, dynamic distortion …
- May be provided in-line or by reference to a library of device descriptions

Grouping

- Traces may be grouped into complex structures
- Any desired purpose:
  - To share context, semantics
  - Composite objects, analysis results, truth annotation
- **`<traceGroup>`** build structures
- **`<traceView>`** build structures by reference

Trace Group

```
<traceGroup>
    <traceGroup>
        <trace> ... </trace>
        <trace> ... </trace>
    </traceGroup>
    <traceGroup>
        <trace> ... </trace>
        <trace> ... </trace>
    </traceGroup>
</traceGroup>
```

Trace View

- Allows nested, overlapping, or partial trace groupings
- Allows objects to be built from traces in other files
-

```
<traceView>
    <traceView>
        <traceView traceRef="#stroke1"/>
        <traceView traceRef="#stroke3"/>
    </traceView>
    <traceView traceRef="#housePicture"
        from="1:3:20" to="1:5:400"/>
</traceView>
```

Semantic Labelling

- **`<metadata>`** XML annotation.
  E.g. recording session characteristics (a la Unipen)
- **`<desc>`** Text annotation.
- **`contentCategory="C1/…/Cn"`**
  Describe content of particular traces or groups of traces

```
<traceView contentCategory="Text/en"
           traceRef="#para1" …>
```

Streaming vs Archival Ink

- Stylized InkML suitable for different applications.
- Streaming:
  - <trace> data and context switches sent as collected
  - Expect continuation traces and backward references
  - E.g. Collaboration via smart whiteboards
- Archival:
  - Expect rich use of references to definitions
  - More assembly, grouping, annotation
  - E.g. Document storage and retrieval

Open Issues

- Few remain
- Unify common aspects of `<traceFormat>` and `<captureDevice>` ?
- Simpler, more uniform, use of mappings ?
- Richer or minimalistic **`contentCategory ?`**

Next Steps

- For further information see
  `http://w3.org/2002/mmi/ink`
- Provide feedback
  - Talk to SriG and myself here at ICDAR.
  - Email **<www-multimodal@w3.org>**
- Expect (last?) working draft this fall.
- If you have trial applications let us know.

InkML Applications

Example Applications

1. Pen-based mathematics
2. Smartphone ink collaboration
3. Multi-channel document processing
4. Rep. and annotation of online handwriting

Components for Pen-Based Mathematical Interfaces

Stephen M. Watt *et al*

University of Western Ontario

watt@uwo.ca

Western 125 YEARS est. 1878

ORCCA Ontario Research Centre for Computer Algebra

Ausdruck unter der Klammer rechts

E_L des bewegten Massenpunktes

ergie.

erhält für

$$E_L = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad (28)$$

Long-Term Goals

- Enter and manipulate math naturally by pen
- Support high-powered math transformations
- Support collaboration
- Do so portably, across applications and platforms

Approach

- Architect for a large problem,
  with many interacting components

- Recognize that each component requires research related to mathematical nature

- Use CA for expression transformation
- Recognize on-going hardware evolution

- Project with Western, Waterloo, Maple, Microsoft

Pen-based Mathematics System

- Mathematics input:
  -- character recognition,
  -- layout parsing,
  -- linear parsing

- Mathematics editing:
  -- subexpression selection,
  -- searching and linking,
  -- expression re-arrangement,
  -- expression transformation
  e.g. **`expand(sin(a+b))`**
  or **`factor(p)`**

- Sketching

- Re-winding and re-playing derivations

- Visual scenario/case organization

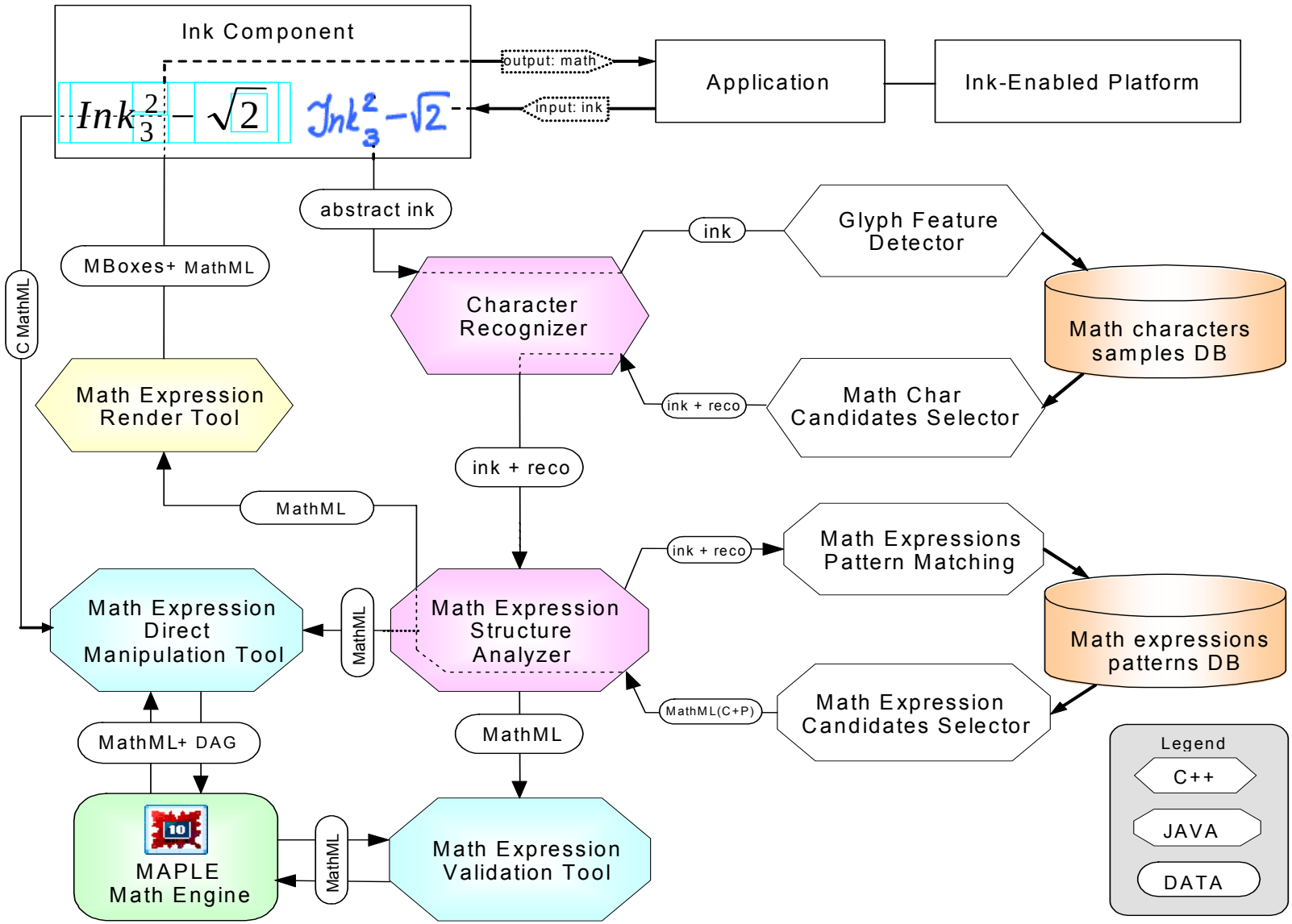
- Spreadsheet-like recalculation

- Collaboration

Early Projects at ORCCA

- **CrossPad**
  w Louie (2000) Off-line analysis
- **Pocket PC**
  w Wan (2001) Elastic matching, alternative prompting
- **Single-line expression grouping**
  w So (2003)
- **Notation Selection Tool**
  w Liu, Smirnova (2000-2003)
- **Expression Transformation**
  w Huerter Li Rodionov Smirnova So (1999-2004)
  TeX ↔ MathML ↔ OpenMath ↔ Maple

Components and relations



Character Recognition

- Large vocabulary of mathematical symbols
- Usual trade off is #symbols vs accuracy

  α vs α vs proportionality …

- ~ 2000 named entities in MathML
- Trick of using special alphabet doesn't work

- Stronger feature identification (w X. Xie)
- Heavier use of context (w So)

Ambiguities

Ambiguities

$$\sum_{i} i^2$$

Ambiguities

Ambiguities

$$\dot{z} + z = \sin \omega t$$

Data Collection

- Math survey
  - IBM Cross Pad Data
  - Tablet PC Data
- UniPen Data
- 240 symbols and a number of formulas.
- TeX sources of 20,000 papers from ArXiv

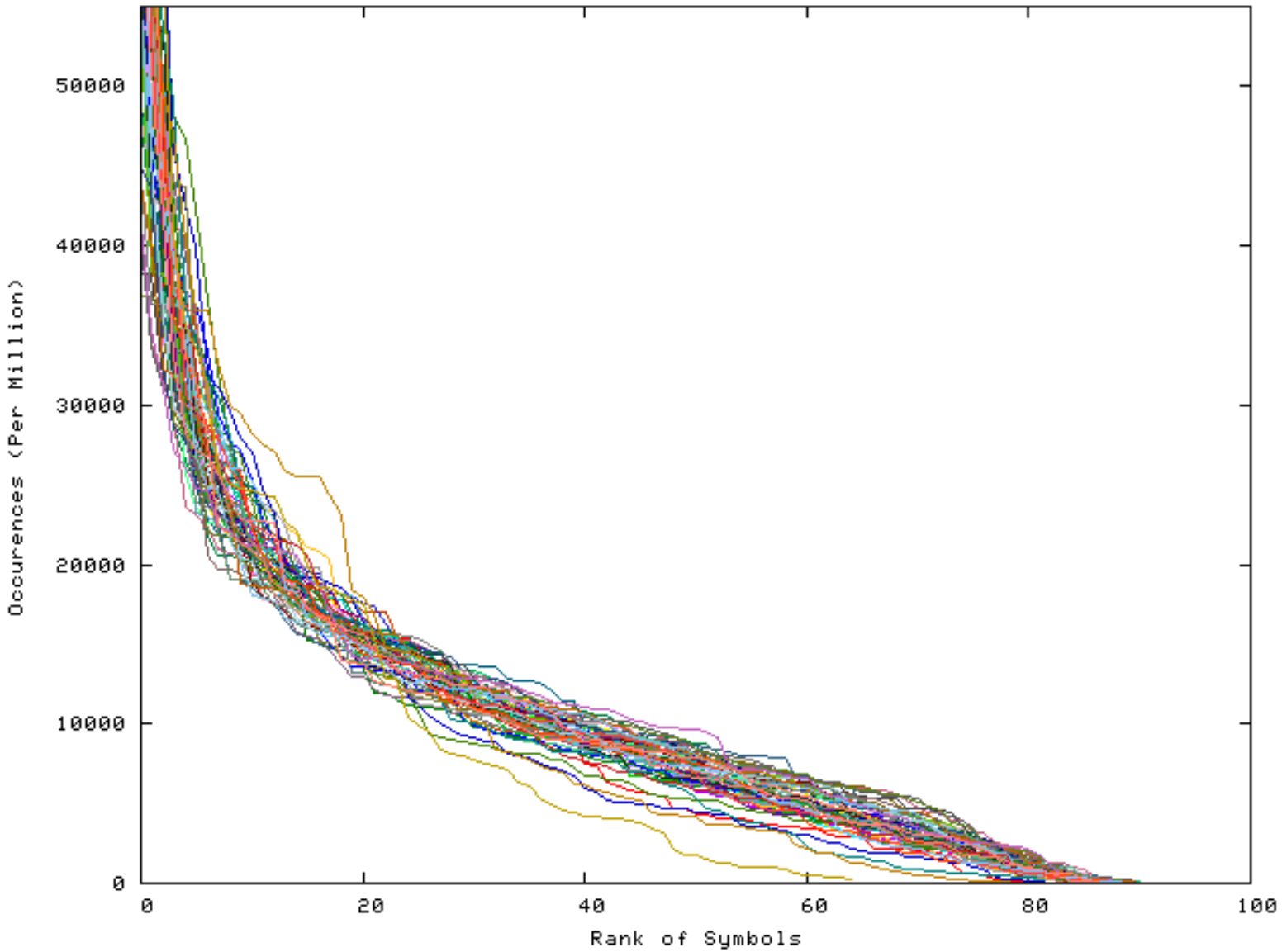
Id Frequencies in 3 Classifications

03 – Logic		
Ucode	Id	Freq
0069	i	51,565
006E	n	48,239
0078	x	41,042
0058	X	33,862
0041	A	29,845
0070	p	26,292
03B1	α	24,604
006B	k	24,374
0066	f	22,671
0061	a	22,030
0047	G	21,983
006D	m	19,893
006A	j	18,062
03C9	ω	18,015
004D	M	17,256
0053	S	17,122
0043	C	17,107
0046	F	16,773
0079	y	16,764
0074	t	15,693

11 – Num. Th.		
Ucode	Id	Freq
006E	n	58,186
0070	p	40,302
006B	k	38,230
0078	x	35,294
0069	i	35,100
0061	a	25,301
006D	m	23,642
0064	d	22,302
0071	q	21,797
0073	s	21,319
006A	j	21,153
0072	r	19,695
0074	t	19,654
0047	G	19,620
0058	X	19,535
0041	A	19,107
004B	K	18,905
0066	f	18,126
0046	F	16,524
004C	L	15,921

35 – PDE		
Ucode	Id	Freq
0078	x	51,773
0074	t	49,859
0075	u	39,841
006E	n	35,705
006B	k	29,924
0069	i	28,941
0073	s	25,234
006A	j	24,968
0064	d	24,095
004C	L	21,094
03B5	ϵ	20,740
03BB	λ	20,189
0070	p	19,107
0043	C	17,450
03B1	α	17,087
0072	r	16,834
0076	v	16,820
0061	a	15,931
0079	y	15,920
0066	f	15,215

Id Freq from All Classifications



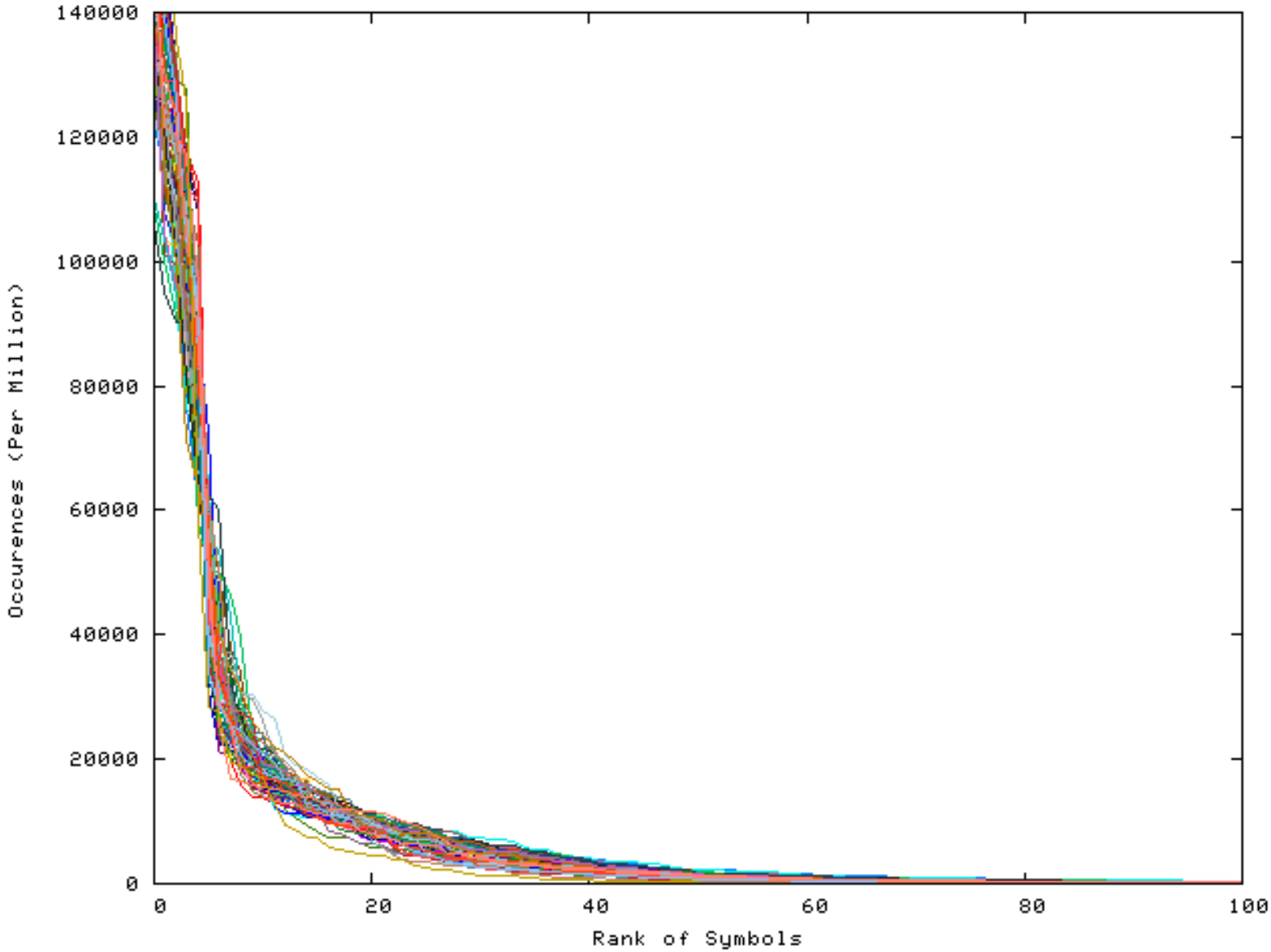
Op Frequencies in 3 Classifications

03 – Logic		
Ucode	Op	Freq
003D	$=$	121,806
2061		115,262
002C	,	100,880
2208	$\ni$	77,021
002D	$-$	60,732
002B	$+$	60,121
002A	$*$	32,796
003C	$<$	28,345
02C9	$\bar{}$	25,805
2192	$\rightarrow$	24,370
2264	$\leq$	24,242
002F	$/$	14,626
2026	$\ldots$	13,495
222A	$\cup$	12,654
2229	$\cap$	12,483
2286	$\subseteq$	12,330
003E	$>$	11,784
2223	$\mid$	9,883
22EF	$\cdots$	9,781
02DC	$\sim$	9,428

11 – Num. Th.		
Ucode	Op	Freq
003D	$=$	130,735
002D	$-$	128,330
2061		112,484
002C	,	104,964
002B	$+$	94,172
002F	$/$	40,239
2208	$\ni$	39,319
2211	$\sum$	20,165
2264	$\leq$	19,574
2192	$\rightarrow$	18,481
002A	$*$	17,757
00AF	$\bar{}$	14,708
221E	∞	14,627
003E	$>$	12,926
22EF	$\cdots$	12,358
02DC	$\sim$	12,209
2265	$\geq$	11,963
2113	ℓ	10,997
003C	$<$	10,151
00D7	$\times$	10,144

35 – PDE		
Ucode	Op	Freq
002D	$-$	138,603
002C	,	111,176
2061		103,527
003D	$=$	103,376
002B	$+$	97,579
2208	$\ni$	38,370
2264	$\leq$	34,575
2202	∂	28,815
002F	$/$	25,985
221E	∞	23,460
222B	$\int$	23,196
02DC	$\sim$	19,545
003C	$<$	16,453
2207	∇	15,387
003E	$>$	15,256
002A	$*$	14,470
2192	$\rightarrow$	14,381
22C5	$\cdot$	12,669
2211	$\sum$	12,394
2265	$\geq$	11,531

Op Freq from All Classifications



Most Popular Expressions of Size 7

03 – Logic (Sz: 7)	
#	Expr
86	ϕ_{m+4i-4}
69	$\nu_0, \ldots, \nu_k$
62	ϕ_{m+4i-2}
32	$\widetilde{y}_{i-1}^{-1}$
29	$(r_\nu \colon \nu \in pos(t))$
28	ϕ_{m+4i-1}
28	$(17\, Gen\, r)$
24	$(b_j \mapsto f_{ij})_j$
24	ϕ_{m+4i-3}
23	$h + d_1 + d_2$

11 – Num. Th. (Sz: 7)	
#	Expr
107	$\sum_{k=1}^{n-1}$
97	$\sum_{k=0}^{n-1}$
76	$\sum_{i=0}^{n-1}$
71	$n + m - i - j$
69	$T', \lambda'_{T'}$
68	$\widetilde{G}_{k,n,d}$
66	$B^{\dagger,s}_{rig,K}$
64	$\sum_{j=0}^{n-1}$
61	$\prod_{j,k=1}^{n}$
59	$\left(\frac{n+m}{n}\right)^{-1}$

35 – PDE (Sz: 7)	
#	Expr
445	$\frac{n+2}{n-2}$
194	$\frac{n+4}{n-4}$
110	(x', ξ', μ)
96	$p-1, q-1$
90	$-(a+1)p + c$
88	$\sum_{i,j=1}^{n}$
75	$j_1, j_2 \geq 0$
75	$(g(t), K(t))$
70	$u^{\frac{2n}{n-2}}$
69	$(t, x; \tau, \xi)$

Expression Analysis and Transformation

- Understanding expression arrangement and re-arrangement

$ax + by +$
$cz + wt$

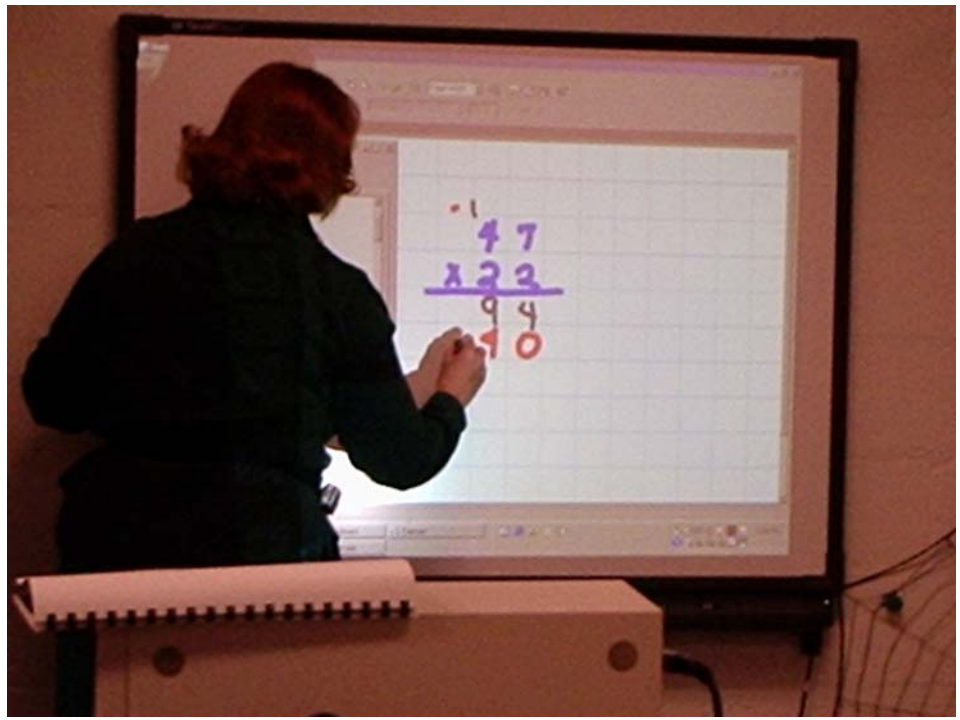
$$\begin{bmatrix} ax & by \\ cz & wt \end{bmatrix}$$

$$a\,x = b\,y$$
$$= c\,z - wt$$

$u_2v_1(a+b+c+z)$ $u_2F_1(a, b, c; z)$

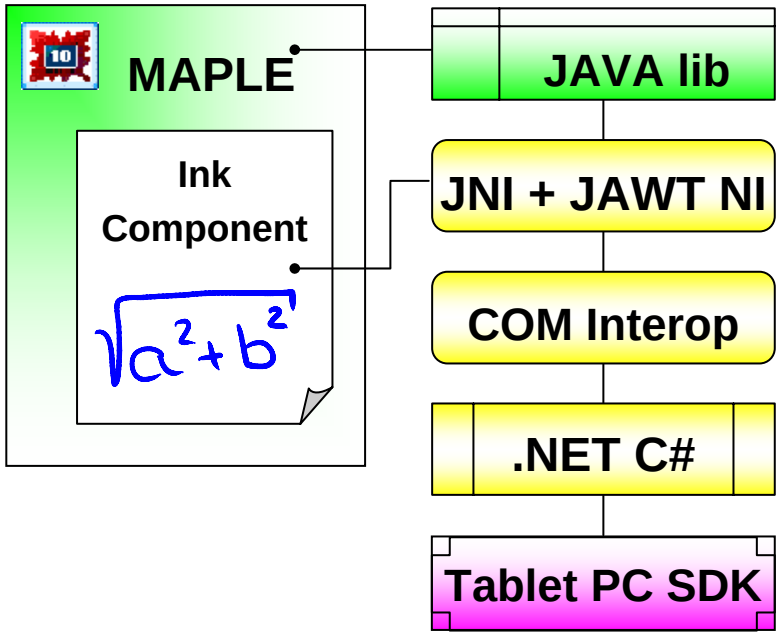
$(x+y)^2$

Portability

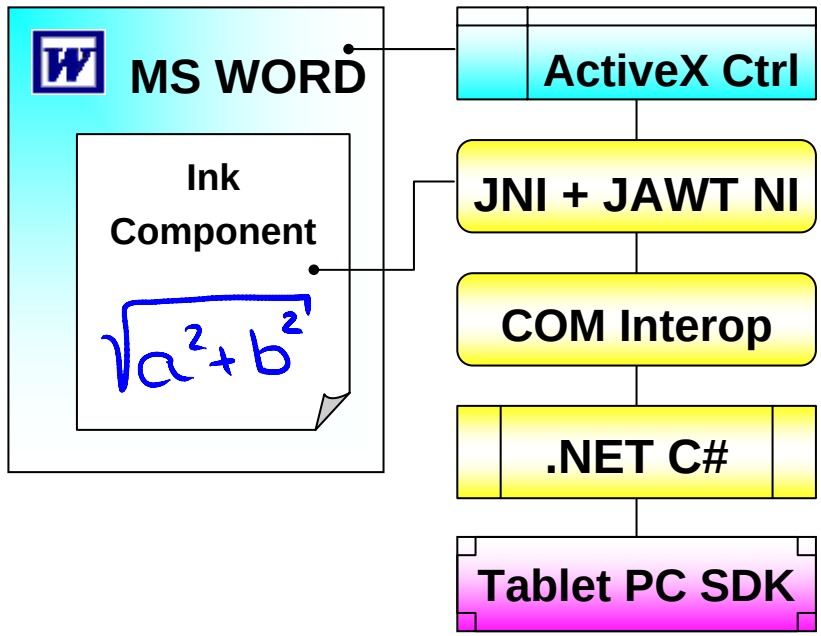


- Platform Portability
  - *Across platforms* and applications
  - *Over time* for evolving platforms and applications
- Digital Ink Portability
  - Can be achieved with InkML
  - Wrappers for device-specific ink interfaces
- Mathematical Data Portability
  - OpenMath
  - MathML

Interface to Host Application

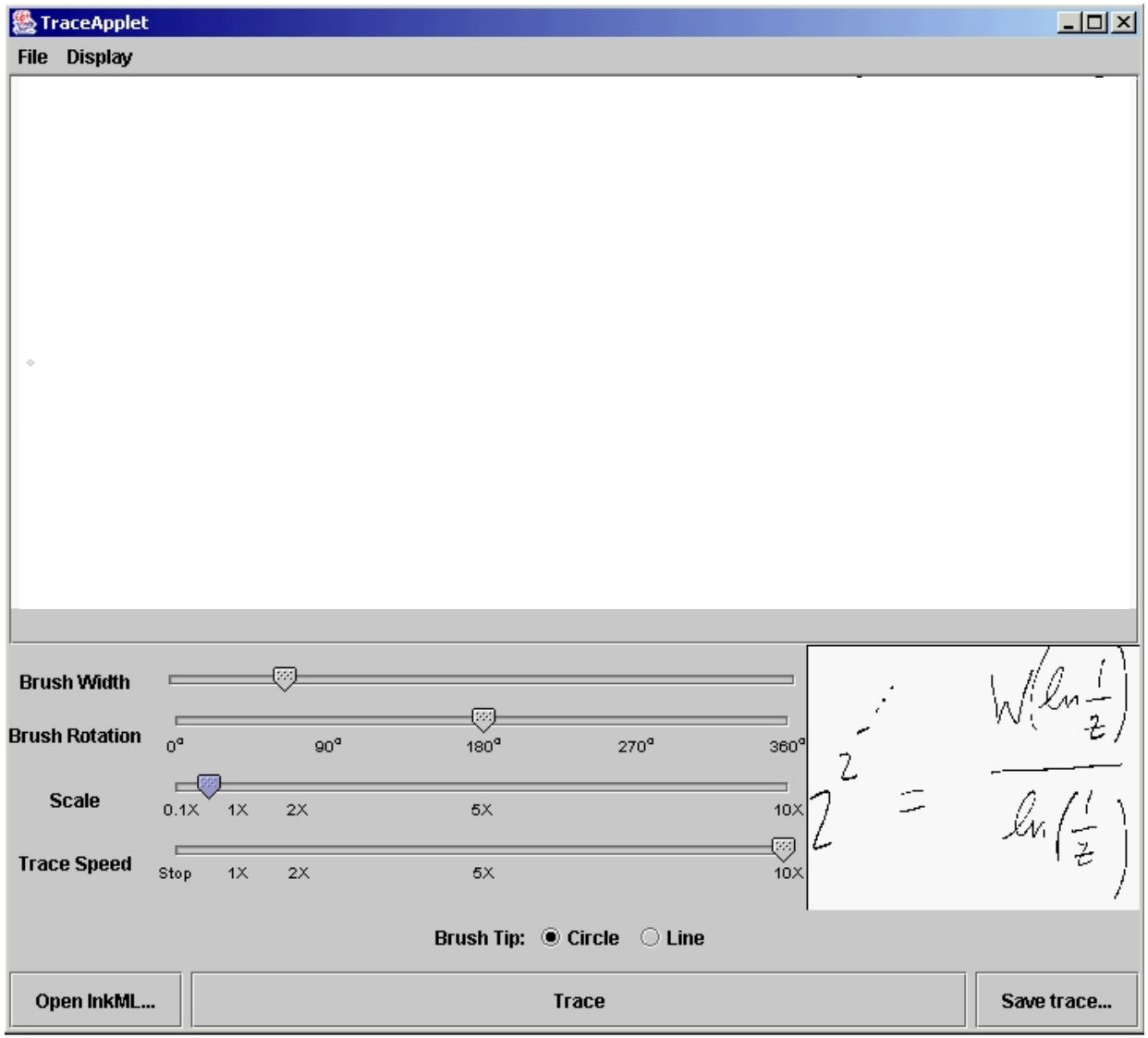


- Java library
- accessing .NET control
- through JNI

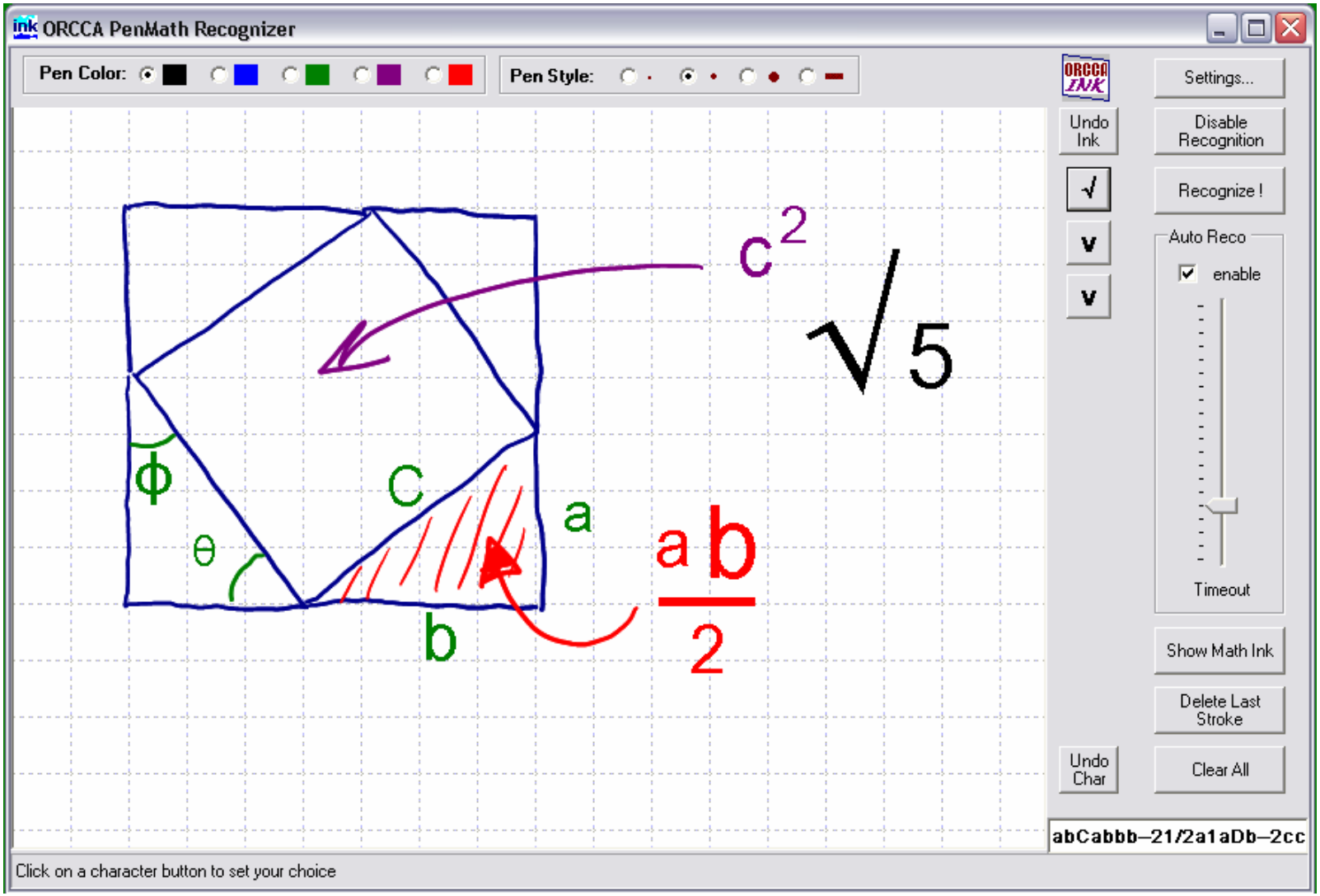


- ActiveX control
- accessing .NET control
- via Win32 C++ Wrapper

InkML Viewer



Prototype


ORCCA PenMath Recognizer
Pen Color:
Pen Style:
Settings...
Undo Ink
Disable Recognition
Recognize !
Auto Reco
enable
Timeout
Show Math Ink
Delete Last Stroke
Undo Char
Clear All
abCabbb−21/2a1aDb−2cc
Click on a character button to set your choice
c²
√5
Φ
θ
c
a
b
a b / 2

Example:

Suppose again that we have two tasks, T_1 and T_2, with

$$p_i(t) = a_i \lambda_i e^{-\lambda_i t} \qquad q_i(t) = (1 - a_i) \lambda_i e^{-\lambda_i t}$$

for $0 \le a_i \le 1, \ \lambda_i > 0.$

For both tasks let the time allotment function be $v_i(t) = \frac{t}{2}$

Then we have

$$P_i(t) = a_i \left(1 - e^{\left(-\frac{1}{2}\lambda_i t\right)}\right) \tag{1}$$

$$Q_i(t) = (1 - a_i) \left(1 - e^{\left(-\frac{1}{2}\lambda_i t\right)}\right) \tag{2}$$

Which implies

$$P_A(t) + Q_A(t) = 1 - (1 - a_i) \left(1 - e^{\left(-\frac{1}{2}\lambda_2 t\right)}\right) - (1 - a_2) e^{\left(-\frac{1}{2}\lambda_1 t\right)} \tag{3}$$

$$+ (1 - a_{1 - a_2}) e^{\left(t \cdot \left(\frac{1}{2}\lambda_1 + \frac{1}{2}\lambda_2\right)\right)}$$

Example:

Suppose again that we have two tasks, $\mathbf{T}_1$ and $\mathbf{T}_2$, with

$$p_i(t) = a_i\ \lambda_i\ e^{-\lambda_i t} \qquad q_i(t) = (1 - a_i)\lambda_i\ e^{-\lambda_i t}$$

for $0 \le a_i \le 1$, $\lambda_i > 0$. For both tasks let the time allotment function be $v_i(t) = t/2$. Then we have

$$P_i(t) = a_i(1 - e^{-\lambda_i t/2}) \qquad Q_i(t) = (1 - a_i)(1 - e^{-\lambda_i t/2})$$

which implies

$$P_*(t) + Q_*(t) = 1 - (1 - a_1)e^{-\lambda_2 t/2} - (1 - a_2)e^{-\lambda_1 t/2} + (1 - a_1 - a_2)e^{-t(\lambda_1 + \lambda_2)/2}$$

Using (3.17), the above expressino yields

$$E_{par}(t) = 2 \cdot \left[\frac{1 - a_2}{\lambda_1} + \frac{1 - a_1}{\lambda_2} + \frac{1 - (a_1 + a_2)}{\lambda_1 + \lambda_1} \right]$$

Requirements for Improvised Synchronous Collaboration

Werner Krandick

Department of Computer Science
Drexel University, Philadelphia, USA

Why improvised synchronous collaboration?

Use improvised synchronous collaboration to

- capture creativity,
- make decisions on the spot,
- access information on the go.

Why smartphones?

Smartphones are the medium of choice for improvised synchronous collaboration.

Ubiquity: A growing percentage of users carry their phone at all times. A growing percentage of phones are smartphones.

Functionality: PDA, Internet, phone, camera, location sensor,…

Limitations: small size, low power --- affects input, output, and processing power.

Why involve computer algebra researchers?

For the benefit of smartphone research:

Mathematics can be used to explore---and push---the limits of representation; see the role of mathematics in typesetting (TeX, LaTeX).

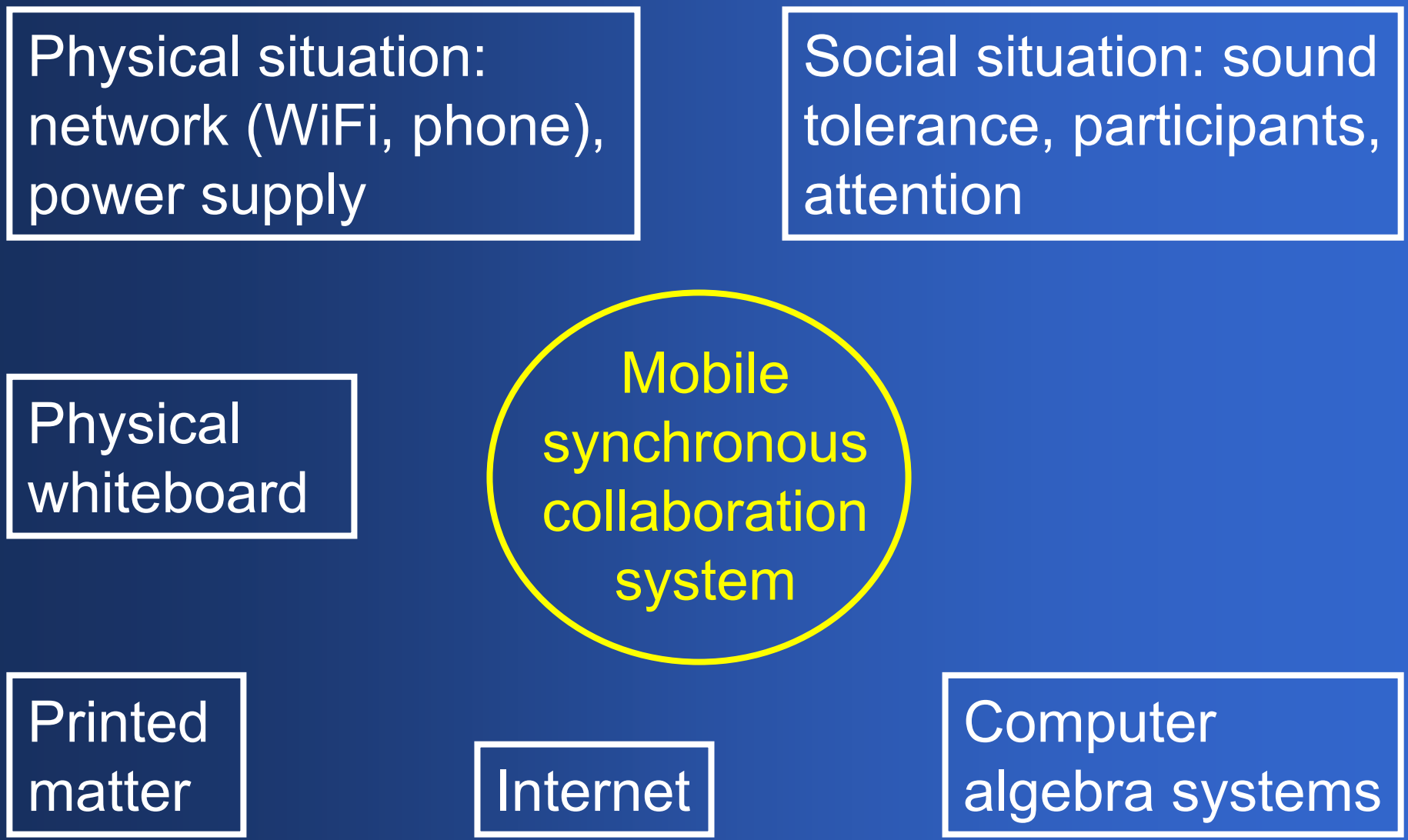
For the benefit of computer algebra:

Ubiquitous computer algebra systems, support for teamwork, create computer algebra services.

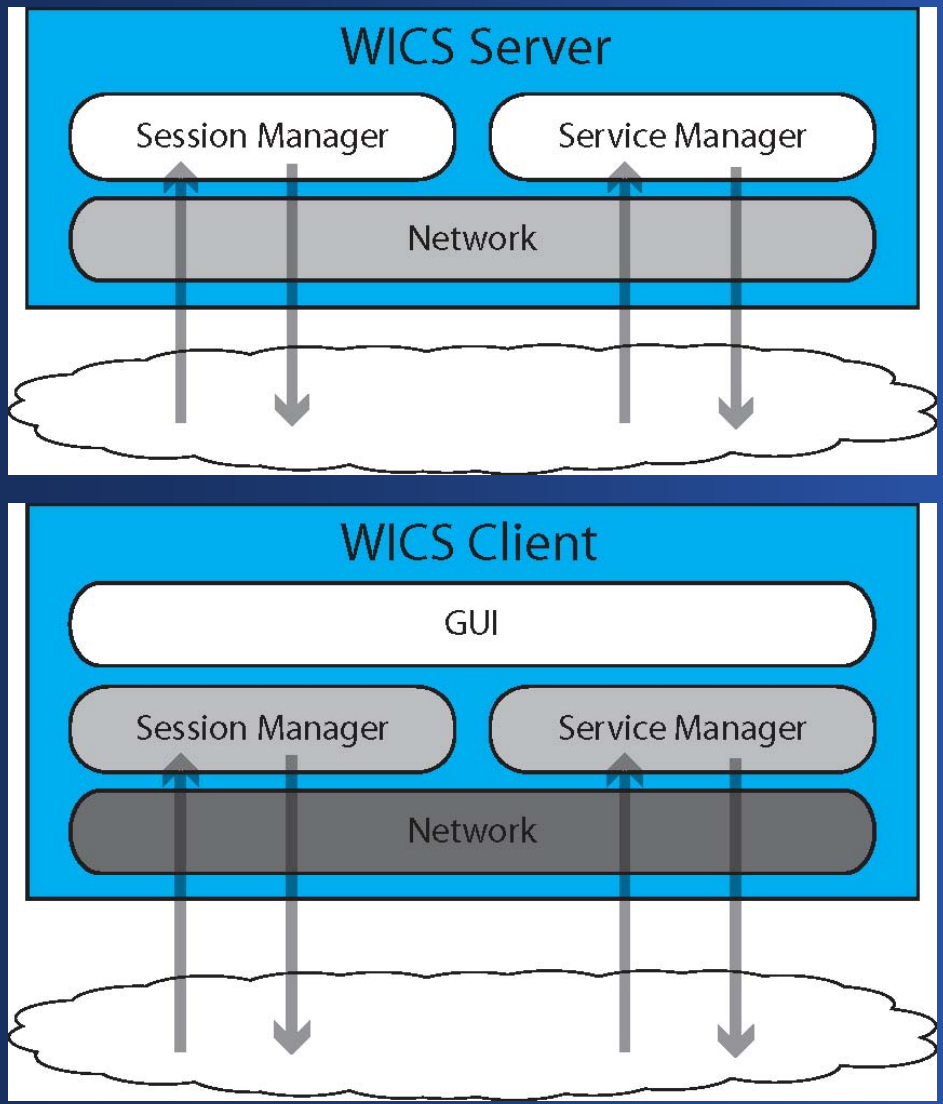
For the benefit of mathematics collaboration:

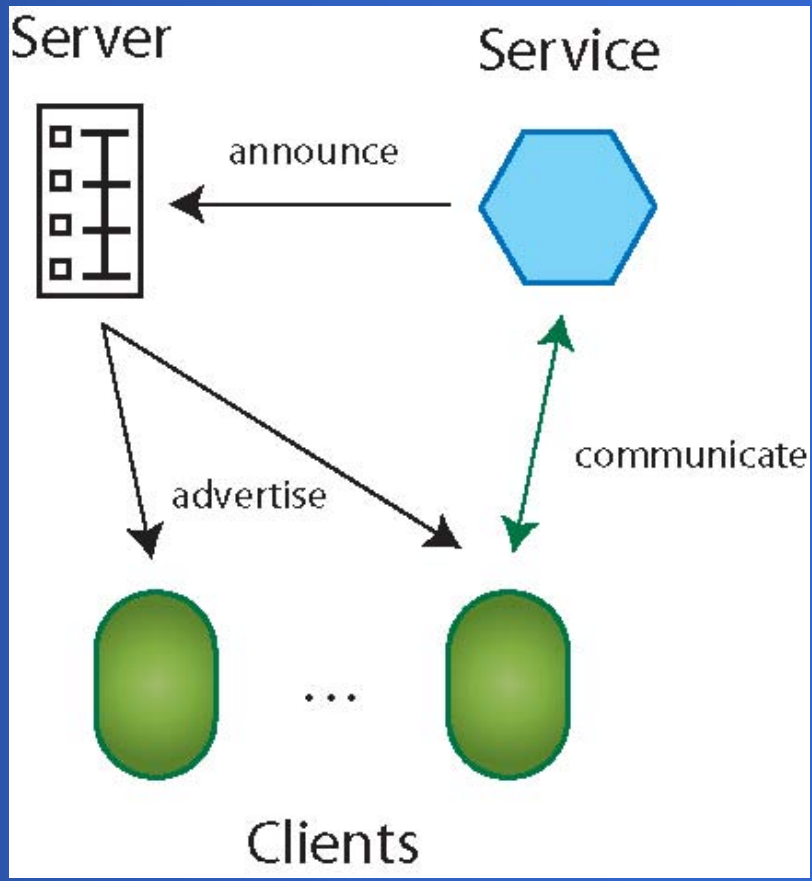
Computer algebra researchers are domain experts.

Context of Mobile Synchronous Collaboration



WICS architecture





Conclusion

Current applications do not exploit the potential of smartphones.

Mathematical collaboration can be used as a paradigm for smartphone-based collaboration.

Smartphone-based mathematics collaboration poses new challenges for computer algebra systems.

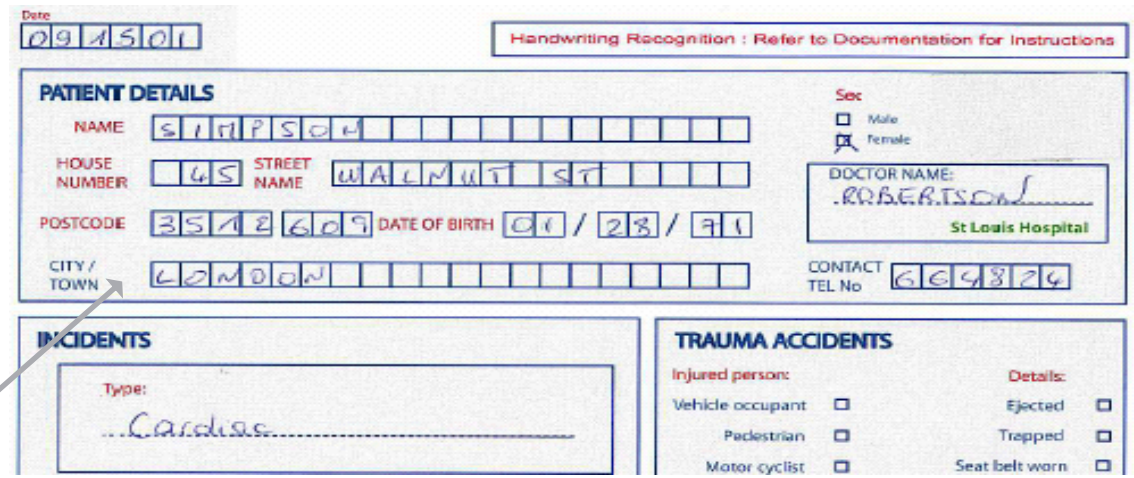
The InkML (language) applied to a multi-channel document processing system.

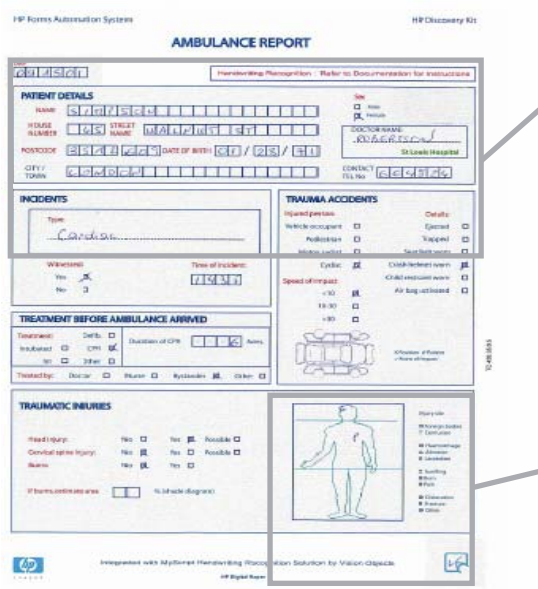
José Antonio Magaña Mesa (jomag@hp.com)
R+D Software Engineer
HP – Barcelona Division
November 2004

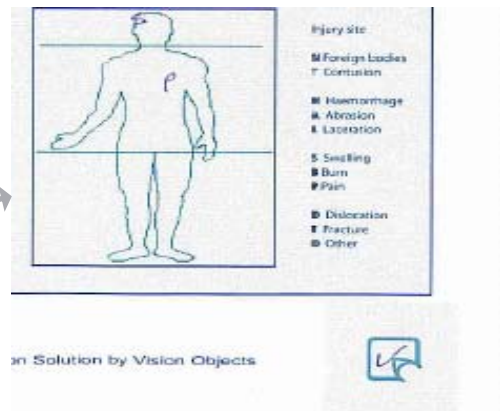
Application:
HP Forms Automation System
Workflow Connect 250

HP Forms Automation System

Fill out the printed form using the HP Digital Pen 250







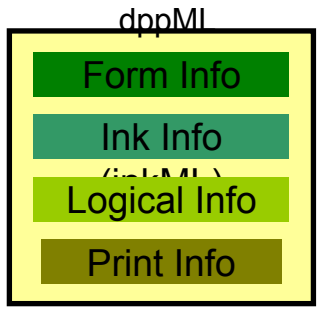
HP Forms Automation System Form Processing Workflow

1. User opens digital form using Acrobat Reader.
2. User issues print command, using Digital Paper Printer driver, from Acrobat Reader.
3. Digital Paper Printer driver sends requests, along with form id, to Service Controller for unique dot pattern. Service Controller responds with dot pattern information.
4. Digital Paper Printer sends print command to printer.
5. HP LaserJet prints paper form with unique dot pattern.
6. User writes on digital paper form and pen simultaneously captures written pen stroke data. User marks Send box when done filling form.
7. User docks digital pen to cradle to being data upload to appropriate form processing application server.
8. Digital Pen software contacts Service Controller for upload request. Service controller checks for appropriate application server and responds with contact information. (Secured comm.)
9. Digital Pen software contacts application server to begin upload. (Optional secured comm.)
10. Forms processing application performs business logic to process form data; e.g. interfaces with IT applications, etc.

dppML – XML language for HP FAS

dppML is the key point in XML inter-application connectivity for Forms Processing:

- Developed using inkML for the definition of digital ink
- Schema based: **http://h30195.www3.hp.com/schemas/dpp/dppML1.0.0/dppML.xsd**
- Several different layers defined so that applications only send/receive the minimum set of information needed: LAYOUT, INK, HWR, LOGICAL (Processing Results), PRINT,...



FAS - dppML structure flexibility

Contains information about:

- Document structure (forms)
- Information filled by the user (digital ink)
- Process results at different stages: ICR, data correction, data validation
- System information
- User information

This information is separated in different blocks so that only the required layers are transmitted.

FAS - dppML samples
Text Field information (partial)

```
- <multiBoxTextArea name="PersonalInformation.Name">
    <isPrefilled>false</isPrefilled>
  - <bounds>
      <x>130.0</x>
      <y>218.0</y>
      <width>529.0</width>
      <height>22.0</height>
    </bounds>
  - <textLayout>
      <isMultiline>false</isMultiline>
    - <color>
        <red>255</red>
        <green>0</green>
        <blue>0</blue>
      </color>
      <alignment>LEFT</alignment>
      <fontSize>12.0</fontSize>
    </textLayout>
  + <boxes>
  - <recognitionAttributes>
      <language>English</language>
      <languageCode>en</languageCode>
      <dataFormat>FIRST-NAME</dataFormat>
      <dataFormatVersion>1.0</dataFormatVersion>
      <country>United Kingdom</country>
      <countryCode>GB</countryCode>
      <isDFCustom>false</isDFCustom>
      <segmentationCriteria>SEGMENTATION_NON_OVERLAPPING_WIDENING</segmentationCriteria>
    </recognitionAttributes>
  - <recognitionResults recognitionEngine="Vision Objects">
    - <candidateList>
      - <textCandidate confidence="49.40033" text="ALICIA">
          <characterCandidate character="A" confidence="87.449646" pos="0" />
          <characterCandidate character="L" confidence="85.88104" pos="1" />
          <characterCandidate character="I" confidence="18.03894" pos="2" />
          <characterCandidate character="C" confidence="58.43048" pos="3" />
          <characterCandidate character="I" confidence="18.431091" pos="4" />
          <characterCandidate character="A" confidence="99.60785" pos="5" />
        </textCandidate>
      </candidateList>
    </recognitionResults>
  - <correctionAttributes>
      <rejectionThreshold>100.0</rejectionThreshold>
    </correctionAttributes>
</multiBoxTextArea>
```

Layout information

Recognition attributes

Recognition results

Correction attributes

Forms Automation System
Multi-channel environment

dppML

dppML

External Form Processing Apps

dppML

dppML

Forms Automation System

dppML

Form data repository

dppML

dppML

XML

WORLD WIDE WEB consortium

XML.ORG

Representation and Annotation of Online Handwriting Data

Sriganesh Madhvanath <SriG@hp.com>

- Issue
  - Annotated datasets of handwriting essential for design, training and evaluation of data-driven HWR algorithms
  - Lack of such datasets in “neglected” scripts a significant barrier to HWR R&D

- Proposed remedy
  - A standard representation and freely available tools to promote creation and sharing of datasets

Handwriting Datasets
Not Only for Recognition

- Recognition of gestures and other forms
  - Annotation and Shorthand
  - Editing gestures (online)
  - Artificial symbols such as Unistroke and Grafiti
- Analysis
  - Forensic applications
  - Writer identification
  - Profiling/Graphology
- Matching
  - Signature verification
    - Authentication
    - Detection of forgeries
  - Retrieval of handwritten documents
- Script identification

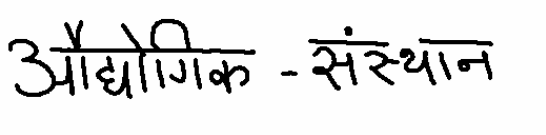
Annotation of Handwriting Data

- Meta-data – about the “writing act”
  - who wrote, where, how, with what device(s) ?
- Interpretation(s)
  - ground truth, at different levels of granularity
  - style
  - quality
  - script
  - other

Creation of Handwriting Datasets

- What data do we collect ?
- Who do we collect from ?
- Using what device(s) ?
- How do we annotate the data ?
- How do we represent ink and its annotation ?
- What tools do we use ?
- How do we validate the dataset ?

WORD



COMPOSITE CHARACTER (??)

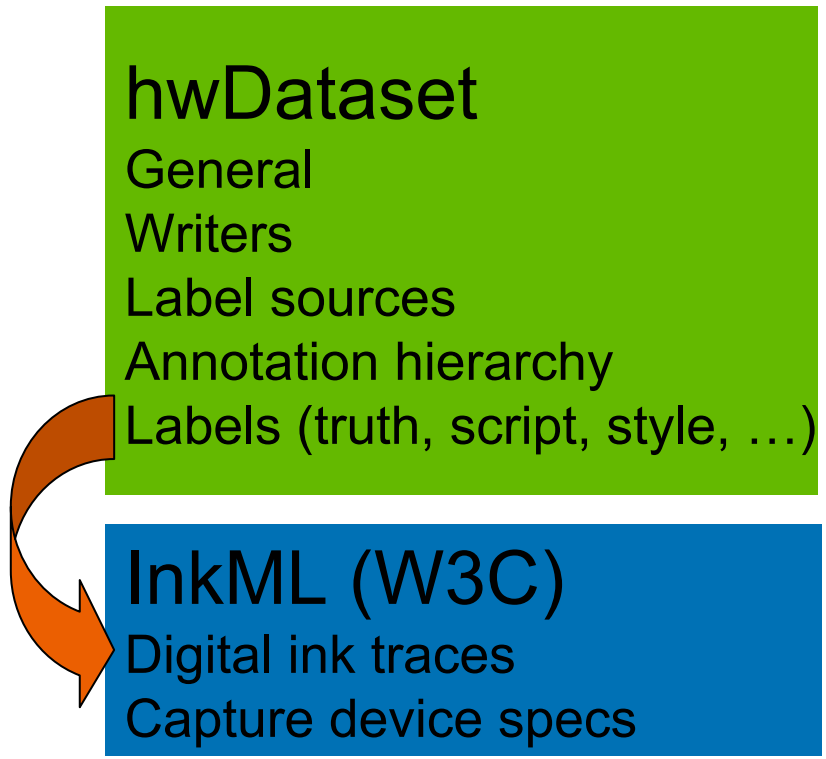
SYMBOL (?????)

Requirements of a Standard Representation

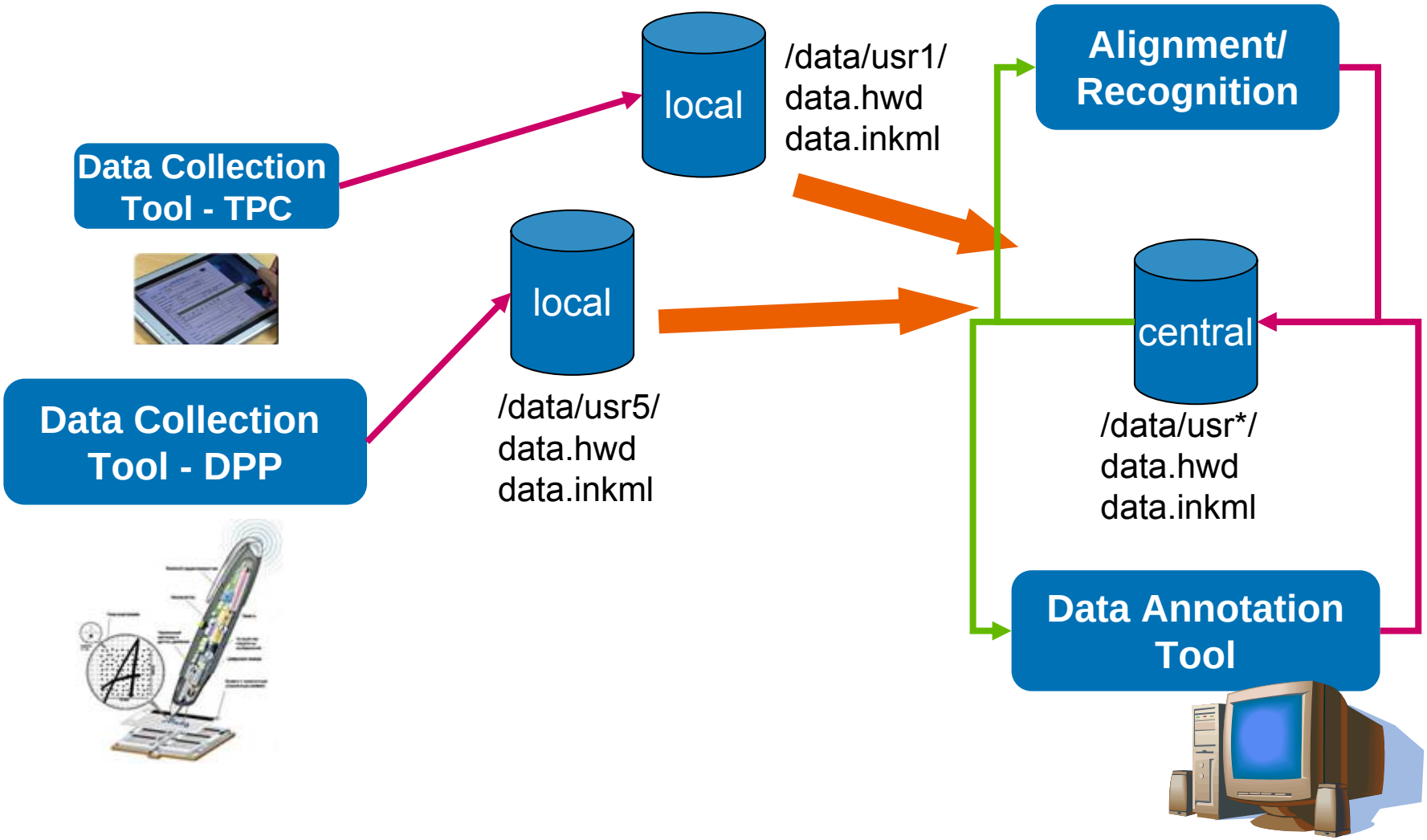
- Script-independence, multi-script documents
- User-definable annotation hierarchy (for syllabic scripts)
- Essential annotation: script, writing style, quality, ground truth
- Multiple writers and data capture environment
- Multiple sources of annotation, semi-automated generation of annotation
- Separation of handwriting (signal) from semantic interpretations
- Designed as well as casual data collection

hwDataset

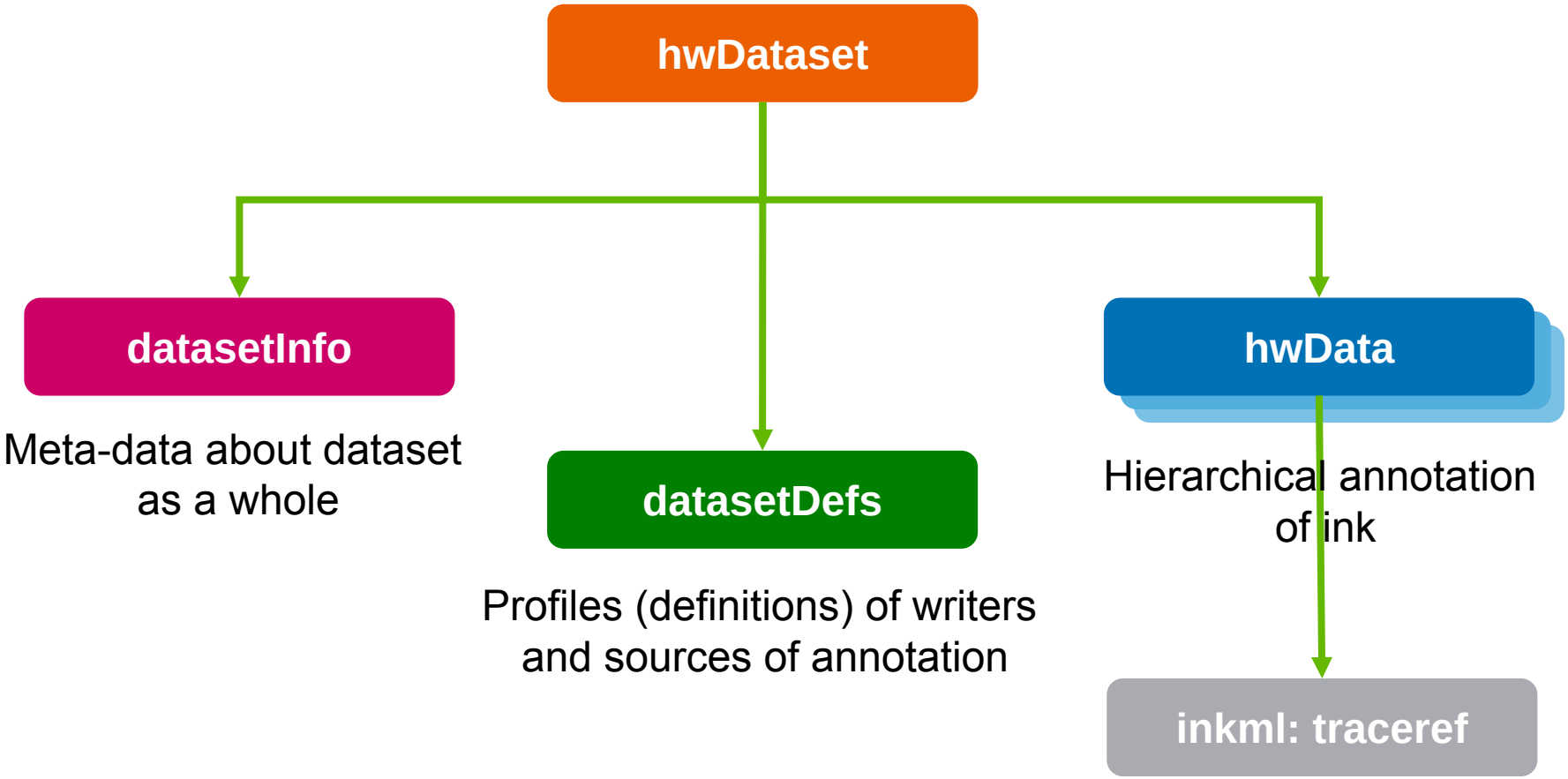
- XML representation of annotation of handwriting for handwriting corpora
  - platform-independent, hierarchical and extensible
- Targeted primarily at Online HWR
- Tags inspired largely by UNIPEN
- Refers to digital ink traces and device specs represented as InkML



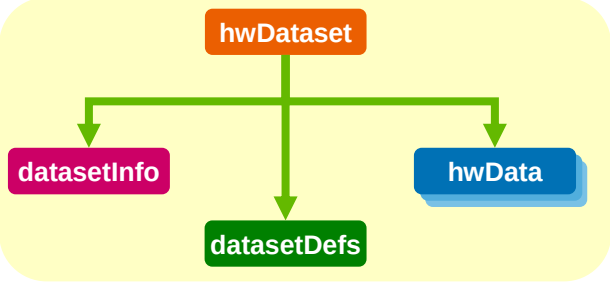
Creation of Handwriting Datasets



hwDataset


hwDataset
datasetInfo
Meta-data about dataset as a whole
datasetDefs
Profiles (definitions) of writers and sources of annotation
hwData
Hierarchical annotation of ink
inkml: traceref

hwData: Hierarchical Annotation



hwData
- id: "Trial01"
- desc: "Trial 1 of 5 trials"

H1
- writerRef: "WRITER01"

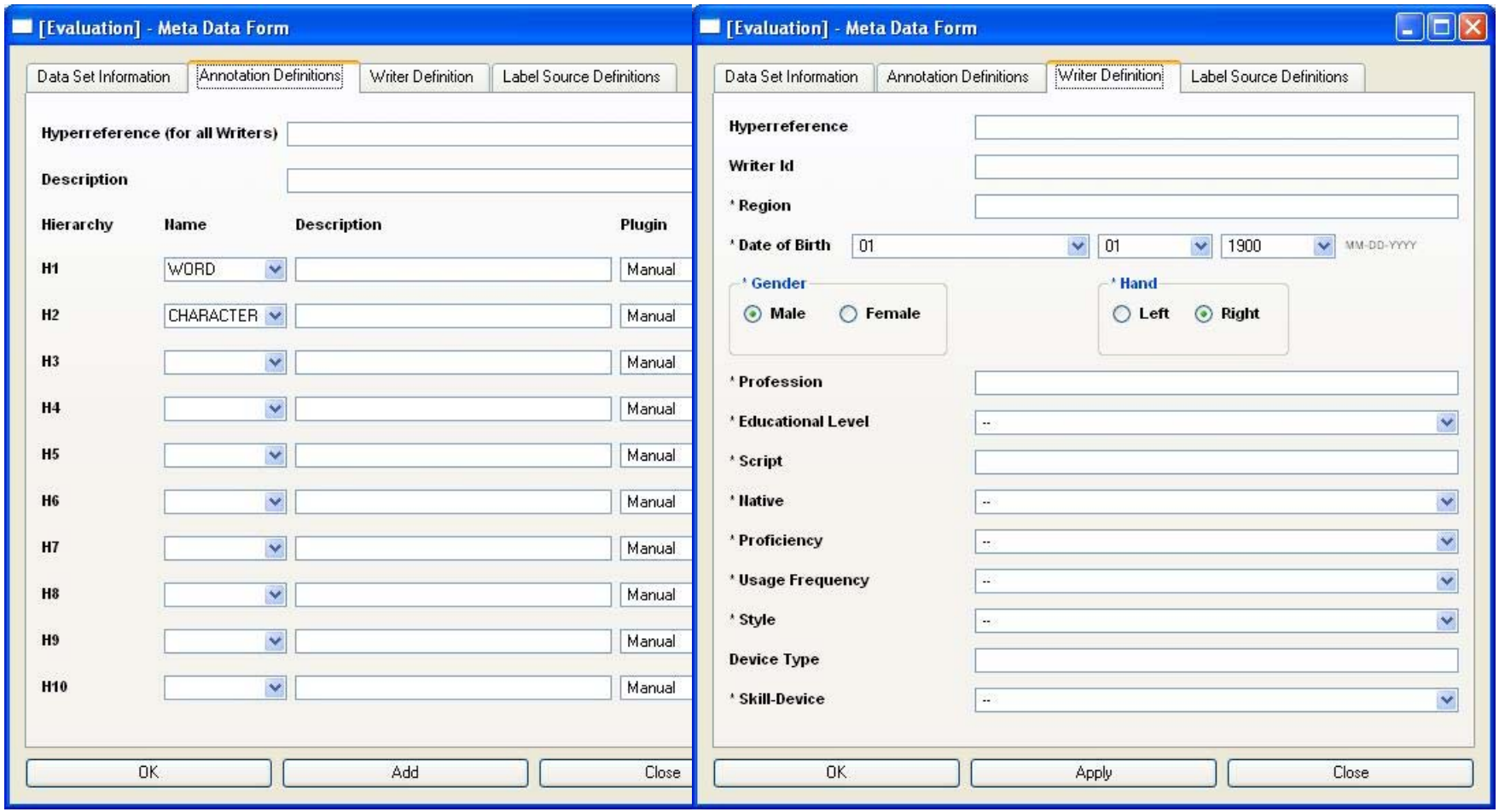
H2
- label
  - labelSrcRef: "HPLRECO"
  - category: "truth"
  - timeStamp: "1073026800000"
  - alternate: "Hello"
    - rank: "1"
    - score: "0.90"

hwTraces
- Inkml:traceRef

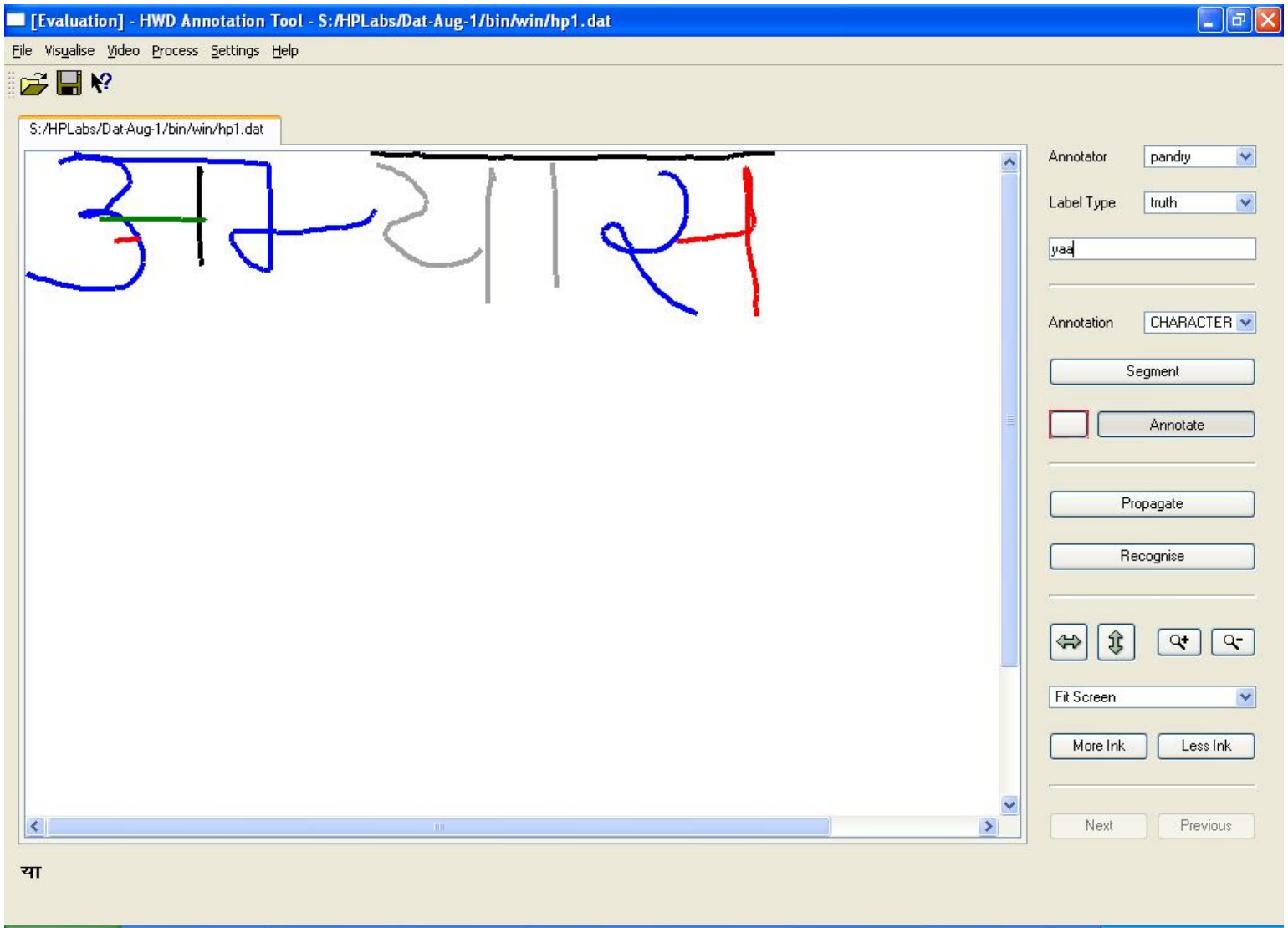
Handwriting Annotation Tool

- Input:
  - Raw traces (ASCII)
  - UNIPEN
  - InkML
  - hwDataset (+InkML)
- Output: hwDataset (+InkML)
- Functions
  - Viewing
  - Editing of all annotation (metadata, defs, hierarchy, labels)
  - Semi-automated processing (segmentation, alignment) using plug-ins
  - Propagation across trials
- Library of access functions

User Interface



User Interface



Status

- Representation
  - UPX (UNIPEN XML) is a new standard XML representation for handwriting datasets proposed by the International Unipen Foundation (IUF)
  - UPX will be based on hwDataset and InkML
  - Collaborating with IUF on converting UNIPEN 1.0 datasets to hwDataset, identifying and resolving issues
- Tools
  - Version 1.0 of annotation tool developed at IIIT-Hyd
  - Undergoing testing and final changes
  - Will be part of LipiTk v1.0

Summary

- Annotated datasets of handwriting a critical need for HWR development
- hwDataset is an XML representation for hierarchical annotation of handwriting data
- hwDataset in turn uses W3C InkML for representation of digital ink
- hwDataset is the basis for UPX, a new standard proposed by Intl Unipen Foundation
- The Handwriting Annotation Tool supports annotation of handwriting data captured as UNIPEN or InkML
- The tool will be part of LipiTk 1.0

- For more information, see:
  - http://www.hpl.hp.com/india/research/pen
  - Poster on UPX in this conference

Discussion of Community Needs

Conclusions and Summary

Next Steps

- For further information see `http://w3.org/2002/mmi/ink`
- Provide feedback
  - Talk to SriG and myself here at ICDAR.
  - Email **<www-multimodal@w3.org>**
- Expect (last?) working draft this fall.
- If you have trial applications let us know.