Digital SI

Machine-interpretable, unambiguous digital representation of metrological information and factual data
Analogue world

SI brochure, Section 5: Writing unit symbols and names, and expressing the values of quantities

• “As a consequence, there now exists a general consensus […]”

• “Compliance with these rules and style conventions […] supports the readability of scientific and technical papers.”
Analogue world

SI brochure, Section 5:
**Writing unit symbols and names, and expressing the values of quantities**

- **upright type** regardless of the type used in the surrounding text
- **printed** in lower-case letters unless they are derived from a proper name
- multiplication must be indicated by a space or a half-high (centred) dot (⋅)
- avoid ambiguities, for example by using brackets or negative exponents

“The use of the correct symbols for SI units, and for units in general, as listed in earlier chapters of this brochure, is **mandatory**. In this way ambiguities and misunderstandings in the values of quantities are avoided.”
M. P. Foster

From a computer science point of view, the SI system is not sufficiently systematic and consistent to be transferred directly to a form suitable for machines.

- Proposes fundamental changes in the SI

D. Pražák

Sees major weaknesses in Foster's proposals: partly contradictory in themselves, fundamental change of SI unrealistic, no future change and further development of SI taken into account.

Both are overlooking an important fact:
There is no need to change the human-readable part to achieve unambiguous machine-readability in digital data formats.
Digital world

Microsoft Word documents: XML-based

PDF documents: machine language

Websites: HTML-based

→ Human-readable digital content is generated by translating machine-readable information and data.
Machine-interpretable, unambiguous digital representation of metrological information and factual data
Digital, machine-readable units

- SI prefix (optional)
- SI unit
- Exponent (optional)

Decimal numbers only with dot as separator

Integer exponent in scientific format allowed

EMPIR
The EMPIR initiative is co-funded by the European Union’s Horizon 2020 research and innovation programme and the EMPIR Participating States
# Standardised digital data format

### Components (of the real quantity type)

<table>
<thead>
<tr>
<th>Type</th>
<th>Label</th>
<th>Value</th>
<th>Unit</th>
<th>DateTime</th>
<th>expandedUnc</th>
<th>coverageInterval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic real with expanded measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>uncertainty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic real with coverage interval</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(probabilistic-symmetric)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(S) sub type</th>
<th>mandatory</th>
<th>optional</th>
</tr>
</thead>
</table>

### Standards
- GUM
- SI unit format
- VIM
- ISO 80000
- CODATA
- IEEE 754
- RFC 362 (UTF-8)
- ISO 8601
Example use case

- Digital certificates (calibration, conformity assessment, etc.) based on XML format
- Standards based on XML format (ISO, IEC already started)
- Metadata in research data management
- Communication in the internet of things
- Scientific publications based on XML format