The Importance of Locale Information

M. Hosken NRSI, SIL International
Locale Information

Collation tailoring

Key Translated strings: Language, script, region names, days, months, etc.

Formatting: Time & date, currency, etc.

Exemplars: Main characters, Dictionary head characters, etc.

Resource locations: Fonts, keyboards, spelling, transcription, etc.
Who benefits directly?

Local Community
- Appropriate language support in tools

Computing Industry
- Feeds back into better language support

Language Documentation
Industry Driven

Minority language support

No financial incentive

Need to reduce costs per language

Industry wants data not code

Data doesn't have security implications, yet!

CLDR
LDML

XML language

Main Data

Display Names, Numbers, Calendars, Collation, Segmentations, Delimiters,

Measurements, Units, Characters, List Patterns

Supplemental Data - Territory Information

Language Use, Currency, Postal Codes, Calendar Preferences
SIL Namespace

Extended:

Collation, identity, exemplars

Resource references

Fonts, keyboards, spell checking, transforms

Quotation marks
Other Resources

Spell Checking

Hyphenation / Line breaking

Fonts

Keyboards

Transformations
Example

LibreOffice Language Extensions

Language tag

Hunspell spell checking

(Locale information in LDML)

Future Dreams

Hyphenation, line breaking

Predictive text keyboarding in Keyman
Language Tag

Unique writing system identifier

Keys Locale information

Keys writing system / language

Language Needs Assessment

DBL translations list
Language Tag

Ing-Scpt-RE-variants-x-extensions

Tag Wisely!

Don't include non-disambiguating information

Likely Subtags

alltags.txt
Where Does Data Go?

SLDR is public

CLDR is Unicode sub project

Google, IBM, Apple, Microsoft

Our tools

FW, PT, Bloom, etc.
Do I Have to Type All That?

Automatic data collection

FW, PT

Please say yes!

Need a language tag
Thank You