

INTERNATIONAL SAMPLE

Unicode

The following sections contain text from various languages using many different kinds of writing systems.

This sample contains unicode characters that may not be available in the font used to display the document. PDFreactor tries to find a replacement for these characters in defined fall-back fonts. If these fonts do not contain the character or the fall-back fonts are not available the characters cannot be displayed.

Arabic

نص حكيم له سر قاطع وذو شأن عظيم مكتوب على ثوب أخضر ومغلف بجلد أزرق.

A wise text which has an absolute secret and great importance, written on a green cloth and covered with blue leather.

Chinese

視野無限廣，窗外有藍天

The view is infinitely wide. There is blue sky outside the window.

Danish

Quizdeltagerne spiste jordbær med fløde, mens cirkusklovnen Walther spillede på xylofon.

The quiz contestants ate strawberry with cream while Walter the circus clown played the xylophone.

German

Victor jagt zwölf Boxkämpfer quer über den großen Sylter Deich.

Victor chases twelve boxers across the great dam of Sylt.

Greek

Θέλει αρετή και τόλμη η ελευθερία. (Ανδρέας Κάλβος)

Liberty requires virtue and mettle. (Andreas Kalvos)

Hebrew

דג סקרן שט לו בים זך אך לפתע פגש חבורה נחמדה שצצה כך.

A curious fish sailed a clear sea, and suddenly found nice company that just popped up.

Japanese

いろはにほへと ちりぬるを わかよたれそ つねならむ うゐのおくやま けふこえて あさきゆめ
みし 糸ひもせず

Even the blossoming flowers / Will eventually scatter / Who in this world / is unchanging? / The
deep mountains of vanity-- / We cross them today / And we shall not see superficial dreams / Nor
be deluded. (from Iroha-uta)

Polish

Pchnąć w tę łódź jeża lub ośm skrzyń fig.

Push into this boat a hedgehog or eight boxes of figs.

Russian

Съешь ещё этих мягких французских булок да выпей же чаю.

Eat some more of these soft French buns and drink some tea.

Demonstration of Bi-Directional text (BiDi) using Left-to-right and Right-to-left languages

Different languages can use different writing systems. In European cultures or cultures influenced by European ones, the direction in writing systems is usually from left to right. This is the case in Latin, German or English, for example. Other languages, especially semitic ones like Arabic, Persian or Hebrew, are written from right to left. Bidirectional script support offers the possibility to display text written in different directions.

Using the CSS property "direction", the base direction of a certain element can be set to "ltr" or "rtl", e.g.

```
.ltr { direction: ltr; }
```

In addition, using "unicode-bidi: bidi-override;" will override the base direction for a certain element:

```
.forcertl { direction: rtl; unicode-bidi: bidi-override; }
```

Example: Demonstration of alternating text direction by setting the direction for the second line to right-to-left.

This text is demonstrating the BiDi capabilities of PDFreactor,

.enil siht rof tfel-ot-thgir gnisu dna noitcerid esab eht gnidirrevo yb

The logical order (order in which the characters are stored in the document) of the second line of this example is:

"by overriding the base direction and using right-to-left for this line"

However, visually, the order is as follows:

"enil siht rof tfel-ot-thgir gnisu dna noitcerid esab eht gnidirrevo yb"

When text from different writing systems is used in one paragraph, the separate characters will keep their base direction if not otherwise defined. Characters originated from left-to-right or right-to-left languages respectively have the corresponding predefined order that can only be overridden by using the style "unicode-bidi: bidi-override;".

Right-to-left text inside left-to-right sections will be displayed as follows:

english CIBARA text

Upper-case letters represent Arabic text while lower-case letters represent English text. The logical order of this text would be:

"english ARABIC text"

The following samples demonstrate the difference between left-to-right and right-to-left languages. The base direction for the separate sections are defined corresponding to their language. English words and abbreviations in right-to-left sections are displayed correctly although the base direction is not overridden for them which proves the predefined direction for characters.

The text excerpts are taken from the original English and the translated Arabic and Hebrew versions of the W3C document "XML in 10 Points".

Left-to-right sample - English text excerpt:

5. XML is a family of technologies

XML 1.0 is the specification that defines what "tags" and "attributes" are. Beyond XML 1.0, "the XML family" is a growing set of modules that offer useful services to accomplish important and frequently demanded tasks. XLink describes a standard way to add hyperlinks to an XML file. XPointer is a syntax in development for pointing to parts of an XML document. An XPointer is a bit like a URL, but instead of pointing to documents on the Web, it points to pieces of data inside an XML file. CSS, the style sheet language, is applicable to XML as it is to HTML. XSL is the advanced language for expressing style sheets. It is based on XSLT, a transformation language used for rearranging, adding and deleting tags and attributes. The DOM is a standard set of function calls for manipulating XML (and HTML) files from a programming language. XML Schemas 1 and 2 help developers to precisely define the structures of their own XML-based formats. There are several more modules and tools available or under development. Keep an eye on W3C's technical reports page.

Right-to-left sample - Arabic text excerpt:

5. إكس إم إل مجموعة تكنولوجيات

XML 1.0 هي المواصفات التي تعرّف ماذا تكون "العلامات" و "الصّفات". حول تلك المواصفات هناك مجموعة متنامية من الوحدات التي تعرّض خدمات مفيدة لإنجاز مهام كثيرة الطلب وهامّة. Xlink يصف طريقة رسمية لإضافة و صلة رابطة لملفّ إكس إم إل. XPointer و XFragments قواعد تركيب نحوية للإشارة إلى أجزاء من وثيقة إكس إم إل. إشارة XPointer تتشابه مع إشارة URL ، لكنّ بدلاً من الإشارة إلى الوثائق على الويب ، تشير إلى قطع صغيرة من البيانات داخل ملفّ إكس إم إل. CSS، لغة النمط والأسلوب، تنطبق على إكس إم إل كما تنطبق على إتش تي إم إل. XSL هي اللّغة المتقدّمة للتعبير عن النمط والأسلوب . أساسها XSLT، لغة تحويل استخدمت لإعادة ترتيب أو لإضافة أو لحذف علامات او صفات. DOM هي مجموعة رسمية لاستدعاءات الوظائف من أجل استغلال ملفّات إكس إم إل (او إتش تي إم إل) و ذلك من داخل لغة برمجة. مخططات إكس إم إل 1 و 2 (XML Schemas 1 and 2) تساعد المطورين بأن يعرفوا بدقّة هياكل أشكالهم الشخصية المؤسسة على لغة إكس إم إل . هناك عدة وحدات و أدوات متاحة أو تحت التطوير. أنظر إلى صفحة تقارير W3C التقنية.

Right-to-left sample - Hebrew text excerpt:

5. XML הוא משפחה של טכנולוגיות

XML 1.0 הנו מפרט שמגדיר מהם "תגיות" ו"תכונות". מעבר לכך "משפחת XML" הינה מערך מודולים ההולך ומתרחב המציעים שירותים שמושיים כדי לבצע משימות חשובות ותכופות. Xlink מתאר כיצד להוסיף קישורים לקובצי XML. XPointer - ו-Xfragments הינם מודולים בפיתוח כדי לאפשר מצביעים בתוך מסמך XML. ה-XPointer הנו מעין URL, אך במקום להצביע על מסמכים ברשת, הוא מצביע על פיסת מידע בתוך הקובץ XML. CSS, שפה לעיצוב גליונות, זמינה ב-XML בדומה ל-XSL. HTML הינה שפה מתקדמת לעיצוב גליונות. השפה מבוססת על XSLT, שפה המשמשת לאירגון, הוספה ומחיקה של תגיות ותכונות. ה-DOM מערך פונקציות המשמשות למניפולציות בקבצי XML (ו-HTML) ע"י תיכנות. פורמט הגדרות מבנה XML Schema 1 ו-2. ישנם עוד מודולים וכלים מוכנים או בפיתוח. שים לב לדף הדיווחים הטכניים של W3C.

International List Numbering

PDFreactor supports several international list-style-types, including the following:

1. decimal	拾玖. japanese-formal
ii. lower-roman	二十. japanese-informal
III. upper-roman	೨೦. kannada
d. lower-alpha	二. katakana
E. upper-alpha	△. katakana-iroha
٦. arabic-indic	២៥. khmer
Է. upper-armenian	໒໕. lao
բ. lower-armenian	八. katakana
ঐ. bengali	൨൭. malayalam
១០. cambodian	ᠮᠯ. mongolian
११. devanagari	၂၉. myanmar
᠒᠔. georgian	୩୦. oriya
ν. lower-greek	۳۱. persian
Ξ. upper-greek	參拾貳. simp-chinese-formal
૧૫. gujarati	三十三. simp-chinese-informal
ੴ. gurmukhi	౩౪. telugu
ぢ. hiragana	๓๕. thai
そ. hiragana-iroha	མུ. tibetan
	۳۷. Urdu

For more information on supported list-style-types, please see PDFreactor manual.