

Chapter 1

Générer des documents PDF avec Pharo

Le format PDF d'Adobe est probablement l'un des formats de fichier les plus répandu. Utilisé quotidiennement, il rend quantité de services en permettant la production de documents graphiques, lisibles et échangeables entre pratiquement toutes les plate-formes matérielles ou logicielles existantes. Si vous recevez une facture, suite à un achat sur un site de e-commerce ou si vous téléchargez un rapport, un livre, une thèse ou encore un formulaire administratif, ces fichiers seront fort probablement des documents PDF consultables sur votre ordinateur, votre téléphone ou votre tablette. Sur l'ensemble de ces supports, ils conserveront une apparence identique. Pour les programmeurs, ce format est devenu incontournable et doit faire parti de la trousse à outils de base d'un outil de développement. Pour Pharo, c'est le framework Artefact qui prend en charge la conception et la génération de documents PDF.

1.1 Aux sources du projet

Artefact est un framework PDF original. Il base la définition d'un document PDF sur une description objet. Pour lui, l'intégralité des éléments composant un fichier PDF se résume à une arborescence d'objets. Le document est un objet contenant d'autres objets correspondant à des pages. Sur chacune des pages, les éléments imprimables ou non visibles sont également des objets. Ceux ci ont ainsi la possibilité d'être réutilisés dans un même document mais aussi dans plusieurs documents. Ces éléments peuvent être aussi simple qu'un texte ou une image mais également, être un composant complexe doté d'un comportement évolué et d'une apparence particulière. Artefact pro-

pose ainsi par défaut des éléments tels que des paragraphes ou encore des tableaux permettant de générer rapidement des rapports. La force de ces éléments est qu'ils sont indépendants les uns des autres. L'ordre dans lequel vous les positionnez dans le document n'influe pas sur leur apparence. Contrairement à de nombreux frameworks PDF qui exploitent une notion de flux dans la définition des styles (un texte bleu sera suivi d'un autre texte bleu en l'absence d'une directive de style différente), Artefact considère que chaque élément intègre une feuille de style qui lui est propre. Si un attribut n'est pas défini dans l'élément, Artefact exploite alors une feuille de style par défaut définie au niveau du document.

Cette autonomie des éléments et cette gestion des styles est une caractéristique forte d'Artefact. Elle permet de générer facilement un document et de le personnaliser rapidement pour une exploitation particulière. Efficacité, productivité et extensibilité ont été les mots clés qui ont guidé le design du produit.

1.2 Concept, caractéristiques et limitations

Les concepts d'Artefact sont maintenant stabilisés et il est déjà utilisé dans le monde professionnel par des entreprises utilisant Pharo. Pourtant, c'est encore un produit jeune et il est en plein développement. Pour cette raison, il est important d'énumérer ses fonctionnalités actuelles et de connaître ses limitations.

Artefact supporte la définition d'un document PDF, permet de définir des meta informations telles que le titre ou encore l'auteur, gère les options d'affichage lors de l'ouverture du document dans un lecteur compatible avec cette fonctionnalité et sait générer des documents PDF compressés.

Chaque page d'un document PDF peut se voir attribuer un format et une orientation particulière. Artefact propose par défaut un ensemble de formats de documents couvrant les besoins courants et il vous est possible de l'étendre pour des usages particuliers. Une page obtient sa position définitive dans un document qu'à partir du moment où elle a été ajoutée au document. Chaque page est autonome et cela permet donc de générer des documents à architecture variable.

Sur chaque page, Artefact dispose des éléments simples ou complexes. Un élément complexe est généralement défini à l'aide des éléments simples ou pourquoi pas, à partir d'autres éléments complexes. C'est le principe de composition qui est ici utilisé. Chaque élément est indépendant et se positionne par rapport au coin supérieur gauche d'une page. Artefact supporte la gestion des niveaux de gris ou l'utilisation de couleurs définies selon le modèle RVB (chaque composante de la couleur est représentée par un octet). Les fontes de caractères sont celles imposées par la spécification PDF mais

Artefact ne supporte pas les fontes true type (TTF). Il est possible d'insérer des images dans un document PDF mais seul le format JPEG est actuellement géré. Artefact ne gère pas la définition de champs interactifs de saisie, l'intégration de code Javascript ou encore les aspects sécurités de PDF tels que les certificats.

Bien évidemment, ces spécifications sont susceptibles d'évoluer au fur et à mesure des évolutions du framework. Artefact possède une architecture simple facilitant l'extensibilité et le support de nouvelles fonctionnalités. Restez connecté, le meilleur reste à venir.

1.3 Prise en main d'Artefact en 10 minutes

Vous programmez en Pharo et vous voulez générer des documents PDF dans votre application. Vous ne savez pas comment faire ? Ne soyez pas inquiet, dans quelques minutes, vous aurez produit votre premier document PDF.

Tout d'abord, il vous faut ajouter le framework Artefact à votre environnement de travail Pharo. La bonne nouvelle est qu'aucune librairie native n'est nécessaire pour faire fonctionner Artefact. Celui-ci est écrit entièrement en Pharo. Quelque soit votre plateforme d'exécution (Microsoft Windows, Mac OS X, Linux, Android, IOS, etc.), Artefact sera disponible et utilisable.

Installation du projet Artefact

Le projet Artefact est disponible sur le site web de dépôt SmalltalkHub. Pour installer Artefact, ouvrez un espace de travail (Workspace) et dedans, faites un copier/coller du script de chargement afin de l'exécuter. Ce script est fourni sur la page du projet Artefact et permet d'automatiser les étapes de l'installation. La configuration déclenche le téléchargement d'Artefact ainsi que celui du projet Unit destiné à la manipulation des différentes unités de mesure. Par défaut, la version récupérée sera une version stable (#stable) destinée à la mise en production.

Vous pouvez maintenant parcourir les principaux paquets et classes composant Artefact. Le paquet Artefact-Demos contient de nombreux exemples d'utilisation et permet une prise en main en douceur du framework. Le paquet Artefact contient les éléments principaux tels que les documents, les pages ou encore les feuilles de styles mais également une documentation électronique accessible via le Help Browser. Dans Artefact-Elements-Basic et Artefact-Elements-Composites, vous trouvez les différents composants (texte, formes géométriques, images, etc) proposés de base par le framework. Les fontes sont définies dans le paquet Artefact-Fonts et les formats de documents (A4, A3, ebook, etc.) dans le paquet Artefact-Formats.

"Hello World!"

Vous allez maintenant créer votre premier document PDF, le plus simple qui soit, un texte sur une page. Pour cela, vous devez définir une instance d'un document PDF contenant une page, sur laquelle vous positionnerez un composant texte.

Ouvrez un espace de travail et exécutez le code suivant pour créer votre document PDF.

```
PDFDocument new exportTo: 'helloworld.pdf' asFileReference writeStream
```

Une fois l'instance de PDFDocument créée, celle ci est exportée à l'aide d'un stream vers un fichier nommé helloworld.pdf. Par défaut, le document PDF produit est placé dans le répertoire de Pharo. L'ouverture de ce fichier à l'aide d'un lecteur PDF indique que le document est vide. C'est un résultat normal puisque vous n'avez pas encore défini et ajouté une page au contenu du document.

Enrichissons le code précédent en ajoutant un page au document.

```
PDFDocument new add: (PDFPage new); exportTo: 'helloworld.pdf' asFileReference  
writeStream
```

L'ouverture de ce fichier donne un résultat différent puisque votre document contient maintenant une page vierge. Artefact suit la philosophie "Convention over configuration". Seuls les paramètres spécifiques doivent être définis. Ainsi, par défaut, Artefact a attribué un format A4 à votre document, fixé l'orientation en portrait et initialisé les méta informations avec des valeurs par défaut.

Nous terminons notre premier document en ajoutant un composant texte sur la page.

```
PDFDocument new add:  
(PDFPage new add:  
(PDFTextElement new text: 'Hello World!';from:10mm@10mm)  
); exportTo: 'helloworld.pdf' asFileReference writeStream
```

Pour contenir le texte, nous avons besoin d'un composant de type PDF-TextElement que nous ajoutons sur la page. Nous définissons ensuite sa position à l'aide de la méthode #from:. L'utilisation du framework Unit autorise l'utilisation d'unités métriques tels que les millimètres (mm), les centimètres (cm) ou encore les pouces (inch). Ces coordonnées sont définies à partir de l'angle supérieur gauche de la page.

Artefact utilise un ensemble de valeurs par défaut afin de réduire au maximum le nombre de lignes de code nécessaires à la production d'un document. Le texte est donc écrit en noir, utilise une fonte (Helvetica) et une taille par défaut.

Ce premier exemple a permis d'introduire les concepts de bases d'Artefact et a démontré la facilité avec laquelle il vous est possible de générer un document PDF à partir de Pharo. Dans la suite de ce chapitre, nous allons approfondir vos connaissances sur le framework Artefact et vous permettre de créer des documents PDF complexes.

1.4 Définir un document

Artefact représente un document PDF à l'aide de l'objet `PDFDocument`. Cet objet est un conteneur destiné à recevoir des pages. C'est l'ordre dans lequel ces pages sont ajoutées qui définit l'organisation des données au sein de l'objet `PDFDocument` et non pas l'ordre dans lequel les pages sont créées. Ce mode de fonctionnement vous permet donc de générer des documents dont le contenu sera dans un premier temps dynamiquement produit puis organisé dans un second temps. L'objet `PDFDocument` dispose également d'options avancées portant sur le format du document, la gestion de la compression ou encore la définition de méta informations.

Ajouter des pages

Pour ajouter des pages dans un document, vous disposez d'une méthode nommée `#add:` qui ajoute la page à la suite de celles déjà présentes dans le document. Le fonctionnement est proche d'une pile FIFO (First In First Out). Lors de la génération du document, Artefact parcourt la liste des pages en partant de la plus anciennement ajoutée jusqu'à la dernière.

Spécifier les propriétés du documents

Un document PDF peut être configuré pour fixer le format utilisé et gérer la compression.

format et orientation du document

Par défaut, un document est généré en respectant le format A4 mais celui ci n'est pas le seul disponible dans Artefact. Dans le paquet `Artefact-Formats`, une liste de formats a été définie et couvre de nombreux besoins. Vous pouvez ainsi produire des documents en A3 (`PDFA3Format`), au format Letter (`PDFLetterFormat`) mais aussi dans un format adapté aux liseuses électroniques (`PDFEbookFormat`). Si vous avez besoin d'un format particulier, il vous suffit de le définir. Un format est simplement exprimé par la valeur de retour de la méthode `#defaultSize`.

L'objet `PDFDocument` possède une méthode `#format`: permettant de spécifier le format de toutes les pages du document. Pour chaque page, cette valeur sera prise par défaut si aucun format particulier n'est défini au niveau de la page. L'exemple suivant génère un document utilisant le format A3:

```
PDFDocument new
  format: (PDFA3Format new);
  add: (PDFPage new); exportTo: 'A3.pdf' asFileReference writeStream
```

Ces formats sont également utilisés pour spécifier l'orientation des pages en choisissant entre le mode portrait ou le mode paysage. Les deux méthodes sont `#setPortrait` et `#setLandscape`.

L'exemple suivant génère un document dans lequel les pages sont au format A3 et orientées en mode paysage.

```
PDFDocument new
  format: (PDFA3Format new setLandscape);
  add: (PDFPage new); exportTo: 'A3landscape.pdf' asFileReference writeStream
```

La sélection par défaut du mode paysage n'exclut pas la possibilité pour une page particulière d'être orientée en mode portrait. Artefact supporte parfaitement des pages de différents formats et de différentes orientations au sein d'un seul et même document.

Gestion de la compression

Le format PDF supporte la compression des données et c'est une bonne chose. Un document PDF peut contenir d'importants volumes de données dont la présentation graphique doit être décrite avec précision. Afin de réduire au maximum le poids des documents générés, Artefact applique par défaut un algorithme de compression. Si par curiosité, vous voulez désactiver cette option et analyser les données produites, vous disposez de la méthode `#uncompressed` qui désactive la compression.

L'exemple suivant génère un fichier PDF non compressé:

```
PDFDocument new
  uncompressed;
  add: (PDFPage new); exportTo: 'uncompressed.pdf' asFileReference writeStream
```

Une autre méthode, nommée `#compressed`, relance la compression si celle-ci a été préalablement désactivée.

Renseigner les méta informations

Chaque document PDF contient un ensemble de champs

1.5 Pages, formats et modèles

Gérer l'apparence avec les feuilles de styles

- Les composants simples

- Les composants complexes

- Créez vos composants