



Contenu [Afficher](#)

Objectif

L'objectif est de produire, à partir des articles d'un groupe Facebook, un magazine PDF destiné à être imprimé .

Un cas d'usage est celui d'une famille qui souhaite éditer un petit journal destiné à des personnes peu connectées sur internet et pour qui le support papier reste un média préférable. Il peut s'agir aussi d'un document rassemblant des reportages sportifs réalisés par les membres d'une association. Par extension, les principes exposés peuvent être utilisés par tout utilisateur Facebook pour la constitution d'un magazine imprimé à partir de ses publications, ou tout simplement pour exporter son contenu vers un système externe.

Approche

Le procédé qui a été choisi est le suivant:

Extraire les publications d'un groupe de façon programmatique en s'appuyant sur l'[API Graph de Facebook](#) afin de récupérer le texte, les images associées ainsi que l'auteur et la date.

Enregistrer les informations dans une database de façon à pouvoir effectuer un post-traitement (filtrage, corrections orthographiques, mise à l'écart de certaines publications selon des critères)

Retraiter éventuellement les images de type paysage afin d'obtenir un ratio différent de celui utilisé par Facebook qui est de 1,3 environ.



Produire un document au format XML ou JSON et leur appliquer un gabarit de mise en page à l'aide d'un moteur d'édition.

Principe et Architecture

- Utilisation de l'API Graph de Facebook (REST api) pour extraire en format JSON les éléments à imprimer (photo + texte)
- Réalisation d'une interface utilisateur avec Oracle APEX pour extraire les articles via l'API Graph, indiquer la période à sélectionner dans le groupe ou la page et produire le fichier XML intermédiaire.
- Traitement spécifique (*cropping*) des images avec ImageMagik afin d'obtenir un ratio largeur/hauteur adapté pour l'impression en A4
- Production du document PDF avec [Apache FOP](#)
- Impression et façonnage du magazine

Les appels REST vers l'api Graph sont lancées depuis du code PL/SQL. C'est donc une procédure stockée dans la database qui coordonne les appels nécessaires pour rassembler les articles contenant la description ainsi que le lien vers l'image associée.

Architecture simplifiée



Dans le schéma d'architecture ci-dessus, les services sont identifiés par la couleur bleue. On constate qu'il y en a trois. Dans mon cas, j'ai groupé les services de traitement d'image et Apache FOP sur une même instance cloud de type *compute* tandis que l'interface de pilotage (Application Oracle APEX) fonctionne sur une autre instance cloud. Il est bien sur possible de tout faire fonctionner sur un simple PC en local.



Réalisation

Créer un groupe Facebook

Pour mémoire: Créer un groupe Facebook ou en utiliser un déjà existant, indiquer une visibilité de type « privé » et enrôler les membres. Seul un administrateur de groupe peut inviter d'autres membres, sous réserve qu'il ait déjà une connexion individuelle avec ceux-ci. Il est cependant possible de constituer un groupe avec des membres qui ne possèdent pas de connexion entre eux.

Créer une application Facebook

Créer une nouvelle application Facebook à partir de la [console de développement](#). Dans le cadre du prototype, l'application a été gardée dans un mode « en cours de développement » et n'a pas été publiée.

Tester le bon fonctionnement des APIs en se connectant à la [console de développement API Explorer](#) (API Graph). Tout appel doit être accompagné d'un jeton d'accès (*Access Token*) obtenu après identification puis consentement quant aux priviléges accordés à l'application. cf [annexes](#)



Tant que l'application n'a pas fait l'objet d'une demande de certification auprès de Facebook, l'appel à certaines méthodes n'est pas possible. C'est le cas en particulier de la récupération du nom de l'auteur d'un article. Cela est évidemment très gênant pour la fabrication de notre magazine, puisque il s'agit d'une information éditoriale essentielle. Pour contourner cette limitation et puisque il n'était pas dans mon intention d'aller jusqu'au bout du processus de déclaration de l'application, j'ai demandé aux contributeurs du groupe d'ajouter systématiquement leur prénom sur la



dernière ligne de l'article. Les prénoms et les photos de profil sont gérés manuellement dans une table.

Créer une application de pilotage

Créer une application Oracle APEX à partir d'une instance APEX existante (depuis une instance *Oracle Free Tiers* par exemple, qui est gratuite). L'instance <https://apex.oracle.com> risque d'être inadaptée car il y a un quota d'appels de services web qui est assez bas. Chaque extraction génère un nombre d'appels représentant le double environ du nombre d'articles

Appliquer un traitement d'image

Cette contrainte est liée au choix de pouvoir disposer deux articles sur une page A4. Il s'agit d'un choix de mise en page purement arbitraire. (cf plus loin sur la partie [Design et mise en page](#))

Les images de type paysage doivent être retraitées afin d'obtenir un ratio 1,7 au lieu de 1,3. Pour cela on a recourt à un traitement de recadrage pour aplatiser l'image tout en gardant les proportions d'origine. Le choix technique est d'installer un service écrit en php avec le module [php_imagemagick](#).

Il aurait été possible d'utiliser les packages Oracle ORD_IMAGE, mais ceux-ci n'étaient pas disponibles dans l'instance *Oracle Free Tiers* que j'ai utilisée.

Fabriquer le document PDF

On utilise [Apache FOP](#). Il faut télécharger le logiciel sous Windows ou Linux. Il est possible de le faire fonctionner en mode commande, *standalone*.

Il est important de bien paramétrier FOP afin de disposer des polices de caractères



souhaitées. Pour cela, on paramètre le fichier de configuration avec la directive auto-detect (fichier de configuration fop.xconf) et on copie les polices de caractères sur le serveur Linux, si nécessaire.

```
<fonts>
...
<directory>C:\Windows\Fonts</directory>

<!-- auto-detect fonts -->
<auto-detect/>
</fonts>
```

Une tâche importante de réalisation a consisté à mettre au point le gabarit de mise en page automatique. Ce *template* se présente sous la forme d'un fichier contenant des déclarations XSLT de formatage. Les données extraites depuis Facebook sont présentées en format XML à Apache FOP qui les fusionne avec le template XSL.

Design et mise en page

Je me suis largement [inspiré du design](#) utilisé par le service payant [Famileo](#) et j'ai utilisé les mêmes principes de formatage.

Leur choix de mise en page consiste à assembler deux articles sur une même page A4. Si l'image est verticale, le texte est placé à droite, sinon au-dessous dans le cas d'une image horizontale. La quantité de texte est, d'après leur FAQ, arbitrairement limitée à 300 caractères.

Une des conséquences est que le rapport de forme (Largeur divisée par hauteur) pour les images horizontales est de l'ordre de 1,7 au minimum alors qu'il est de 1,3, voire



moins pour les images enregistrées dans Facebook. Cela explique que dans ce cas, il faut rogner les images en hauteur (cela fait l'objet du traitement de recadrage qui est effectué à la volée par le module PHP).

Pour le texte, j'ai fixé une limite plus haute que 300, quitte à réduire dans une certaine limite la taille des images selon l'article.



Echantillon de page avec deux articles

Façonnage du magazine imprimé

Il faut faire un choix sur le format de papier et cette décision entraîne un type de reliure correspondant. Le plus simple est de partir sur du format A4 « portrait » et d'imprimer en recto-verso. Les feuilles sont ensuite agrafées en cinq points sur le bord gauche et un ruban adhésif opaque (blanc ou fantaisie) et placé par dessus les agrafes.

L'autre solution consiste à partir sur du format A3 « Paysage » en mode recto/verso et d'imprimer sur deux colonnes. La reliure sera effectuée par deux points de couture au milieu des deux colonnes. L'avantage est de procurer de meilleures sensations à la manipulation puisque les pages A3 seront pliées et le magazine pourra se maintenir en position ouverte sur une table. Le problème à résoudre concerne le « chemin de fer » car les pages devront être assemblées dans un ordre qui n'est pas séquentiel:

Par exemple la première feuille physique comprend 4 pages:

- 2
- n-1
- n
- 1 (couverture)

Je n'ai pas encore étudié cette option A3.



Conclusions et enseignements

La réalisation automatisée d'un journal aboutit à un résultat qui peut être considéré comme correct mais qui n'est peut-être pas aussi attractif que lorsque la composition est réalisée manuellement.

L'avantage d'utiliser Facebook est de pouvoir constituer aisément un pool de participants, indépendamment de la tranche d'âge car FB est très populaire à la fois chez les jeunes et moins jeunes. Il n'est donc pas nécessaire de réaliser un réseau social privé avec les inconvénients de changement d'habitude, confidentialité, perte de mot de passe que cela aurait pu entraîner.

La destination des publications Facebook sur un support imprimé impose une discipline dans la rédaction des articles et notamment une exigence quant à la quantité minimum du texte associé à l'image. Or, sur FB, les usages donnent une plus grande importance à l'image qu'au texte. Cela implique donc une coordination éditoriale pour inciter les participants à contribuer différemment de leurs habitudes.

D'un point de vue budget, il faut :

- disposer d'un ordinateur sur lequel seront installés tous les softs répertoriés plus haut ou bien avoir une instance en cloud
- Coût d'impression des pages A4 en couleur recto-verso : environ 50 c /feuille, soit 4 € pour un magazine de 8 pages
- Coût d'envoi au destinataire: environ 2 €

Par comparaison, l'abonnement mensuel à Famileo et de 5,90€/mois par destinataire, ce qui en fait une solution compétitive s'il n'y a qu'un seul destinataire.



Annexes

Matériel, Code, templates

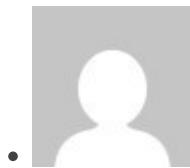
Sera mis à disposition plus tard sur Github.

Manipulations simples sur l'API Graph

- Pour obtenir la liste des groupes:
`/me/groups`
- Liste des posts
`/<group_id>/feed?since=<date_depart au format YYYY-MM-DD>`
- Récupérer les images
`/<feed_id>/attachments`

Remarque: si la publication Facebook est issue d'un partage depuis un espace personnel ou bien une page ou un autre groupe, il faut effectuer un appel supplémentaire car l'élément est considéré comme une story. cf *subattachments*

Author



[Patrick](#)

GPM Factory