

# RAPPORT DE STAGE SI2

## STAGE DEVELOPPEUR SI2 @ IPERNITY

**Sujet :** Indexation plein texte de nombreux formats de documents contenant du texte, et leurs métadonnées

 Tuteur : **T. M.**  
Réalisé du **28 juin** au **21 septembre** 2007



## REMERCIEMENTS

Je tiens à adresser mes remerciements à toute l'équipe d'ipernity pour m'avoir rapidement intégré au sein de l'entreprise. Grâce à eux, j'ai pu découvrir le monde des startups qui m'attire depuis longtemps, le fonctionnement au quotidien d'une entreprise dynamique qui se doit de toujours rapidement répondre aux attentes de ses clients et aussi, la difficulté et les risques que représente sa gestion.

Je souhaite remercier, tout particulièrement, T. M., mon maître de stage, pour l'attention qu'il a portée à répondre à toutes les questions techniques que j'ai pu lui poser et pour la qualité de ses réponses qui m'ont permis de beaucoup apprendre.

Enfin, je ne peux terminer sans faire un petit clin d'œil à Bully, le chien, si célèbre, d'ipernity qui sait apporter régulièrement joie et bonne humeur à toute l'équipe.



*Camille Roux*

## SOMMAIRE

Introduction .....	3
ipernity .....	4
Le travail proposé.....	11
Le travail réalisé .....	14
Conclusion .....	19
Bibliographie .....	20
Annexes.....	21
Table des matières .....	26

## INTRODUCTION

De plus en plus de gens publient leurs articles, photos, vidéos et autres documents sur le web. Ceci leur permet de les partager avec qui ils veulent, mais aussi d'y avoir accès de n'importe où, il suffit d'une simple connexion internet. Aujourd'hui, de nombreux sites Internet proposent d'héberger nos documents (souvent spécialisés pour un certain type : photo, vidéo, ...) et beaucoup sont très célèbres dans le monde du web :

[YouTube](#) [Twitter](#) [Zoomer](#) [DailyMotion](#) [Scribd](#) [Flickr](#)  
[Facebook](#) [Fotolog](#) [MySpace](#) [Google Video](#) [Skyblog](#) [Picasa](#)

Si le fait de mettre ses documents en ligne nous semble évident de nos jours, c'était loin d'être le cas, il y a quelques années, quand ipernity a commencé le projet.

Un des points qui permet à ces sites de se démarquer par rapport aux concurrents est la valeur ajoutée aux documents présentés : tag, géolocalisation, commentaires, affichage des métadonnées... ipernity propose déjà ce genre de fonctionnalités pour les photos et les vidéos, mais a souhaité faire de même avec les documents textes.

Pour ce faire, leur souhait fut de trouver des systèmes pour extraire le texte brut et les métadonnées des documents pour pouvoir faire des recherches, mais aussi de chercher des moyens de les rendre facilement visualisables aux visiteurs et ce pour le plus grand nombre de documents « texte » possible (Microsoft Word et PowerPoint, OpenDocument, code source, ...). C'est pour effectuer ces recherches et pour y trouver des solutions concrètes que cette entreprise m'a pris en stage pour l'été 2007.

Dans la suite de ce rapport, dans un premier temps, je vous présenterai ipernity, son histoire, ses projets, sa situation actuelle, ... Ensuite, je détaillerai l'objectif de mon stage, les contraintes et les moyens dont je pouvais disposer et bien évidemment, la solution que j'ai proposée. Enfin, je décrirai les différentes étapes de mes recherches et du développement ainsi que les problèmes que j'ai pu rencontrer tout au long du stage.

## IPERNITY

### IPERNITY, AUJOURD'HUI

#### L'EQUIPE

Afin d'éviter de proposer une liste sûrement rébarbative des caractéristiques de chacun, j'ai décidé de faire quelque chose de plus original. J'ai proposé un questionnaire à tous les membres de l'équipe avec des questions sérieuses et d'autres un peu moins, ce qui permettra de bien cerner la personnalité de chacun. Vous trouverez l'intégralité des réponses de chacun en annexe [Réponses au questionnaire](#).



Christophe Ruelle

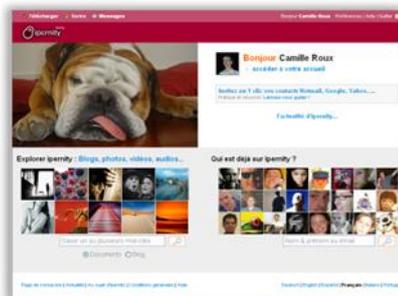


Christian Conti



Léa Cymbalista

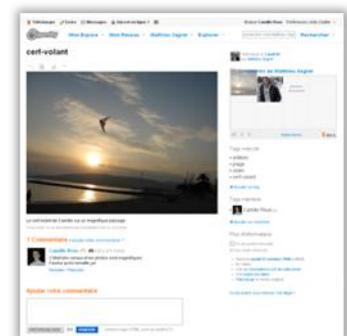
#### LE SERVICE PROPOSE



ipernity est un service grand public de partage multimédia et communautaire. Les utilisateurs peuvent mettre en ligne tout type de documents (blog, photos, vidéos, mp3, ...) dans un même espace (contrairement à la tendance actuelle qui consiste à se spécialiser sur un type précis de documents : Flickr pour les photos, Dailymotion pour les vidéos, ...).

Ce site internet moderne est destiné à un public très large. Il peut être utile aux familles et aux groupes d'amis pour sauvegarder et partager leurs photos de vacances ou de soirées. Il peut servir aussi à la promotion des réalisations de créateurs. Il y a déjà de nombreux photographes qui mettent en ligne leurs œuvres grâce à ipernity.

Ce service ajoute une autre dimension aux documents envoyés. Comme sur la plupart des sites communautaires actuels, il est possible de poster des commentaires sur un document, de partager ses photos avec seulement ses amis ou sa famille, de chatter en ligne ou d'envoyer un mail à un contact, ... Ces fonctions sont très appréciées car elles permettent aux membres de se rencontrer ou encore d'avoir des avis extérieurs sur les documents qu'ils postent et les articles qu'ils écrivent.



ipernity en plus de proposer l'hébergement, permet bien évidemment la consultation des documents en ligne. Pour permettre au visiteur de ne pas être perdu dans la masse de documents

et de lui permettre de trouver ce qu'il recherche, de nombreux outils sont à sa portée. Le site propose un système de classification de documents très utilisé dans le monde du web 2.0 : la folksonomie<sup>1</sup>. Il est possible de tagger les documents (leur associer des mots-clés) que l'on met en ligne. Ceci permet aux gens de facilement trouver des documents grâce à un moteur de recherche. Il est également possible d'associer un lieu (géotag) et donc de faire des recherches par position géographique. Le système de recherche du service est très complet ; il est, en outre, possible d'effectuer des recherches par couleur, licence, utilisateur, ...



Afin de rendre le site plus attrayant et facile à utiliser, les développeurs ont utilisé des technologies web moderne, notamment l'AJAX<sup>2</sup>. Ces technologies permettent par exemple, d'éditer certaines informations sans changer de page ou encore d'écrire des articles de blog avec une interface proche d'un traitement de texte classique.

## LE MODELE DE REVENUS

ipernity est une société anonyme présidée par Christophe Ruelle. Son siège est basé sur la technopole de Sophia-Antipolis, à Mougins-le-haut. Ses activités et son développement sont financés sur fonds propres. ipernity fonde son modèle économique sur 3 sources de revenus :

- Les abonnements de l'offre **pro** et de la vente de services exclusifs à la carte. Voici le comparatif entre compte gratuit et compte pro :

	Compte gratuit	Compte PRO
<b>Volume de contenu supplémentaire publiable chaque mois (Upload)</b>	200 Mo / mois par exemple 200 photos de 1 Mo	1000 Mo / mois par ex. 1000 photos de 1 Mo
<b>Disponibilité du format original des photos, vidéos et audios</b>	NON	OUI
<b>Disponibilité des documents publiés</b>	Les 1000 derniers	TOUS
<b>Insertions publicitaires dans les pages liées au compte</b>	Les pages liées au compte peuvent contenir de la publicité	Les pages liées au compte sont exemptes de publicité
<b>Prix</b>	Gratuit	1, <sup>99</sup> € / mois

- La publicité et plus généralement, la monétisation d'une partie de l'audience.
- La fourniture de services par des tiers (ex : tirage photo, impression sur T-shirts, ...)

<sup>1</sup> Une **folksonomie** est un néologisme désignant un système de classification collaborative décentralisée spontanée. Le concept étant récent, sa francisation n'est pas encore stabilisée, bien que le grand dictionnaire terminologique l'ait ajouté à sa base de données. Le terme folksonomie est une adaptation française de l'anglais folksonomy, combinaison des mots folk (le peuple, les gens) et de taxonomy (la taxinomie). (Wikipédia France - Folksonomie 2005)

<sup>2</sup> **AJAX**, ou **Asynchronous JavaScript And XML** (« XML et Javascript asynchrones »), est un acronyme désignant une solution informatique libre pour le développement d'applications Web. (Wikipédia France - Asynchronous JavaScript And XML 2007)

## QUELQUES CHIFFRES

Je trouvais intéressant de publier des chiffres concrets afin de caractériser la popularité, l'ampleur ou encore la croissance du site. Cependant, il est difficile pour une startup naissante de communiquer là-dessus car ces données peuvent avoir un intérêt pour la concurrence ou peuvent ne pas être très avantageuses du point de vue de l'image (aucun site n'a des millions de visites dès les premiers jours...).

La société m'a gracieusement tout de même autorisé à publier quelques informations chiffrées. Aujourd'hui, ipernity, c'est :

- **4** mois d'activité
- **plusieurs dizaines de milliers** de membres
- **plusieurs millions** de pages vues par mois
- bientôt **1 million** de documents
- des visites d'une durée de **15 minutes** en moyenne
- une croissance à **2 et 3 chiffres** comme beaucoup de startup
- une audience **1/3** allemande, **1/4** nord-américaine et **1/5** français (cela va fortement évoluer lors de l'ouverture du site en **4** autres langues en octobre)

(Conti, Quelques chiffres sur ipernity (mail) 2007)

## IPERNITY, DE LA PREMIERE IDEE A NOS JOURS

### HISTORIQUE



Christophe Ruelle et Christian Conti se connaissent depuis longtemps. Ils sont, par exemple, tous les deux des anciens élèves de l'ESSI. Ils ont également fréquenté la même pépinière d'entreprise à leur début. A ce moment là, ils faisaient partie des quelques rares pionniers français à voir le potentiel que pouvait représenter le web et à avoir quelques idées pour l'exploiter. Ensuite, ils se sont tous les deux lancés dans le monde des startups puis ont été adossés à de grands groupes de télécommunication (France Télécom pour Christophe et LibertySurf pour Christian). Tout a bien marché pour eux pendant plusieurs années, puis ils ont passé le relais en 2001.



Christophe a une problématique en tête depuis longtemps. Il constate que de plus en plus de documents numériques sont créés, mais qu'aucune disposition n'est prise par leurs auteurs pour éviter de les perdre. Une photographie papier va être rangée dans un tiroir et il est fort possible qu'elle soit toujours là et dans un bon état des dizaines d'années après. Si on prend le cas d'une photo numérique sauvée sur disque dur, c'est très différent, il est quasi impossible que les fichiers ne soient pas effacés par mégarde, que le disque dur tombe en panne ou soit formaté, ... De ce fait, une très grosse partie de notre patrimoine numérique va malheureusement être perdu. Christophe réfléchit alors à des manières de sauvegarder ces données, de les organiser ou encore de les mettre

en valeur. Rapidement, il crée ipernity<sup>3</sup> dont le nom est la contraction de IP<sup>4</sup> et « eternity » (éternité). Pendant ce temps, Christian visite le monde entier, les valises à la main.

A cette époque, ipernity n'est pas du tout le site qu'on connaît aujourd'hui. Il s'agit d'une entreprise<sup>5</sup>, qui moyennant finances, recueille des récits, photos, vidéos, ... sur la vie du client afin d'écrire une biographie ou une chronique familiale. Ce service était loin d'être accessible<sup>6</sup> à tous et, de plus, les gens n'étaient probablement pas prêts. Peu de personnes ont aujourd'hui conscience de la fragilité des documents numériques, et donc du fait que dans quelques dizaines d'années il n'en restera probablement aucune trace. Partant de ce constat, Christophe réfléchit à des concepts plus accessibles. Pour baisser les coûts, il faudrait que les gens recueillent et stockent eux-mêmes les données qu'ils souhaitent préserver et partager au lieu d'avoir à payer des gens qualifiés pour le faire. La finalité elle-même, devrait être plus simple qu'un livre ou un CD-ROM. En pensant à tout ça, Christophe est à ce moment là, en train de définir les concepts de base d'ipernity.com. Il est intéressant de remarquer que l'objectif, en revanche, reste le même : *conserver et valoriser le patrimoine de mémoire de tout-un-chacun*.



En 2005, le premier prototype de publication en ligne voit le jour. Cette première version ne convenait pas tout-à-fait à l'idée que Christophe pouvait en avoir. Suite à ça, il parcourt le web à la recherche d'inspiration pour améliorer le service qu'il développe et découvre Flickr<sup>7</sup>. Il y trouve de nombreuses idées et concepts innovant qu'il ajoutera par la suite. Afin de créer un service le mieux adapté possible à la demande du marché, il s'inspire aussi de plusieurs réseaux sociaux tel MySpace<sup>8</sup>.



Christian rejoint le projet en 2006 pour assurer le développement économique et pour apporter son expérience en matière de sites web communautaires. Assisté par T., Christophe redéveloppe entièrement le site en y ajoutant des technologies (Ajax) et des concepts (tags, nuages de tags, ...), apparus entre temps, provenant du web 2.0.

<sup>3</sup> Vous avez probablement remarqué que « ipernity » ne prend jamais une majuscule, ceci, pour des raisons esthétiques.

<sup>4</sup> **Internet Protocol**, protocole utilisé pour le routage des paquets sur les réseaux

<sup>5</sup> Le site internet est encore accessible : <http://ipernity.org>.

<sup>6</sup> Ce service coûtait 990€ et donnait lieu à la réalisation d'un livre et d'un CD-ROM

<sup>7</sup> **Flickr** est un service d'hébergement de photos développé par Ludicorp, une société canadienne de Vancouver fondée en 2002. Ce site est généralement considéré comme un symbole du Web 2.0.

<sup>8</sup> **MySpace** est un service de réseautage social en ligne fondé aux États-Unis par Tom Anderson et Chris DeWolfe en Aout 2003, qui met gratuitement à disposition de ses membres enregistrés un espace web personnalisé, permettant d'y faire un blog, d'y envoyer ses photos et d'y remplir diverses informations personnelles. (Wikipedia France - MySpace 2007)

## LES PROBLEMES RENCONTRES

Comme toute entreprise, ipernity a parfois dû faire face certaines difficultés. On peut tout d'abord citer à nouveau la remise en question en 2005 dont j'ai parlé dans l'historique. Actuellement, la plus grande inquiétude d'ipernity est qu'une société web 2.0 très connue arrive sur leur marché. Ceci rendrait probablement leur site quasi-invisible car il serait extrêmement difficile de faire le poids. ipernity a déjà été confronté à un problème de ce type.

En mai 2007, ipernity a fait traduire la totalité de l'interface par des membres volontaires (qui en échange, ont reçu plusieurs comptes PRO). L'équipe de développement mit rapidement les traductions en ligne par la suite. C'est alors qu'un élément perturbateur arriva, le 12 juin précisément. Ce jour-là, Flickr a lui aussi mis en ligne une version multilingue de son site. Cet événement enleva une partie de l'avance qu'ipernity pouvait avoir sur ce concurrent. Heureusement pour la startup de Mougins-le-haut, juste après ça, Flickr a dû censurer du contenu pour les membres allemands à cause d'une loi qui est apparue peu de temps avant dans leur pays. Grâce cette actualité, ipernity a vu de nombreux membres de qualité (souvent des photographes) venir sur le site.

## IPERNITY DANS UNE BOULE DE CRISTAL

Après avoir vu comment ipernity en est arrivé là, il est intéressant de faire un bond dans le futur et de voir ce à quoi cette société pourrait ou devrait ressembler dans quelques années.

## LES OBJECTIFS

ipernity ne fait que démarrer et donc leur principal souhait, pour le moment, est de faire la démonstration que leur site fait partie de ce qu'il se fait de mieux actuellement. L'important, dans un premier temps, est que les gens aient envie de rester sur le site et qu'ils y fassent venir leurs amis ou des membres de leur famille. Si tel est le cas, alors une « dépendance » va se créer entre les membres et le site, ce qui permettra à ipernity d'avoir un bel avenir.

Un objectif concret et réalisable, d'après Christian Conti, serait de devenir rentable d'ici 3 ans, soit courant 2010. Ceci semble tout à fait possible et il suffit d'étudier rapidement la popularité des services similaires (Flickr, MySpace, Dailymotion, YouTube, ...) pour en avoir le cœur net.

## LES PROJETS FUTURS

Il y a pour le moment aucune réelle révolution du site dans les cartons d'ipernity, ni aucun élargissement d'activité prévu. Ceci n'aurait d'ailleurs pas vraiment de sens. La société a tout intérêt à se concentrer sur l'activité principale vu que le site est tout jeune et que le concept attire de plus en plus de gens.

Cependant, le site va tout de même évoluer pour répondre aux attentes des membres. Voici quelques fonctionnalités qui vont apparaître dans les prochains mois :

- Traduction du site dans de nouvelles langues
- Tirage photo

- Chat à plusieurs (réalisé par Matthieu Sivade, qui a réalisé son stage en même temps que moi)
- Diaporama Flash à exporter
- Et bien sûr, toutes les fonctionnalités apportées par la plateforme de conversion que j'ai réalisée pendant le stage.
- Dans un futur plus lointain, l'équipe d'ipernity imagine proposer une application web permettant la réalisation de films à partir des photos, vidéos et musiques téléchargées sur le site (à la <http://animoto.com>), ou encore des services spécialisés de retouche photo (colorisation, suppression du bruit ou de l'effet yeux-rouges, ...)

## IPERNITY, DANS LA JUNGLE DU WEB 2.0

Aujourd'hui, de nombreux sites proposent d'héberger des documents en ligne, de les partager et d'y apporter une certaine valeur ajoutée. On peut par exemple citer les exemples de l'hébergement photos (ex : Flickr) et de l'hébergement de vidéo (ex : Dailymotion, Youtube,...), qui sont à présent des marchés très saturés où il est extrêmement difficile de s'imposer. Comme beaucoup de sociétés présentes dans le monde du web 2.0, celles-ci sont localisées pour la plupart aux Etats-Unis. Ces sites tiennent un place très importante sur la toile. Par exemple, Youtube représente plus de 13% du trafic internet mondial à lui seul et est le 4eme site le plus visité apres Yahoo, Microsoft Network et Google (source Alexa).



Afin de pouvoir se démarquer, ipernity a choisi de proposer l'hébergement de tous les documents en essayant, pour chaque type, de fournir le plus de valeur ajoutée possible (tag, commentaires, géolocalisation, tag membre, ...). La tendance étant à la séparation des services, ce marché est une réelle niche pour le moment. Aucun grand acteur du web 2.0 n'est sur ce marché. Le marché paraissant intéressant, on peut penser que la société ne va pas rester seule longtemps. Par exemple, Flickr va se rapprocher d'eux dans quelque temps quand il proposera l'hébergement

de vidéo<sup>10</sup>. La grosse différence entre Flickr et ipernity se situera au niveau de la valeur ajoutée à la musique et aux documents textes. Ceci montre bien l'intérêt de mon stage.

Une autre fonctionnalité d'ipernity que je qualifierai de secondaire vis-à-vis de l'hébergement de document, mais qui est tout de même très importante, est la présence d'un réseau social. Et là aussi le marché est très saturé. De grands noms sont déjà présents, tel MySpace qui est le 6eme site le plus visité au monde (source Alexa). La différence entre ces sites et ipernity n'est pas au niveau du réseau social lui-même (qui présente toujours les mêmes fonctionnalités), mais plutôt par rapport à ce qu'il y a autour. MySpace propose l'hébergement de fichiers mais de manière assez limitée. De ce fait, ipernity reste plus intéressant pour l'hébergement et le partage de documents que beaucoup de réseaux sociaux.

Pour résumer, ipernity présente l'avantage de ne pas être limité à certains fichiers et propose un espace de stockage assez important comparé aux autres services. Si ce site devient populaire, il ne serait pas étonnant de voir la concurrence arriver...

## IPERNITY, VUE DE L'INTERIEUR

ipernity est une petite entreprise de 4 personnes. De ce fait, il n'est pas évident de parler de la communication, de l'information ou encore de rituels et de symboles de l'entreprise (pour ces derniers voir les interviews des membres de l'équipe en annexe [Réponses au questionnaire](#)). Ce que je peux dire, c'est que les gens qui travaillent ici sont plus des amis que de simples collègues.

L'entreprise étant à échelle humaine, la communication y est très facile. Il n'est pas rare que le « département développement » demande l'avis du « département marketing » pour des questions importantes qui mettent en jeu l'ergonomie, le design, ... Ceci a probablement dû contribuer à l'unité du site et permet un développement plutôt rapide du site.

Cependant, l'entreprise se doit de répondre parfois en urgence aux attentes de la communauté des membres. ipernity est aussi en lien avec de nombreuses entreprises pour la recherche de partenariats, la demande de conseils, sans oublier les services commerciaux et techniques. De ce fait, il n'est pas rare que des discussions soient interrompues par un coup de téléphone ou un problème survenu sur la plateforme. De plus, comme la quantité de travail est assez conséquente, il est assez fréquent que des personnes de l'équipe soient occupées et donc pas disponibles pour répondre à une question ou discuter sur un point précis. Ceci n'est probablement pas étonnant pour une startup de cette taille.

---

<sup>10</sup> <http://www.techcrunch.com/2007/08/03/reconfirmation-flickr-to-add-video/>

## LE TRAVAIL PROPOSE

Après avoir décrit en détail la société, il est tant de se pencher sur le stage pour lequel j'ai postulé. Je vais définir le travail qui m'a été demandé de réaliser, les objectifs, les moyens à disposition et les contraintes diverses qui m'ont été données.

### LE SUJET

Voilà le sujet original :

*Nous le sentons bien, l'année 2007 sera l'année de l'indexation intelligente de tous vos fichiers. Entre Google Desktop Search, Microsoft avec Vista et Spotlight d'Apple, jamais autant d'applications se sont proposées d'analyser le contenu des fichiers de votre ordinateur.*

*ipernity.com, dans son approche d'organisation et d'archivage de vos fichiers en ligne, se doit de proposer une fonction de recherche (et/ou de prévisualisation) de nombreux formats de fichiers. Nous supportons actuellement tous les formats d'image, de vidéo et les fichiers audio (avec id3).*

*L'objectif du stage est d'implémenter un moteur d'indexation capable d'analyser le contenu de fichiers PDF, PS, WORD, EXCEL, RTF, POWERPOINT, ... ainsi que le plus grand nombre de formats de fichiers contenant du texte, ou d'autres indications susceptibles d'être indexées (dates de création, de modification, auteurs, révisions,...).*

*Dans un environnement essentiellement constitué de logiciels libres (LINUX/PHP/MYSQL,) l'étudiant devra réaliser un moteur d'indexation aux meilleures performances possibles (qualité de l'indexation / temps CPU). Les données d'indexation seront stockées en base MySQL, la recherche sera faite au moyen de scripts PHP.*

*Connaissances particulières requises : PHP, PERL, SQL mais aussi C, C++, algorithmique*

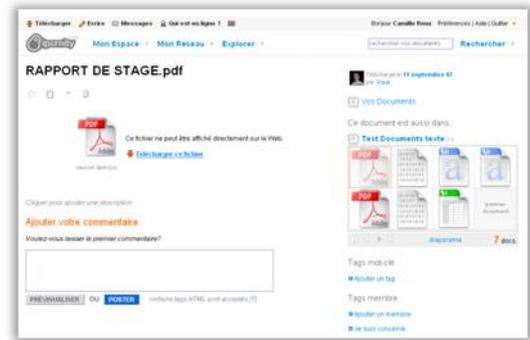
### QUELQUES PRECISIONS

Comme précisé dans le sujet, mon but au départ était de proposer une plateforme d'indexation. Je devais arriver à extraire le texte brut du plus grand nombre de documents possibles puis le stocker avec un système d'index pour pouvoir faire des recherches facilement dessus, tout ceci en tenant compte du temps de calcul et de la qualité de l'indexation.

Avant la mise en ligne de ma réalisation, la seule métadonnée accessible sur des documents mis en ligne est le nom du fichier. Souvent peu explicite, le membre doit donc ajouter lui-même des informations (descriptions, tags, changer le titre, ...) pour faciliter l'accès à ses documents. Celui-ci n'avait donc que peu de valeur ajoutée.

Par la suite, le sujet a évolué pour répondre un peu mieux aux attentes que pouvaient avoir les membres et pour mieux coller à la philosophie du site. Au fur et à mesure des discussions et de mes recherches, nous avons mis en exergue quelques points.

J'ai rapidement pris la décision de transformer tous les fichiers téléchargés sur le site dans un format central unique afin d'appliquer un traitement quasi-identique à chaque fichier. Le choix d'un format unique m'a non seulement simplifié l'opération d'extraction du texte brut mais nous a permis d'avoir facilement accès à d'autres traitements comme la génération d'aperçu, la génération d'animation flash, l'extraction des images, ... (pour plus d'information voir l'annexe ()). Avant mon stage, les documents étaient représentés par l'icône du format du fichier d'origine et parfois son titre. Ceci rendait très difficile le classement des documents par album car on ne pouvait pas bien identifier les documents. De ce fait, nous avons décidé d'aller plus loin que ce qui avait été prévu au départ pour mon stage et de profiter du format central unique pour générer des miniatures afin de remplacer les icônes. J'ai donc dû réaliser une plateforme qui serait non seulement capable d'indexer le contenu des documents mais aussi de convertir les documents en des formats plus facilement visualisables que le fichier d'origine sur un navigateur classique.



## LE CAHIER DES CHARGES

Je viens de détailler le travail qui m'était demandé pendant ces 13 semaines, voici à présent les moyens dont je disposais et les contraintes qui m'ont été imposées.

## CONTRAINTES TECHNIQUES

Comme ma plate-forme allait faire partie intégrante du site actuel, j'ai dû m'adapter à certains choix qui avaient été faits. Par exemple, le site étant développé en PHP/Perl, j'ai dû utiliser ces langages pour qu'ils puissent interagir avec ma partie.

J'ai dû m'adapter également à l'architecture serveur du site. Comme tout site web ayant un trafic important, la charge est répartie sur plusieurs serveurs. Par exemple, quand une requête est faite, elle arrive tout d'abord sur un répartiteur de charge qui va l'envoyer sur un serveur web choisi de manière aléatoire. Il devait pouvoir être fait de même avec la plateforme que j'avais à réaliser. Les temps de conversion peuvent être très longs (jusqu'à une minute pour certains fichiers). Ma plateforme devait donc être « clusterisable », c'est-à-dire qu'elle devait permettre de lancer des conversions simultanées sur plusieurs serveurs en même temps.

Bien évidemment, il était attendu que la plateforme soit la plus rapide et la plus stable possible. Comme j'allais utiliser de nombreuses applications externes, la question de la stabilité était un point très important. Ce paramètre m'a aussi demandé de penser mon application de manière totalement modulaire. Il devait être simple de remplacer une application externe par une autre si cela permettait, par exemple, d'améliorer la stabilité ou la vitesse de conversion.

## CONTRAINTES UTILISATEURS

Après avoir expliqué les moyens dont je disposais et les contraintes techniques qui leur étaient liées, je vais détailler les contraintes en rapport, cette fois, avec l'utilisateur.

Le point, probablement le plus important, était que je sois capable de générer des documents de qualité. La mise en page, le format, ou encore, le style devaient être préservés quel que soit le support de destination et le fichier d'origine. Je devais aussi prendre soin de la qualité des supports non vectoriels (images, animations Flash, ...).

J'ai dû aussi tenir compte d'un paramètre important : la compatibilité. Effectivement, comme les fichiers que j'avais à générer seraient affichés directement dans les pages du site, je devais m'assurer qu'ils seraient visualisables par tous. Par exemple, il ne serait pas question de générer des PDF qui ne sont visualisables qu'avec la toute dernière version d'Acrobat Reader.

J'avais également certaines contraintes au niveau de la pertinence des résultats lors d'une recherche plein texte sur ces documents. Il fallait qu'il soit facile de retrouver un document téléchargé même parmi un grand nombre.

## LE TRAVAIL REALISE

### CARACTERISTIQUES GENERALES DE LA PLATEFORME

Le meilleur moyen de comprendre le résultat final de mon travail est de regarder le schéma de la plateforme de conversion.



Une fois le fichier en PDF, la plateforme va générer différents fichiers qui permettront de faciliter la visualisation du document. Tout d'abord, la plateforme génère des miniatures de plusieurs tailles différentes (voir annexe [Exemple de miniatures](#)) avec l'extension du fichier d'origine dans un coin. Ensuite, elle extrait le texte brut du document PDF.

Enfin comme l'indique le schéma, j'ai également fait des recherches sur d'autres formats de sortie. J'ai notamment trouvé des moyens de convertir des PDF en animations Flash (ce qui permet de consulter le document directement dans une page, contrairement à Acrobat Reader) et aussi d'extraire les images contenues dans le fichier. Cependant, ces solutions n'ont pas été intégrées à la plateforme, mais le seront peut-être dans un futur plus lointain.

Pour tous mes choix, j'ai tenu à faire attention à plusieurs points qui me paraissaient importants pour garantir la stabilité, la fiabilité ainsi que sa capacité à évoluer.

### DEROULEMENT DU STAGE

Je vais à présent décrire ma démarche pour le développement de la plateforme de conversion. Je vais présenter les grandes étapes. Je ne vais en revanche pas parler ici des travaux annexes qui ont été développés très rapidement (1-2 au maximum), ni de l'écriture du rapport technique, ni des tâches concernant la plateforme mais qui ont été de courte durée.

<b>18 juin</b>	Recherches préliminaires
<b>25 juin</b>	Recherches préliminaires + Recherches sur l'indexation
<b>2 juillet</b>	Recherches sur l'indexation (+travaux annexes)
<b>9 juillet</b>	Recherches sur l'indexation + Script PHP pour la plateforme
<b>16 juillet</b>	Script PHP pour la plateforme (+travaux annexes)
<b>23 juillet</b>	Script PHP pour la plateforme + premiers tests
<b>30 juillet</b>	Test et débogage de la plateforme (+travaux annexes)
<b>6 août</b>	Test d'un système d'indexation
<b>13 août</b>	Vacances
<b>20 août</b>	Test et débogage de la plateforme + Amélioration de la génération des aperçus
<b>27 août</b>	Test et débogage de la plateforme
<b>3 septembre</b>	Finalisation de la plateforme + Ecriture de tests pour l'API
<b>10 septembre</b>	Rédaction du rapport de stage + Ecriture de tests pour l'API
<b>17 septembre</b>	Rédaction du rapport de stage + Ecriture de tests pour l'API

## TRAVAUX ANNEXES

En plus de développer la plateforme de conversion, il arrivait régulièrement que Christophe ou T. me fassent travailler sur des tâches très diverses sans liens particuliers avec le sujet du stage. J'ai beaucoup apprécié chercher des réponses à ces questions qu'ils se posaient. Elles étaient diverses, mais très instructives.

Voici quelques sujets sur lesquels j'ai été amené à réfléchir:

- Etudes de solutions existantes d'indexation de documents
- Conversion de format HTML vers PDF
- Etude pour la réalisation d'un script GreaseMonkey (plugin Firefox permettant d'ajouter des scripts JavaScript externes dans les pages visitées) pour importer les photos de Flickr facilement
- Extraction des images des fichiers uploadés
- Protection des PDF hébergés
- Extraction des métadonnées (notamment les données EXIF) d'une image
- Etude de failles de sécurité Flash/JavaScript (afin de voir les possibilités d'afficher les animations Flash téléchargées par les membres directement dans des pages du site)
- Bloquer immédiatement l'upload de fichiers trop importants (et ne pas attendre la fin du téléchargement pour vérifier la taille)
- Protection du site contre les attaques DoS<sup>13</sup>
- Extraction du texte et des images des documents openXML<sup>14</sup>
- Indexation et recherche de documents
- URL rewriting pour rediriger les adresses de la forme [www.xxx.ipernity.com](http://www.xxx.ipernity.com) vers [xxx.ipernity.com](http://xxx.ipernity.com)
- Personnalisation du Player vidéo en Flash (tâche annulée par l'équipe après un jour de travail)
- Ecriture de tests unitaires pour l'API



---

<sup>13</sup> Le **déni de service** ou **Denial of Service (DoS)** est, d'une manière générale, l'attaque qui vise à rendre une application informatique incapable de répondre aux requêtes de ses utilisateurs. Une machine serveur offrant des services à ses clients (par exemple un serveur web) doit traiter des requêtes provenant de plusieurs clients. Lorsque ces derniers ne peuvent en bénéficier pour des raisons délibérément provoquées par un tiers il y a déni de service.

<sup>14</sup> Format utilisé par Microsoft Office 2007 (format qui n'est pas actuellement géré par OpenOffice.org)

## CONNAISSANCES UTILISEES ET ACQUISES

Le but de ce stage n'était pas d'écrire une grosse application avec une structure complexe, mais plutôt de relier des applications, écrites dans des langages assez divers (C, C++, Python, Perl, Java, ...), réalisant chacune une tâche particulière nécessaire à la plateforme. De fait, j'ai passé beaucoup de temps à compiler et tester des applications, à écrire des scripts pour effectuer des benchmark<sup>15</sup>, ... qu'à écrire du code structuré.

Le sujet du stage ne m'a pas permis de mettre en application certains formalismes et connaissances appris en 2eme année comme l'UML ou encore les schémas de base de données. En revanche, j'ai été amené à apprendre ou améliorer mes connaissances sur de nombreux langages comme le Perl, le Python, le PHP, le script Shell, le Java, Lex/Yacc, XHTML, ... De plus, afin d'être le plus proche possible de l'environnement de production, j'ai travaillé sur la même distribution linux que sur les serveurs d'ipernity, CentOS. Grâce à ça, j'ai dû utiliser de manière très intensive le Shell Linux et donc de mettre en pratique les cours linux de 1ere année.

J'ai également mis en pratique certaines connaissances en réseaux afin de faire communiquer par socket TCP un programme Java avec un script PHP. Aussi, j'ai eu à compiler et paramétrer un serveur web Apache + PHP, ce qui m'a permis de mettre en application les quelques bases fournies à ce sujet par l'école et d'en apprendre beaucoup plus. Dans cette liste de connaissances utilisées, je pourrais ajouter, la compilation. En effet, j'ai compilé des programmes sous Linux, ce que, curieusement, je n'avais jamais fait auparavant.

Les travaux annexes ont été, eux aussi, des occasions d'apprendre ou parfaire certaines connaissances. Ne pouvant pas toutes les citer, je ne donnerai que les plus importantes: test unitaires (pour tester l'API qui va être rendue publique), JavaScript, configuration et ajout de module Apache, Action Script, ...

Les apports du stage sont au final très variés, mais concernent tous d'assez près le monde du web, domaine vers lequel je souhaite me diriger. L'école néglige ces connaissances alors qu'elles sont exigées par de nombreuses propositions de stage (environ 1/3 des stages). Il m'a toujours paru étonnant que le PHP et le Perl ne fasse pas parti du tronc commun, alors que des cours qui me paraissent vraiment secondaires pour une formation d'ingénieur en informatique en font partis. A mon sens, tous les élèves devraient sortir de l'école en étant capable de construire un site web complet en utilisant un Framework, des bibliothèques Ajax, ... Il est vrai qu'une option, CLAW, propose cet enseignement, mais les technologies qui y sont étudiées, sont loin d'être en phase avec celles utilisées aujourd'hui par des entreprises comme ipernity.

---

<sup>15</sup> Un **test de performance** ou **benchmark** est un test dont l'objectif est de déterminer la performance d'un système ou un logiciel informatique.

Pour conclure, je suis très heureux d'avoir pu faire ce stage qui, comme les projets que j'ai réalisés en 2ème année, a pu combler les lacunes que l'école peut avoir à propos des technologies de l'internet.

## RELATIONS DANS L'ENTREPRISE

Le fait que l'entreprise soit de petite taille facilite grandement la communication en comparaison à une grande entreprise. Il est facile de discuter avec une personne, poser une question à n'importe qui, ... Le petit nombre de personnes permet aussi de rapidement connaître les compétences et fonctions de chacun et donc de savoir à qui poser une question.

Cette proximité m'a permis de bien connaître les gens qui travaillent ici. J'ai pu avoir de nombreuses informations sur leur parcours et aussi comprendre comment ils en sont arrivés là. Ils travaillent dans un domaine qui me passionne, le web, et sont dans une structure que j'affectionne tout particulièrement, la start-up. Cette forme d'entreprise permet de réaliser un projet en suivant toutes ses étapes et d'avoir des retours directs de ce que l'on fait (ici de la part des membres de la communauté). Une start-up, c'est aussi un défi personnel. On prend des risques. Il y a de gros enjeux. Tout ceci donne un intérêt supplémentaire au travail qu'on fait. Tout ceci pour dire, que j'ai beaucoup appris des discussions que nous avons pu avoir.

Cependant, une start-up représente aussi un travail considérable pour chacun des employés. De ce fait, il arrivait qu'on ne puisse pas avoir une réponse à une question dans l'instant, ce qui a pu ralentir quelque peu mon travail par moment. Aussi, ce problème a pu causer également un certain manque d'attention à mon égard et au travail que je réalisais.

## DIFFICULTES RENCONTREES

Je pense que la plus grosse difficulté de ce stage a été le fait d'avoir à tester puis utiliser de nombreuses applications externes. La conversion de documents étant une opération bien souvent complexe, j'ai été confronté de nombreuses fois à des documents qui étaient mal convertis alors que beaucoup marchaient. Il n'est pas vraiment étonnant que ce genre de problèmes arrive vu la complexité de certains formats. J'ai aussi dû avoir à faire à des problèmes de stabilité de la part de certains programmes. Pour garantir une certaine stabilité à la plateforme, j'ai pris certaines mesures assez draconiennes.

Il a également été très difficile pour moi de m'assurer que la plateforme fonctionnait parfaitement quelque soit le fichier d'entrée. Le nombre d'applications externe mises en jeu et le nombre de format de fichiers gérés ont rendu la tâche plutôt complexe. Pour la tester, je l'ai faite marcher sur les quelques milliers de documents textes téléchargés sur le site, puis j'ai vérifié les fichiers générés un par un. Il n'y avait aucun bon moyen d'automatiser le test...

Enfin, j'ai souvent été confronté à des comportements hasardeux de certains programmes sur quelques rares fichiers. Il est, en effet, arrivé que des documents parfois très simples fassent planter une application () de la plateforme pour des raisons difficiles à trouver. J'ai pu, la plus part

du temps, les résoudre en contactant les développeurs du logiciel en question, en changeant de version ou encore, en modifiant moi-même le code du programme dans certain cas. Quoi qu'il en soit, il y a de forte chance de voir apparaître de nouvelles erreurs de ce type quand la plateforme sera en production.

Ces nombreux aléas durant toute la durée du développement m'ont beaucoup ralenti. De plus, il était de ce fait, difficile de donner des prévisions sur les temps de développement. Ces problèmes ont été assez souvent ennuyeux et longs à résoudre. Cela a tout de même été instructif. C'était la première fois que je réalisais une application reposant sur de nombreux programmes externes, souvent instables. J'ai appris à gérer les comportements aléatoires et/ou imprévisibles d'applications qui peuvent se révéler instables.

## CONCLUSION

Quand j'ai passé mon entretien pour ce stage, j'avais dit que ce que je souhaiterais faire à la sortie de l'école, serait de réaliser des sites ou applications web dans des structures de petites tailles pour avoir la possibilité de suivre de près le projet tout au long du développement. En toute logique, je suis allé chez ipernity car cette entreprise semblait correspondre parfaitement à cette image.

Lors de ce stage, qui était pour moi le premier dans le domaine vers lequel je me dirige, j'ai pu analyser et comprendre le fonctionnement global d'une entreprise de ce genre. Il était intéressant de voir comment l'équipe gérait leur image, le stress à l'approche de deadline, les imprévus, les retards, les critiques de la communauté, les relations avec les partenaires et la presse, ...

Ce stage m'a permis d'apprendre tout ce que je souhaitais connaître à propos de ce genre d'entreprise et de leur activité. En discutant avec les membres de l'équipe et en les observant travailler, j'ai pu comprendre comment et pourquoi ils en étaient arrivés là, quelles sont les qualités qu'il faut avoir ou encore comment ils font pour avoir autant d'ambition, de motivation et d'énergie. Ils représentent, en quelque sorte, mon objectif.

Cette expérience m'a permis de voir de l'intérieur ce monde qui m'attire de plus en plus. Cela a été également pour moi l'occasion de confirmer mon envie de travailler dans la sphère très fermée du Web 2.0. J'aimerais avoir rapidement l'occasion de me retrouver dans une entreprise similaire et j'espère que nos routes seront amenées à se croiser à nouveau.

## BIBLIOGRAPHIE

Conti, Christian. *Communiqué de presse*. Sophia Antipolis, 25 Avril 2007.

Conti, Christian. *Dossier de presse ipernity*. Sophia Antipolis, avril 2007.

Conti, Christian. *Quelques chiffres sur ipernity (mail)*. Sophia Antipolis, 18 Septembre 2007.

Conti, Christian, interviewer par Camille Roux. *Questions diverses sur ipernity* (7 Septembre 2007).

*Wikipédia France - Asynchronous JavaScript And XML*. 5 Septembre 2007.

[http://fr.wikipedia.org/wiki/Asynchronous\\_JavaScript\\_And\\_XML](http://fr.wikipedia.org/wiki/Asynchronous_JavaScript_And_XML) (accès le Septembre 5, 2007).

*Wikipédia France - Folksonomie*. 25 Juillet 2005. <http://fr.wikipedia.org/wiki/Folksonomie> (accès le Septembre 7, 2007).

*Wikipedia France - MySpace*. 6 Septembre 2007. <http://fr.wikipedia.org/wiki/MySpace> (accès le Septembre 10, 2007).

*Wikipédia France - Portable Document Format*. 6 Septembre 2007.

[http://fr.wikipedia.org/wiki/Portable\\_Document\\_Format](http://fr.wikipedia.org/wiki/Portable_Document_Format) (accès le Septembre 18, 2007).

## ANNEXES

### REPONSES AU QUESTIONNAIRE

#### CHRISTIAN CONTI

**Quelles études as-tu faites?**

*ESSI option informatique temps réel + DESS ingénierie des ressources humaines*

**Qu'as-tu fait avant de venir chez ipernity?**

*Entrepreneur / Globetrotter ;-)*

*- 1991 consultant en NTIC*

*- 1994 fondateur d'une Web agency (Ove interactive)*

*- 1998 co-fondateur de Respublica.fr (plate-forme de communautés virtuelles francophones rachetée par le groupe LibertySurf en 1999)*

*- 2000 membre du comité de direction de LibertySurf puis Tiscali, en charge du pôle communautés et web radios*

*- 2001 globetrotter*

**Quel est ton rôle dans l'entreprise?**

*En charge du développement économique... plus tout le reste qui fait le charme des micro-entreprises (gérer le quotidien administratif, vider les poubelles, donner des interviews...)*

**Décris ipernity en une phrase**

*"Your home for digital life", ça résume bien et pour une fois c'est de moi ;)*

**Quel est, pour toi, le meilleur argument pour convaincre quelqu'un de prendre un compte PRO?**

*Un jour, vous en aurez marre de courir après la gratuité sur internet. Un jour, vous aurez besoin de vous poser et de construire. Un jour vous comprendrez que pour moins de 2€ par mois, vous pouvez constituer durablement et valoriser votre patrimoine de mémoire numérique. Nous ne sommes pas pressés, ce jour viendra !*

**Pour toi, quel est le rituel (quotidien ou hebdomadaire) le plus important dans l'entreprise?**

*Communiquer tous les matins les chiffres de la veille.*

*S'interroger collectivement vers midi sur les possibilités de déjeuner ;-)*

**A ton avis, ipernity dans 5 ans, c'est...?**

*Google, Yahoo... ou plus rien du tout !*

**Donne-moi un secret ou une info en exclusivité mondiale sur ipernity**

*Hé hé bien tenté ! Niet. Il faudra attendre un article dans Nice Matin !*

**Si tu étais une entreprise web 2.0, laquelle serais-tu ? Pourquoi?**

*Mais nous sommes nous-mêmes une entreprise Web 2.0, pourquoi envier les autres ?*

**Que t'as apporté notre présence (nous, les stagiaires) cet été?**

*De la fraîcheur, de la diversité, de la compétence aussi ;-)*

**LEA CYMBALISTA****Quelles études as-tu faites?**

*IUT info-com (sophia) puis DESS ingénierie de la communication et nouvelles technologies (Nice)*

**Qu'as-tu fait avant de venir chez ipernity?**

*Postes occupés : responsable projet web-marketing*

*Dernier emploi : Reed Exhibitions - salon du livre de paris, Fiac, Paris photo - Sites Internet des salons  
+ supports de communication*

**Quel est ton rôle dans l'entreprise?**

*Support, web-marketing, modération*

**Décris ipernity en une phrase**

*En français : partagez ce que vous voulez avec qui vous voulez*

*En anglais : home for digital life*

*:)*

**Quel est, pour toi, le meilleur argument pour convaincre quelqu'un de prendre un compte PRO?**

*Tout simplement la meilleure offre du marché !*

**Pour toi, quel est le rituel (quotidien ou hebdomadaire) le plus important dans l'entreprise?**

*Les pauses café, la distribution de gâteaux... :)*

**A ton avis, ipernity dans 5 ans, c'est...?**

*IN-CON-TOUR-NABLE !!!*

**Donne-moi un secret ou une info en exclusivité mondiale sur ipernity**

*Combien tu payes ?*

*Bon ok, prix d'amis : Bully est gay !*

**Si tu étais une entreprise web 2.0, laquelle serais-tu ? Pourquoi?**

*www.picnik.com : retouche d'images en ligne.*

*Pas encore très connu mais c'est le seul service web pour lequel je suis prête à payer :) Excellente intégration de la tendance "web 2.0" : utilité du service (un Photoshop simplifié mais tout aussi performant), ergonomie/design (simplicité, esthétique, intuitivité)*

**Que t'as apporté notre présence (nous, les stagiaires) cet été?**

*A part plus de fonctionnalités pour ipernity ?*

*De la bonne humeur, l'obligation (pour nous) d'arriver à l'heure le matin :))*

**T. M. (MAITRE DE STAGE)****Quelles études as-tu faites?***Bac S (spécialité physique)**Classes Prépa : PCSI**Ecole d'ingénieur en informatique***Qu'as-tu fait avant de venir chez ipernity?***Des maths, du code, des start-up, de la planche à voile et du vtt :)***Quel est ton rôle dans l'entreprise?***Concevoir, développer, optimiser des solutions techniques***Décris ipernity en une phrase***Home for your digital life :)***Quel est, pour toi, le meilleur argument pour convaincre quelqu'un de prendre un compte PRO?***Vivre pleinement cette aventure communautaire.***Pour toi, quel est le rituel (quotidien ou hebdomadaire) le plus important dans l'entreprise?****A ton avis, ipernity dans 5 ans, c'est...?***- un site communautaire indépendant qui a trouvé sa voie, son marché, son créneau et qui mène sa barque avec ses quelques milliers de membres payants.**- un site communautaire énorme racheté récemment par une grosse entreprise avide de start-up**- du passé? ;)***Donne-moi un secret ou une info en exclusivité mondiale sur ipernity***Heu... On est les meilleurs... mais chut!, ça doit rester secret!***Si tu étais une entreprise web 2.0, laquelle serais-tu ? Pourquoi?***Ben, ipernity, c'est bien du web 2.0?**Et parce que c'est la nôtre!***Que t'as apporté notre présence (nous, les stagiaires) cet été?***Du bon boulot... Une bonne ambiance... et qui sait peut-être de futurs collègues de travail?***Merci de m'avoir accordé un peu de ton temps si précieux ;)***Merci à toi pour tout ce que tu nous as apporté*

**CHRISTOPHE RUELLE****Quelles études as-tu faites?**

*1991-1992 : Prépa math sup/math spé + deug*

*1993-1995 : ESSI option calcul scientifique*

*1995-2000 : Echo interactive, CTO*

*2000-2003 : Wanadoo Portails, CTO*

*2001-now : Ipernity*

**Qu'as-tu fait avant de venir chez ipernity?**

*Echo interactive 96-2000*

*Wanadoo Portails / CTO / 2000 – 2003*

*ipernity 2001 -> now*

**Quel est ton rôle dans l'entreprise?**

*Chef d'orchestre*

**Décris ipernity en une phrase**

*A home for your digital life*

**Quel est, pour toi, le meilleur argument pour convaincre quelqu'un de prendre un compte PRO?**

*Je ne crois pas au modèle du tout gratuit financé par la pub. Lorsqu'un service est de qualité il doit pouvoir se permettre d'être payant, à un juste prix. C'est je crois aussi une demande de la part des utilisateurs qui préfèrent souvent payer que d'être inondés de publicité.*

**Pour toi, quel est le rituel (quotidien ou hebdomadaire) le plus important dans l'entreprise?**

*??? pas de rituel à mon avis mais un certain état d'esprit.*

**A ton avis, ipernity dans 5 ans, c'est...?**

*Flickr ;)*

**Donne-moi un secret ou une info en exclusivité mondiale sur ipernity**

*ipernity est entièrement codé en assembleur. Non franchement je ne vois rien de particulier à indiquer pour le moment, car je ne préfère rien annoncer qui ne soit pas déjà prêt. Rien en dehors de vos sujets de stage qui viendront bientôt étoffer l'offre ;)*

**Si tu étais une entreprise web 2.0, laquelle serais-tu ? Pourquoi?**

*On serait nous-mêmes. On n'est jamais aussi bien servi que par soi-même non ?*

**Que t'as apporté notre présence (nous, les stagiaires) cet été?**

*Au delà de l'amélioration de certaines fonctionnalités sujets de stage, un appui sur certaines opérations coup de poing et le plaisir de travailler et discuter avec des têtes bien faites.*

**Merci de m'avoir accordé un peu de ton temps si précieux ;)**

*N'exagons rien ;)*

## EXEMPLE DE MINIATURES

(Les images ont été rétrécies)

**Camille Roux**  
 Le Cavalier  
 81340 Saint-Cirq

**Camille Roux**  
 Né le 09/02/1986 à Antibes  
 Nationalité française  
 Permis B  
 Véhicule personnel

+33 5 63 56 46 09  
 +33 6 31 06 33 97  
 contact@camilleroux.com  
 http://camilleroux.com

**Formation**

2005-2007	<b>Polytech'Nice-Sophia-Antipolis</b> département Sciences Informatiques
2003-2005	<b>CPSE physique-chimie</b> au Lycée Lapérouse à Antibes (BAC +2)
2003	<b>Bac S option SI (Sciences de l'Informatique)</b> spécialité Mathématiques (mention Assez Bien)

**Compétences Informatiques**

<b>Systèmes d'exploitation</b>	Unix, Linux, Windows
<b>Langages</b>	Java, C/C++, PHP, JSP, SQL, Python, Ruby, sh (shell scripts), Scheme lex & yacc, Maple, SciLab, Postscript, VBA, XML, XSL, XSLT
<b>Technologies web</b>	(X)HTML, CSS, Javascript (+ frameworks AJAX), frameworks (Code Igniter, TurboGears, Ruby on Rails), Apache
<b>Base de données</b>	MySQL, PostgreSQL, SQLite
<b>Logiciels/Outils</b>	Eclipse, Visual Studio, Rational Rose, SVN, CVS Microsoft Office, OpenOffice.org, Photoshop, Dreamweaver

**Expériences professionnelles et projets d'études**

2007	<b>Projet d'études de 1 mois</b> : Développement d'une application web open source (Cero Repository) pour la gestion d'un dépôt de logiciels
2006	<b>Stage informatique de 2 mois</b> : Showroom assistant chez Peters BookSelling Service à Birmingham (Angleterre) <b>Projet de 2006</b> : Développement d'un jeu de cartes pour ordinateurs via Java (Cero Cards) <b>Projet personnel</b> : Développement d'un manager de jeu de cartes open source en Java (Cero Project)

**Compétences linguistiques**

<b>Français</b>	Langue maternelle
<b>Anglais (courant et technique)</b>	Bon niveau (TOEIC : 685 en décembre 2006)
<b>Espagnol</b>	niveau Bac

**Loisirs**

<b>Arts/sport</b>	Musique (Batterie, percussions, accordéon), cinéma Bâton de diable, cert-volant acrobatique Jeu de cartes
<b>Informatique</b>	Gestion et soutien de projets open source
<b>Robotique</b>	2 participations à la coupe de France de Robotique (2002 et 2003)

**Camille Roux**

**Informations**

**Adresse**

**Coordonnées**

**Informations professionnelles**

**Informations linguistiques**

**Loisirs**

**Camille Roux**

**Informations**

**Adresse**

**Coordonnées**

**Informations professionnelles**

**Informations linguistiques**

**Loisirs**

**Camille Roux**

**Informations**

**Adresse**

**Coordonnées**

**Informations professionnelles**

**Informations linguistiques**

**Loisirs**

```
<?xml version="1.0" encoding="UTF-8" ?>
<xml xmlns:o="urn:schemas-microsoft-com:office:office">
  <o:MainFile HRef="..tableauvraits.htm"/>
  <o:File HRef="themedata.thmx"/>
  <o:File HRef="colourschememapping.xml"/>
  <o:File HRef="image005.gif"/>
  <o:File HRef="image001.jpg"/>
  <o:File HRef="image018.jpg"/>
  <o:File HRef="image003.jpg"/>
  <o:File HRef="image019.jpg"/>
  <o:File HRef="image020.png"/>
  <o:File HRef="image021.gif"/>
  <o:File HRef="image022.gif"/>
  <o:File HRef="image007.jpg"/>
  <o:File HRef="image023.jpg"/>
  <o:File HRef="image009.jpg"/>
  <o:File HRef="image024.jpg"/>
  <o:File HRef="image011.jpg"/>
  <o:File HRef="image025.jpg"/>
  <o:File HRef="image027.gif"/>
  <o:File HRef="image028.jpg"/>
  <o:File HRef="image014.jpg"/>
  <o:File HRef="image029.jpg"/>
  <o:File HRef="image016.jpg"/>
  <o:File HRef="image030.jpg"/>
  <o:File HRef="image031.gif"/>
  <o:File HRef="image032.gif"/>
  <o:File HRef="image033.gif"/>
  <o:File HRef="filelist.xml"/>
</xml>
```

```
<?xml version="1.0" encoding="UTF-8" ?>
<xml xmlns:o="urn:schemas-microsoft-com:office:office">
  <o:MainFile HRef="..tableauvraits.htm"/>
  <o:File HRef="themedata.thmx"/>
  <o:File HRef="colourschememapping.xml"/>
  <o:File HRef="image005.gif"/>
  <o:File HRef="image001.jpg"/>
  <o:File HRef="image018.jpg"/>
  <o:File HRef="image003.jpg"/>
  <o:File HRef="image019.jpg"/>
  <o:File HRef="image020.png"/>
  <o:File HRef="image021.gif"/>
  <o:File HRef="image022.gif"/>
  <o:File HRef="image007.jpg"/>
  <o:File HRef="image023.jpg"/>
  <o:File HRef="image009.jpg"/>
  <o:File HRef="image024.jpg"/>
  <o:File HRef="image011.jpg"/>
  <o:File HRef="image025.jpg"/>
  <o:File HRef="image027.gif"/>
  <o:File HRef="image028.jpg"/>
  <o:File HRef="image014.jpg"/>
  <o:File HRef="image029.jpg"/>
  <o:File HRef="image016.jpg"/>
  <o:File HRef="image030.jpg"/>
  <o:File HRef="image031.gif"/>
  <o:File HRef="image032.gif"/>
  <o:File HRef="image033.gif"/>
  <o:File HRef="filelist.xml"/>
</xml>
```

```
<?xml version="1.0" encoding="UTF-8" ?>
<xml xmlns:o="urn:schemas-microsoft-com:office:office">
  <o:MainFile HRef="..tableauvraits.htm"/>
  <o:File HRef="themedata.thmx"/>
  <o:File HRef="colourschememapping.xml"/>
  <o:File HRef="image005.gif"/>
  <o:File HRef="image001.jpg"/>
  <o:File HRef="image018.jpg"/>
  <o:File HRef="image003.jpg"/>
  <o:File HRef="image019.jpg"/>
  <o:File HRef="image020.png"/>
  <o:File HRef="image021.gif"/>
  <o:File HRef="image022.gif"/>
  <o:File HRef="image007.jpg"/>
  <o:File HRef="image023.jpg"/>
  <o:File HRef="image009.jpg"/>
  <o:File HRef="image024.jpg"/>
  <o:File HRef="image011.jpg"/>
  <o:File HRef="image025.jpg"/>
  <o:File HRef="image027.gif"/>
  <o:File HRef="image028.jpg"/>
  <o:File HRef="image014.jpg"/>
  <o:File HRef="image029.jpg"/>
  <o:File HRef="image016.jpg"/>
  <o:File HRef="image030.jpg"/>
  <o:File HRef="image031.gif"/>
  <o:File HRef="image032.gif"/>
  <o:File HRef="image033.gif"/>
  <o:File HRef="filelist.xml"/>
</xml>
```

```
<?xml version="1.0" encoding="UTF-8" ?>
<xml xmlns:o="urn:schemas-microsoft-com:office:office">
  <o:MainFile HRef="..tableauvraits.htm"/>
  <o:File HRef="themedata.thmx"/>
  <o:File HRef="colourschememapping.xml"/>
  <o:File HRef="image005.gif"/>
  <o:File HRef="image001.jpg"/>
  <o:File HRef="image018.jpg"/>
  <o:File HRef="image003.jpg"/>
  <o:File HRef="image019.jpg"/>
  <o:File HRef="image020.png"/>
  <o:File HRef="image021.gif"/>
  <o:File HRef="image022.gif"/>
  <o:File HRef="image007.jpg"/>
  <o:File HRef="image023.jpg"/>
  <o:File HRef="image009.jpg"/>
  <o:File HRef="image024.jpg"/>
  <o:File HRef="image011.jpg"/>
  <o:File HRef="image025.jpg"/>
  <o:File HRef="image027.gif"/>
  <o:File HRef="image028.jpg"/>
  <o:File HRef="image014.jpg"/>
  <o:File HRef="image029.jpg"/>
  <o:File HRef="image016.jpg"/>
  <o:File HRef="image030.jpg"/>
  <o:File HRef="image031.gif"/>
  <o:File HRef="image032.gif"/>
  <o:File HRef="image033.gif"/>
  <o:File HRef="filelist.xml"/>
</xml>
```

## TABLE DES MATIERES

Remerciements .....	1
Sommaire .....	2
Introduction .....	3
ipernity .....	4
ipernity, aujourd’hui .....	4
L’équipe.....	4
Le service proposé .....	4
Le modèle de revenus.....	5
Quelques chiffres .....	6
ipernity, de la première idée à nos jours .....	6
Historique.....	6
Les problèmes rencontrés .....	8
ipernity dans une boule de cristal.....	8
Les objectifs .....	8
Les projets futurs .....	8
ipernity, dans la jungle du web 2.0.....	9
ipernity, vue de l’intérieur .....	10
Le travail proposé .....	11
Le sujet.....	11
Quelques précisions.....	11
Le cahier des charges .....	12
Contraintes techniques.....	12
Contraintes utilisateurs.....	13
Le travail réalisé .....	14
Caractéristiques générales de la plateforme .....	14
Déroulement du stage .....	14
Travaux annexes .....	15
Connaissances utilisées et acquises.....	16
Relations dans l’entreprise .....	17
Difficultés rencontrées.....	17

Conclusion.....	19
Bibliographie .....	20
Annexes.....	21
Réponses au questionnaire.....	21
Christian Conti.....	21
Léa Cymbalista .....	22
T. M. (maître de stage).....	23
Christophe Ruelle.....	24
Exemple de miniatures .....	25
Table des matières.....	26