



Création d'un site institutionnel web 3.0 : de la communication à l'information

[outil] Le projet Spectacle vivant en Bretagne a été l'occasion d'imaginer un site web « outil » de communication qui soit aussi un site ressource. Trois axes structurent ce projet : développer un outil numérique offrant des services aux usagers, s'appuyer sur des technologies et des informations existantes, inscrire le site dans l'ère du web 3.0, plus particulièrement du web sémantique et de la datavisualisation.

Spectacle vivant en Bretagne est un établissement public de coopération culturelle (EPCC). Ses deux fondateurs et contributeurs sont la Région Bretagne et le ministère de la Culture et de la Communication/Drac Bretagne. Son cœur de mission est d'accompagner les stratégies de développement et de diffusion, en France et à l'étranger, des équipes artistiques professionnelles implantées en Bretagne. Ses domaines d'intervention couvrent toutes esthétiques de musique, danse, théâtre, arts du cirque pour la scène, l'espace public, la piste. Ses axes de travail se déploient autour du développement de la diffusion, de la coopération et de la prospective. Ses outils sont le conseil sur mesure, l'aide financière à la diffusion et à la mobilité, la mise en visibilité du travail artistique, la mise en lien.

Le besoin : une base de connaissance professionnelle

L'ambition fondatrice du nouveau site web de Spectacle vivant en Bretagne¹ était de capitaliser les informations relatives à l'activité de l'établissement afin de constituer une base de connaissances de la diffusion artistique et de ses réseaux, en France et à l'échelle internationale.

Le projet devait répondre à plusieurs problématiques.

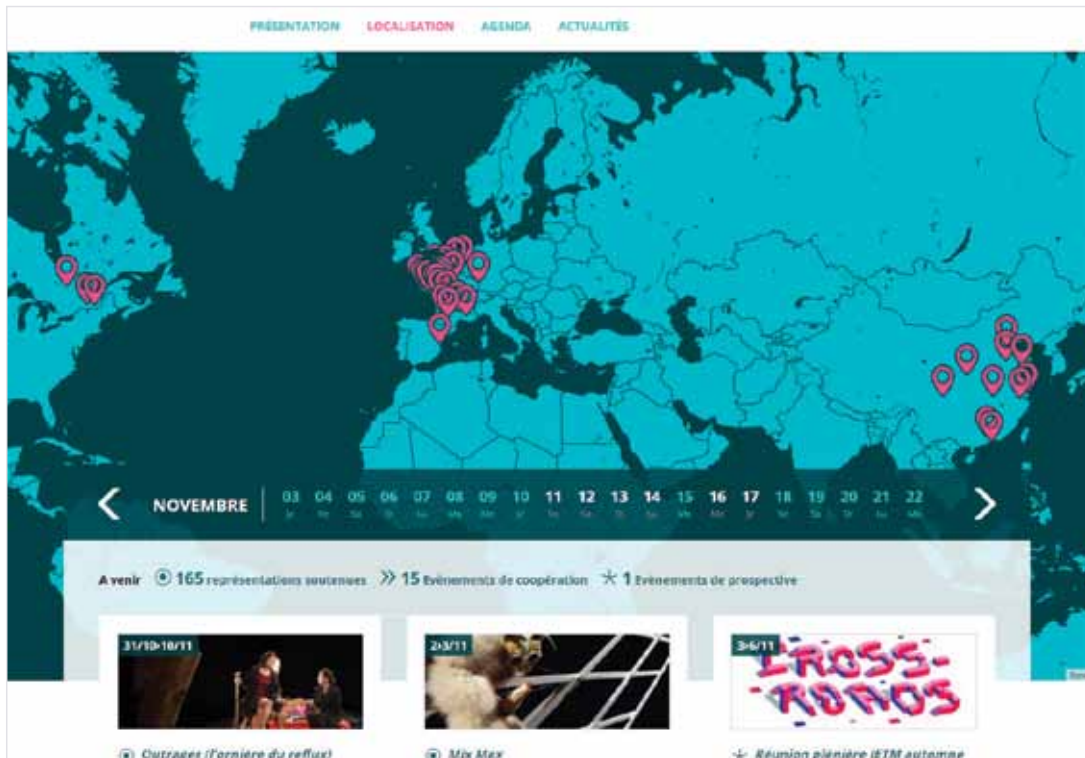
Sur le plan de la communication :

- mobiliser les professionnel(le)s sur les dates de spectacle aidées financièrement par Spectacle vivant en Bretagne et sur les actions de l'établissement ;
- rendre lisibles et visibles ces mêmes dates et actions et leur amplitude géographique ;

1. www.spectacle-vivant-bretagne.fr

Figure 1

Page d'accueil du site avec visualisation géographique et temporelle générée en temps réel. Les chiffres en vert indiquent qu'un événement est lié à cette date.



- favoriser l'interaction et la mise en réseau des professionnels.

Sur le plan de la ressource :

- valoriser les données associées à ces événements (structure de programmation, lieu de diffusion, spectacle, équipe artistique, garantie financière accordée) ;
- conserver l'historique de ces actions et constituer une banque d'expériences publique sur la diffusion ;
- analyser ces données d'activité à des fins d'adaptation stratégique et de prospective ; disposer d'un rapport d'activité en temps réel.

Sur le plan de la gestion de l'information :

- optimiser sa saisie, sa mise à jour et sa diffusion ;
- agréger des données de nature différente : financières, administratives, fichier de contacts.

Afin de permettre l'optimisation et l'exploitation des informations saisies relatives à l'activité de la

structure, le projet a été orienté « web de données ». La première étape a consisté à identifier les types de contenus du site, conformément à la ligne éditoriale, puis à analyser et décomposer leur structure (nature et format des données) : titre/nom, personne morale ou physique liée, notion de date, données chiffrées, données textuelles, liens (URL), données de géolocalisation (adresse), etc. pour déterminer *in fine* les liens entre les entités.

L'addition des technologies pour donner du sens à l'information

Les technologies et les standards du Web évoluant très vite, la philosophie du projet a privilégié l'usage, voire le détournement, d'outils existants et reconnus. Le CMS WordPress proposé par

Artwai², l'agence retenue pour le projet, entrain dans ces critères. Toutefois, la cartographie de l'information, le cycle de vie des ressources ainsi que les fonctions de requêtes, spécifiques à ces environnements technologiques, ajoutaient de forts besoins de performance et d'optimisation des temps de réponse pour garantir un confort de navigation.

Le système de gestion de base de données finalement adopté est composé de Docalist, module spécifique intégré à WordPress (voir encadré). Il permet de structurer et lier finement les données et s'appuie sur la technologie ElasticSearch (moteur de recherche *open source*) pour la partie *datamining*. Initialement développé pour des besoins documentaires, l'usage de Docalist a donc été élargi et adapté au projet et aux typologies d'informations de Spectacle vivant en Bretagne. // //

Figure 2

Tableau de correspondance à trois entrées. Extrait d'une fiche de type « structure »

Champs type "organisation" de schema.org			Champs entité "l'organisme" de la base type Le3vivre en français			Champs Docalist (*) - (médiatisé)		
nom	Type	Description	nom	Unité	Description	nom	Type	Description
name	Text	The name of the item. On a aussi : legalName: The official name of the organization, e.g. the registered company name.	nom (id ?)	1	texte	titre	Text	Table d'autorité associée à Type à 1-axe
description	Text	A short description of the item.	champ notes/description?	n	texte	contexte	(Type: Text?)	Table d'autorité associée à Type à 1-axe
logo	URL or ImageObject	An associated logo.	logo	1	image	image_fonctionnel	image	image à la une standard de WP
member employee founder	Person	A member of an Organization, someone working for this organization. A person who founded this organization.	PP "liés"	n	table "lg" (cf. pour voir gérer des fonctions comme pour le champ "auteur" de une liste documentaire (auteur principal, secondaire, prof., coord., etc.))	auteur	nom, prénom, «de»?	nom nommé (person ?) Table d'autorité associée : main relations, voir à de + gérer relations
member	Organization	A member of an Organization.	car asso ou affiliation	n	table "organisation" (destiné pour une ou plusieurs organisations, structure membre d'un réseau, états instance (N°))	organisation	nom, sigle, ville, pays, «de»?	Table d'autorité associée : main relations, voir à de + gérer relations
address	PostalAddress	Physical address of the item.	adresse + région d'impression	n	adresse postale (spécification) (1 ?)	adresse (usage à créer)	(Type: PostalAddress?)	Type d'adresse (ville), adresse, CP, CP, ville, région, pays, indice ?
telephone faxNumber contactPoint	Text Text ContactPoint	The telephone number. The fax number. A contact point for an organization.	contonées tel	n	texte	class adresse ? champ ? à part ? (ville, st, pr, etc) dans number ?	(Type: Text, tel?)	tel au format international +33-200
?		Il y a quelques champs numéros connus dans, tel?4, mais nous avons rien utilisable pour nous. On a aussi fax(D), url(D), globalLocationNumber	SRET + APE	1	1	number	Type: value?	Table d'autorité associée "Types de numéros" à créer : sret, sret, app, et tra. ...
url	URL	URL of the item. On a aussi : mainEntityOfPage	site web	n	url	link	Type: url, label, date?	Table d'autorité associée "Types de liens" à compléter. Pour chaque lien possibilité d'indiquer un libellé et date de consultation du lien.
startDate foundingDate	Date Date	The date that this organization was founded.	date de création / d'activité	1	date	date	Type: date?	Table d'autorité associée "Types de dates" à compléter

//// Visibilité, accessibilité et pérennité des données

Pour répondre aux enjeux d'interopérabilité des données et faciliter le travail de structuration, Schema.

2. www.artwai.com

3. Initiative lancée conjointement par Bing, Google et Yahoo! en 2011, Schema.org est un schéma de micro-données utilisé sur le Web. Source : Wikipédia, <https://fr.wikipedia.org/wiki/Schema.org>

org³ a été choisi comme modèle pour établir la liste des propriétés des différentes entités de la base de données. Non seulement le vocabulaire SchemaOrg, initiative commune de quatre géants de la recherche d'information sur Internet, est devenu une référence internationale, mais il améliore le référencement et répond aux enjeux du SEO (Search Engine Optimization). À terme, cela permettra aussi d'enrichir les données du site d'un balisage sémantique.

Un rapprochement et une table de correspondance ont été établis entre des types de contenus et leurs propriétés (la liste structurée est disponible sur le site dédié schema.org), puis entre les entités

(familles d'informations) de la base de données du site (figure 2). Ce travail a permis de créer le modèle conceptuel de données (MCD), phase de modélisation (ou maquettage) de la base de données, avant de la développer.

Intégration, saisie, recherche et exploitation des contenus

La base de contacts et une grande partie des données d'activité de l'établissement sont gérées dans une base de données interne. Ces informations ont été importées dans la base de connaissances et sont ainsi mises à disposition sur le site. Elles sont enrichies, par exemple d'images ou de liens hypertextes, via des formulaires spécifiques à chaque entité.

La dimension stratégique de l'outil se situe dans sa capacité à croiser l'information, à géolocaliser et agréger des données chiffrées dont la mise à jour est dynamique. Il permet d'accéder aux contenus

Docalist, module intégré à WordPress

« Docalist apporte à WordPress une autre manière de publier de l'information : là où WordPress permet essentiellement de publier des articles (du texte non structuré dans lequel on met ce qu'on veut), Docalist impose dès le début des masques très structurés permettant de typer l'information, de la rendre répétable et de l'associer à des listes de termes prédéfinies (faisant office de tables d'autorité comme un thésaurus, une ontologie, etc.). Pour cela, Docalist se base sur des entités prédéfinies (des « grilles ») qui modélisent les objets (réels ou virtuels) que nous manipulons et qui ont été élaborées en tenant compte de plusieurs standards existants (schema.org, [marc](http://marc.unimarc.org), [unimarc](http://unimarc.org), [onix](http://onix.org) et d'autres). »

Daniel Ménard, développeur de Docalist

du site par les entrées géographique, temporelle (figure 1), par corpus (figure 3) ou par entité de la base de données, avec génération dynamique de chiffres clés liés à un résultat de recherche (figure 4). La liste peut ensuite être affinée par facettes.

De l'information à la connaissance

Le projet a débuté en septembre 2015 et une première version du site a été mise en ligne fin avril 2016. Si le versant « fouille de données » est bien développé et ouvre des perspectives de data-visualisation poussées (extension de l'affichage géolocalisé ou temporel aux pages de résultats entre autres) – un site stabilisé pour début 2017 est envisageable –, des fonctionnalités orientées réseau social et mise en relation devraient être développées dans un second temps. Mais tous les contenus du site sont désormais capitalisés et exploités. Les données saisies ou intégrées sont valorisées : du statut d'« actualité », une information structurée devient ressource.

Le web de données à l'échelle d'un organisme permet ainsi, dans cet exemple, d'étendre le concept d'économie circulaire au domaine de la gestion de l'information. ■

> Flora Even

Responsable Information
Communication. E ressources
documentaires
Spectacle vivant en Bretagne
f.even@svbretagne.fr

Figure 3

Contenus de la base de connaissance avec données chiffrées générées en temps réel

Source : www.spectacle-vivant-bretagne.fr rubrique « écosystème »

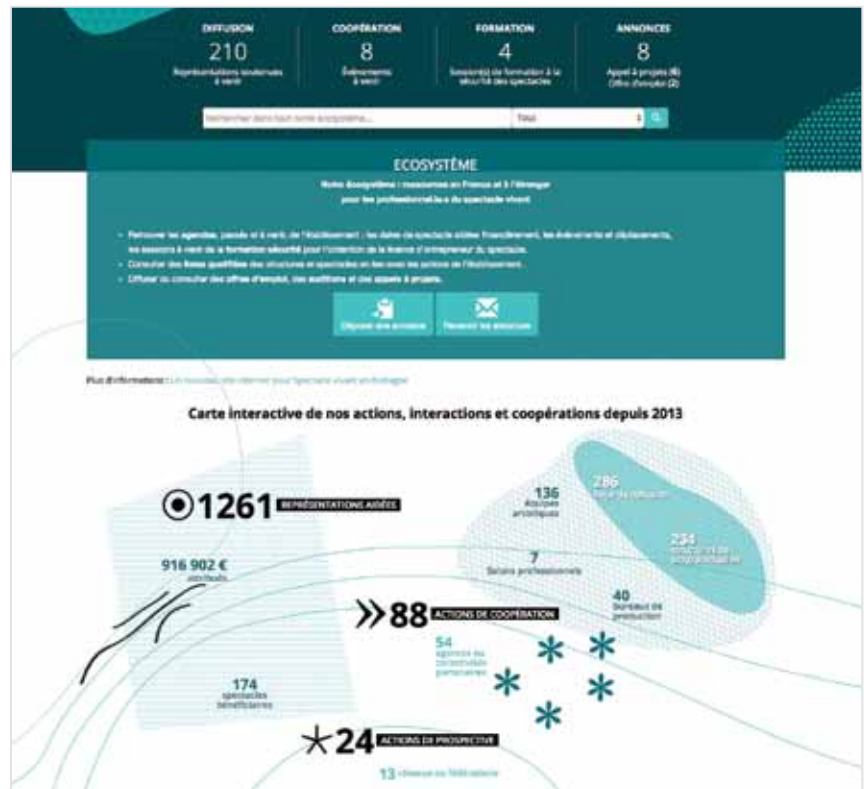


Figure 4

Liste de résultats sur l'entité « date soutenue » de la base, avec options de tri, facettes pour affiner la requête et chiffres clés dynamiques liés au résultat

