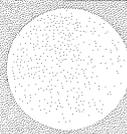


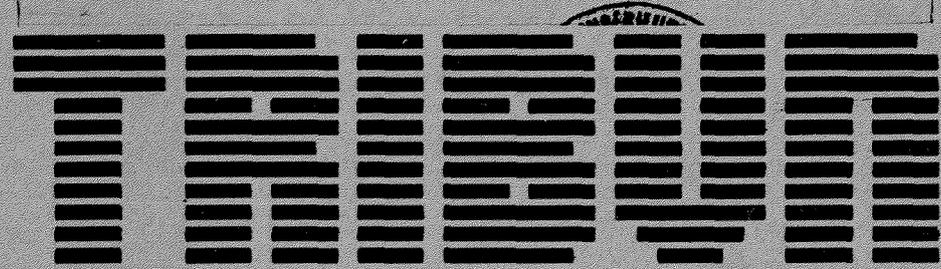
0879

1988
Stages
6



CROZET Marc

CREATION D'UNE BASE DE DONNEES
DOCUMENTAIRE ASSOCIEE A UNE
PHOTO THEQUE "ON LINE"



RAPPORT DE STAGE

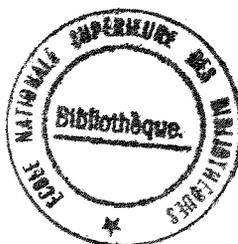
D.E.S.S. d'informatique documentaire
Université LYON 1 - Claude Bernard

1988

1988
Stages
6

CROZET Marc

CREATION D'UNE BASE DE DONNEES
DOCUMENTAIRE ASSOCIEE A UNE
PHOTOTHEQUE "ON LINE"



<u>RAPPORT DE STAGE</u>
D.E.S.S. d'informatique documentaire
Université LYON 1 - Claude Bernard
1988

1988
Stages
f

SOMMAIRE

AVERTISSEMENT.....p.2

LE CADRE DU STAGE.....p.3

- Présentation de TRIBVN p.4
- Présentation du Messager d'images p.7
- Présentation de KIPA.....p.9

LE CONTENU DU STAGE p.10

- Missions p.11
- Analyse de l'existant.... p.12
- Configuration de la base de données.. .p.16
- Intégration de la base de données.... .p.22

CONCLUSION p.25

ANNEXES.....p.27

- Synoptique du serveur KIPA. p.28
- Agence KIPA : composition du fonds.. ..p.29
- Codes pour le champ "Nature du reportage..p.30
- Codes pour le champ "Descripteur".....P.31

AVERTISSEMENT

Le stage s'est déroulé du 06 juin au 30 septembre 1988, auprès de la société TRIBVN, 102 avenue du Maine à Paris 14ème, sous la responsabilité de M. Klossa, son directeur général.

L'essentiel du travail a toutefois porté sur une réalisation à l'agence de photographies de presse KIPA, 12 rue Martel à Paris 10ème, dans le cadre d'un projet de "Messager d'images" dont TRIBVN assure la maîtrise d'oeuvre. Ce projet étant en cours de réalisation, un certain nombre des développements exposés supra n'ont pas encore, à ce jour, eu de traduction concrète.

La bonne marche du stage n'a été possible que grâce au soutien de Jacques Klossa et d'Annemiek Veldman, directrice de l'agence Kipa; que soient également remerciés pour leurs conseils André Agabalyan, d'Electronic research et Ursule de la Vaissière, d'ID France.

LE CADRE DU STAGE

PRESENTATION DE TRIBVN

Créée en 1985 par deux ingénieurs issus de la D.G.T., la société TRIBVN (avec un V comme vidéo !) est une SSCI à caractère industriel positionnée sur l'image. Un positionnement reconnu performant et sanctionné par de nombreuses distinctions : Mercure des banques d'images (Musée de la mode - IMACOM 1987), Champion de l'innovation (FIEE-COOPEREL 1987), Prix national des banques populaires ...

I - SSCI

Ses prestations couvrent la quasi-totalité des besoins en terme de banques d'images :

- maîtrise d'oeuvre d'applications clés en main : RATP, musée de la mode...
- développement de logiciels, applications interactives.
- prestations de transfert d'images sur supports optiques (ou autres supports)
- conseil et assistance technique.

II - PRODUCTION INDUSTRIELLE

Produits élémentaires :

- carte imageur "HD-RAM", une carte mémoire d'image PC/AT qui, associée aux logiciels documentaires Turbo thesaurus, Texto ou BRS, transforme un PC en un puissant outil de recherche visuelle (comparaison et manipulation d'images de référence en sus de la pleine image couleur issue d'un vidéolecteur).
- carte d'incrustation "CD-RAM" : le graphisme en incrustation pour le sous-titrage des banques d'images.
- carte imageur "BVN 4.2.2", une carte mémoire 16 millions de couleurs simultanées, mode 4.2.2, pleine définition broadcast, option numérisation temps réel.

Produits intégrés :

- gamme de stations Imageur documentaire à vidéodisques (l'imageur permet de visionner par planches de 4, 9 ou 16, les images issues d'un vidéodisque, et numérisées en temps réel).
- stations locales de gestions d'images "fraiches" (vidéolecteurs numériques).

TRIBVN s'est associée en 1988 avec Lyonnaise Communication, filiale de Lyonnaise des eaux spécialisée dans les réseaux cablés, pour créer la société IMAGE DIRECTE, chargée de développer et de commercialiser des services à valeur ajoutée utilisant des banques d'images et notamment un service de diffusion, en direction des journaux et des magazines, de photographies de presse numérisées : le messageur d'images.

III - PRODUCTION VIDEO

TRIBVN est aussi une société de création audiovisuelle (films d'entreprise, génériques, dessins animés ...) utilisant les techniques les plus élaborées en la matière (image de synthèse, incrustation, palettes graphiques, image 3D ...). Sa prestation sur Antenne 2 lors des soirées électorales de mai et juin 1988 a été particulièrement remarquée.

TRIBVN : CARTE D'IDENTITE

STATUTS

S.A. au capital de 700 000 francs, en cours d'augmentation à 3,5 MF .

Code APE : 7703 Cabinets d'études en informatique

CHIFFRES D'AFFAIRES

Exercice 85-86 : 5 MF (recherche pure)
Exercice 86-87 : 5 MF (début d'industrialisation)
Exercice 87-88 : 9,5 MF (intégration d'un service commercial)
Prévision 88-89 : 18 MF

LISTE DU PERSONNEL : 20 personnes

- 3 ingénieurs - chefs de projets, origine Télécommunication
- 2 ingénieurs experts systèmes vidéo
- 1 ingénieur confirmé
- 4 techniciens
- 1 cableur
- 1 designer, directeur artistique
- 3 graphistes
- 1 responsable de gestion
- 1 responsable des relations publiques
- 2 attachés commerciaux
- 1 secrétaire

Cette équipe agit souvent en maître d'oeuvres et est amenée à compléter ses forces à l'aide de deux PMI sous-traitantes travaillant essentiellement pour TRIBVN : Electronic research, Siros France.

PRESENTATION DU MESSAGEUR D'IMAGES

I - UN SERVICE NOUVEAU

Le développement par FRANCE TELECOM du Réseau numérique à intégration de services (RNIS) crée les conditions pour la mise en oeuvre d'applications novatrices, notamment dans le domaine de la diffusion des images fixes. La vocation du Messageur d'images est de proposer l'accès "on line", via le RNIS, aux grandes banques d'images, leur permettant de trouver ainsi de nouveaux débouchés commerciaux.

II - L'OPERATION PILOTE

La première application du Messageur porte sur les fonds de l'agence KIPA, sur la base d'un partenariat à trois : IMAGE DIRECTE (= TRIBVN + Lyonnaise Communication), FRANCE TELECOM ET KIPA.

Un serveur d'images numérisées sera mis en place chez KIPA et raccordé au RNIS; des stations de consultation seront mises à la disposition de 5 clients pilotes : France-Soir TV magazine, VSD, Le Pèlerin, Burda, et un des journaux du groupe EDI 7. L'utilisateur final pourra consulter les images sur son terminal, d'abord par planches de 9 imasettes puis, après une présélection, plein écran. Dans un premier temps, seules les photos fraîches (photos du jour) seront télédéchargées chez le client; pendant cette période débutera la phase de digitalisation des photos d'archives, au rythme de 1000 à 2000 par jour. A moyen terme, le client pourra interroger le fonds d'archives, par l'intermédiaire des documentalistes de KIPA qui chargeront sur son terminal une sélection de photos répondant à sa demande.

TRIBVN assure la maîtrise d'oeuvre de la mise en place de l'application, qui s'appuie sur les unités suivantes, reliées entre elles par un réseau Ethernet (voir le schéma en annexe):

- banque d'images
- base de données textuelles
- unité de gestion
- unité d'étiquetage
- poste de numérisation / compression
- frontal de télécommunication, raccordé au RNIS
- terminaux locaux : texte seul, texte et image

Le fichier numérique obtenu après digitalisation servira à produire 3 éléments :

- une archive de l'image pleine taille 0,8 Mo sur streamer ou disque optique
- une image réduite au 1/9 et éventuellement redressée si le document est vertical, puis compressée
- une image pleine taille compressée

Les stations de consultation à distance auront la possibilité de conserver sur leur disque environ 6000 images; elles seront dotées de fonctions de manipulation des panneaux d'images, la demande d'images plein écran ou de documentation textuelle supplémentaire se faisant en léger différé. Une imprimante couleur sera éventuellement proposée, quoique les performances actuelles ne permettent une qualité optimale que sur de petits formats. L'hypothèse retenue est que, une fois leur sélection achevée, les journaux continueront à exiger la communication de l'image sur un support chimique .

III - PERSPECTIVES

L'application devrait rapidement s'étendre à d'autres agences de presse (Gamma dès mars 1989), puis à des fonds culturels, au cinéma (montage de castings), à la cartographie... L'utilisateur abonné au Messageur recevrait bien sûr toutes ces images sur le même terminal.

Si la vocation d'IMAGE DIRECTE est d'assurer la gestion et l'extension du service Messageur, TRIBVN, pour sa part, songe à d'autres développements possibles : édition électronique, montage de produits vidéo à partir d'images fixes ,etc...

PRESENTATION DE KIPA

Avec un effectif de 31 personnes, KIPA se place au quatrième rang des agences de photographies de presse en France. En 1987, le chiffre d'affaires de 11 MF a été réalisé pour 75 % en France (60 % à Paris, 40 % en province).

Les principaux secteurs d'activité de KIPA se situent dans le domaine de la télévision (60 % de la production de l'agence) et du spectacle.

Ses principaux clients sont au nombre de 80 et se situent parmi les quotidiens, les magazines spécialisés (cinéma, musique), les journaux de télévision. De plus, KIPA est présente par l'intermédiaire de ses revendeurs ou de ses correspondants dans 12 pays européens, aux Etats-Unis et au Japon.

En plus de la production journalière (environ 200 photos), KIPA dispose d'un fonds de plus de 3 millions de photos. La diffusion auprès des clients représente 14 000 photos par mois, dont 2400 seront publiées (et donc payées). Elle s'articule autour de deux principes : la duplication des photos et le déplacement des vendeurs auprès des clients pour présenter la production du jour.

La mise en exploitation du Messageur permettra à l'agence d'alimenter plus rapidement ses clients (notamment en province), de proposer un choix plus riche et mieux ciblé, enfin de faire diminuer sensiblement le nombre de duplicatas, puisque le client ne demandera que des photos qu'il aura déjà visualisées et choisies.

LE CONTENU DU STAGE

MISSIONS

L'objet du stage était de trouver les solutions adéquates pour gérer les informations textuelles associées aux photos dans la future banque d'images KIPA, en tenant compte de deux impératifs :

- Le système devait être opérationnel le 13 septembre, jour de l'inauguration officielle du RNIS (cette date a finalement été repoussée par FRANCE TELECOM).
- Le rythme à tenir était de 2000 images à intégrer par jour, afin d'atteindre 1 million en 2 ans.

Le travail a alors suivi plusieurs phases :

- Etude de l'existant, définition des objectifs.
- Rédaction d'un cahier des charges documentaire, choix du logiciel, suivi de la réalisation de la base de données.
- Etude des problèmes de manipulation des photos lors de la saisie / numérisation. Spécification, avec les autres intervenants du projet, de la chaîne de traitement.
- Participation à la définition du dialogue utilisateur, pour le terminal texte / image.

ANALYSE DE L'EXISTANT

I - CLASSEMENT - NUMEROTATION - FICHIERS

KIPA ne disposait jusqu'à présent d'aucun outil informatique pour gérer les quelque 3 millions d'images de son fonds. Le système de classement actuel épouse l'historique, la provenance et la nature des photos, c'est à dire qu'il est extrêmement disparate. On peut distinguer une quinzaine de divisions différentes, chacune ayant son classement propre, voire plusieurs, suivant le support (diapositives, planches contact, négatifs, tirages ...); tous les grands types de classement sont représentés : alphabétique, séquentiel, systématique, thématique, dictionnaire ...

Traversant ces divisions, mais ne les recouvrant pas, on compte 11 numérotations, certains fonds anciens n'étant par ailleurs pas numérotés ! Les numéros utilisés qualifient non pas une photo, mais plutôt un groupe de photos, le plus souvent un reportage, mais parfois une planche ou encore un regroupement de reportages.

La gestion et la recherche des photos se fait à l'aide de cahiers "chrono", décrivant sommairement les reportages par ordre d'arrivée, et de fichiers manuels, inégalement tenus et donc peu performants (à quelques exceptions près). Retrouver des photos d'archives est donc avant tout question de mémoire et de bonne connaissance du fonds; cette dernière qualité est, il faut le dire, remarquable chez le personnel de KIPA.

En ce qui concerne les fonds vivants (les fonds morts ne seront intégrés à la base qu'en dernier lieu), on peut distinguer 4 catégories :

a) *Les reportages ordinaires*

Ils constituent la part la plus importante de la production : émissions de télévision, tournages de film, événements mondains, photos de studio, pièces de théâtre, etc ... D'une façon générale, on peut définir un reportage comme la résultante de trois facteurs : un photographe + un lieu + une date.

Les reportages sont numérotés par ordre d'arrivée puis rangés, suivant leur sujet ou leur provenance, dans 7 divisions différentes (un même reportage pouvant être éclaté dans plusieurs divisions) :

- *ALPHA (portraits de personnalités). Classement alphabétique
- *CLASSEURS (reportage divers). Classement séquentiel
- *VIE MONDAINE . Classement alphabétique titre
- *MODE . Classement alphabétique titre
- *ILLUSTRATION GENERALE . Classement systématique
- *L.G.I. (du nom d'une agence américaine) . Classement alphabétique titre
- *ALBANE NAVIZET (photographe indépendante) . Classement alphabétique

Le fichier correspondant à ces reportages contient essentiellement, en une seule séquence alphabétique, des fiches "titre" (mais tous les reportages n'ont pas de titre...) et des fiches "personnalités".

Les diapositives sont rangées par panodias de 20, numérotés de 1 à n au sein d'une unité de classement. Les diapos de très bonne qualité reçoivent un point rouge et sont rangées devant.

b) *Les photos de film*

Il s'agit des photos de production (4 à 10 par film), que KIPA récupère auprès des distributeurs; elles sont reproduites d'une part sur négatifs noir et blanc, d'autre part sur diapos couleur.

Chaque film reçoit un numéro de négatif et, s'il est en couleur, un numéro de planche couleur (numéros tous deux chronologiques mais distincts). Le rangement de ces deux supports est séquentiel, un fichier alphabétique titre permettant de retrouver les numéros d'un film donné.

KIPA faisant une part importante de son chiffre d'affaires avec l'illustration des programmes TV, ce secteur est important (20 000 films actuellement).

c) *L'opéra*

Ce secteur, ainsi que le suivant, possède une organisation très spécifique, avec un système documentaire précis, quoique compliqué. On peut presque parler d'agence dans l'agence.

Chaque production lyrique y fait l'objet d'un grand nombre de fiches : titre, compositeur, metteur en scène, chorégraphe, chef d'orchestre, lieu, livret, décors, costumes, chanteurs, etc... Il est à noter que ces notices concernent l'objet représenté plutôt que le contenu réel des photos.

Trois numérotations entrent en jeu :

- Nos des planches noir et blanc
- Nos des planches couleurs
- Nos de dossiers (un dossier regroupant plusieurs événements)

d) *La danse*

L'organisation est absolument identique à celle qui précède, mais les séries de numéros sont propres à la danse.

II - CRITERES DE RECHERCHES

KIPA est essentiellement une agence de photos de personnalités. Le critère de recherche le plus courant est donc le nom de personne; l'organisation actuelle de l'agence permet de répondre, plus ou moins bien, à une requête portant sur un nom ou sur un titre (d'émission, de film, de pièce...).

En revanche, une requête telle que, par exemple, "des chanteurs de variétés avec leur chien", sera, faute de moyens documentaires, traitée de manière totalement empirique, en passant en revue un grand nombre de photos et en faisant appel à la mémoire des documentalistes.

Les critères de recherche non anthroponymiques, quoique très divers, peuvent néanmoins être regroupés en quelques catégories identifiables, le même type de questions revenant en effet régulièrement. Ils concernent :

- l'approche du sujet
- le look (détails vestimentaires, allure générale)
- les accessoires (cigare, téléphone, piano ...)
- le type de lieu (rue, jardin, intérieur ...)
- le genre, pour les films
- l'action (rire, baiser ...)

III - LA VENTE DES PHOTOS FRAICHES

En ce qui concerne les photos du jour, KIPA démarché directement ses clients, sans attendre d'être sollicitée. Plus qu'une photo, c'est alors tout un sujet que l'on vend, en suggérant au journal d'écrire un article que la ou les photos illustreraient. 3 ou 4 sujets sont choisis quotidiennement parmi la production du jour et font l'objet d'un "écho" : les meilleures prises sont dupliquées en plusieurs exemplaires et sont emportées, avec un court texte d'accompagnement, par les vendeurs qui se chargent de les caser auprès des journaux.

Les vendeurs négocient également avec les journaux des contrats d'exclusivité, ou des garanties de diffusion pour des reportages commandés par un client. Cette "force de frappe" commerciale que représentent les vendeurs ne peut être négligée : il ne suffit pas de présenter des photos sur un terminal, encore faut-il convaincre le client de leur intérêt !

Bien que l'intégralité du fonds soit régulièrement sollicité, KIPA est plus un agence d'actualité qu'une agence d'archives; ce qui explique peut-être l'absence de système documentaire vraiment efficace.

CONFIGURATION DE LA BASE DE DONNEES

I . LES PRINCIPES

a) *une indexation à 2 niveaux*

Une image n'est jamais seule; elle est insérée dans un réseau de rapports qui permet d'organiser des regroupements. La base de données doit faciliter ces regroupements, tant pour la saisie des descripteurs, que dans sa structuration même. Pour KIPA, deux unités documentaires ont été définies :

* Une unité de premier niveau, constituée par un lot homogène de photos; l'organisation actuelle des fonds, quoique peu satisfaisante, a servi de base à la définition de ce lot (car il n'était pas question de reclasser 3 millions de photos!); on a tenu compte des numérotations existantes et de l'organisation physique des photos. Un lot sera donc défini par le croisement de deux données indépendantes :

- un numéro d'ordre dans une série (4 séries : reportages ordinaires, films, opéra, danse)
- un code de classement correspondant au rangement des photos en rayon (11 possibilités)

* Une unité de deuxième niveau, constituée par une photo. Il est en effet fondamental de pouvoir identifier sans équivoque une photo donnée (ce qui n'est pas le cas aujourd'hui).

b) *une double numérotation*

Chaque photo est identifiée dans la base par 2 numéros :

* Un numéro séquentiel absolu, chronologique, délivrée par la base de données; ce numéro permettra d'adresser l'image, aussi bien dans la banque d'images que dans la base de données.

* Un numéro logique complexe, composé des éléments suivants

- Code de classement (2 lettres)
- Numéro de série (5 chiffres)
- Numéro d'ordre de la photo dans le lot (3 chiffres)

par exemple : AL-15078-018 : 18ème photo du lot portant le numéro de reportage 15078 et classé dans le rayon ALPHA.

Pour certaines séries, qui disposent aujourd'hui de plusieurs numérotations suivant le support d'archivage (même s'il s'agit de la même image), on a choisi le numéro le moins ambigu : numéro de négatif pour les films, numéro de dossier "redivisé" pour la danse et l'opéra.

Le numéro séquentiel de photo n'apparaît que sous la forme d'un code à barre, le numéro logique est inscrit en clair sur la cache de la photo, afin d'aider au rangement.

c) une indexation légère

Dans les grands fonds d'images de presse visités (EDI 7, Gamma, BBC), l'indexation est en général très approfondie, ce qui s'explique facilement lorsque la recherche, que ce soit en texte intégral ou sur thesaurus, se fait entièrement sur des données textuelles. Or le fait de disposer de l'image en ligne, de plus par panneaux de 9 aisément manipulables, permet d'alléger considérablement le texte associé. La sélection visuelle vient compléter une première sélection textuelle dont le résultat peut être assez large (l'oeil peut balayer plus de 100 images en une minute).

On a donc choisi pour KIPA d'axer l'analyse sur l'identification de l'image, plus que sur sa description ; les éléments purement visuels (cadrage, exposition) ont été écartés et le nombre de mots-clés a été limité. La politique d'indexation doit être pensée uniquement en fonction de la recherche : la question que devront se poser les documentalistes n'est pas "Que représente cette image ?" mais "Comment va-t-on accéder à cette image?".

On trouvera en annexe une première liste de mots-clés correspondants à des requêtes très fréquentes chez KIPA.

d) un vocabulaire contrôlé

Comme nous l'avons vu, le critère déterminant pour caractériser une photo chez KIPA est la ou les personnalité(s) représentée(s) ou intervenant, en tant que créateur(s), sur ce qui est photographié (film ou spectacle).

On trouvera donc des noms propres dans un grand nombre de champs ; chacun de ces champs sera indexé séparément, mais afin de les contrôler et d'éviter les erreurs de graphie, on disposera d'un dictionnaire reprenant en une séquence alphabétique unique tous les noms de personnes utilisés dans la base.

La formule du thesaurus "a continuo" a été choisi pour les descripteurs. Il eut été couteux, long et inutile d'établir pour KIPA un thesaurus a priori, compte tenu du nombre limité de descripteurs envisagé. Les documentalistes disposeront donc dans un premier temps d'un simple dictionnaire des termes déjà utilisés, à charge pour elles d'établir, petit à petit, des relations entre ces termes.

II - LES DIFFERENTS CHAMPS

** Identification d'un lot*

- Code de classement
- Numéro de série
- Autres numéros
- Date
- Photographe
- Titre
- Nature du reportage
- Lieu
- Chaîne (pour les émissions de télévision)
- Production (pour les séries, téléfilms, feuilletons)
- Commentaires
- Nombre de photos

** Informations complémentaires sur un lot (optionnelles)*

- Réalisateur
- Metteur en scène (sauf film)
- Compositeur
- Chef d'orchestre
- Chorégraphe
- Livret
- Musiciens
- Compagnie
- Décors
- Costumes
- Comédiens
- Chanteurs
- Danseurs
- Styliste

** Identification d'une photo*

- Numéro séquentiel
- Numéro d'ordre dans le lot
- Support
- Format
- Point rouge
- Confidentialité
- Sujet (personnes présentes sur cette photo)
- Descripteur

D'autres champs seront éventuellement rajoutés au niveau "photo", lorsqu'il sera possible de les gérer automatiquement :

- Nombre de demandes "plein écran" sur le Messager
- Nombre de duplicatas disponibles
- Diffusion

A certains champs seront attribuées des valeurs par défaut correspondant au cas de figure le plus fréquent :

- Support = Couleur
- Format = 24x36
- Point rouge = non
- Confidentialité = non

III - LES FONCTIONNALITES

Outre les fonctionnalités classiques des logiciels documentaires, la base de données doit assurer les fonctions suivantes :

- Gestion de notices à deux niveaux (voir infra)
- Délivrance de plusieurs numéros et gestion des liens entre ces numéros :
 - * un numéro séquentiel absolu par photo
 - * un numéro d'ordre par photo au sein d'un lot
 - * un numéro de série lorsque celui-ci n'existe pas (cas des photos fraîches ou de certains reportages d'archives)
- Saisie plein écran, avec plusieurs masques disponibles
- Subdivision d'un champ en sous-champ et d'un sous-champ en éléments, avec liaisons des éléments d'un même sous-champ (pour créer un lien entre, par exemple, un comédien et son rôle)
- Gestion de lexiques de noms de personnes et de descripteurs, avec contrôle avant validation lors de la saisie
- Utilisation, pour certains champs, de codes alpha-numériques qui seront remplacés, à la saisie, comme à la recherche, par les valeurs qu'ils représentent. Cette procédure a pour but d'accélérer la saisie. Les champs "photographe" et "nature du reportage" seront entièrement codés, le champ "descripteur" partiellement (voir annexes).
- Affichage des résultats de la recherche sur deux niveaux, avec éventuellement un tri sur le champ "date".
- Utilisation, pour la saisie d'informations de 2ème niveau (niveau photo), d'une procédure permettant d'éviter de saisir plusieurs fois les mêmes données : la documentaliste saisit une information au clavier, puis valide pour ce critère les photos pertinentes, par simple

lecture du code à barres. Si toutes les photos du lot répondent à ce critère, la documentaliste le spécifie par une touche de fonction.

IV - LA MISE EN OEUVRE

Le logiciel choisi pour l'application KIPA est BASIS, progiciel commercialisé en France par la société ID France, filiale de Battelle. Ce choix a été fait pour des raisons d'opportunité, mais aussi suivant des critères techniques :

- La structure relationnelle de BASIS lui permet de gérer des notices à deux niveaux, donnée capitale de l'application, tout en proposant les performances habituelles des logiciels documentaires.
- Sa structure modulaire permet de modifier facilement l'application, sans devoir reparamétrer toute la base.
- BASIS est assez bien implanté dans la presse (Reuters et George Outram en Grande-Bretagne, Axel Springer en RFA, Le Monde, Bayard-Presses et bientôt EDI7 en France); son acquisition récente par le ministère italien de la Culture lui ouvre les portes des fonds culturels.

Mais BASIS a également des défauts :

- Son prix, relativement élevé : 64 000 F sur MicroVAX 2000
- Pas de mise à jour en temps réel
- L'organisation en modules est parfois contraignante (pas d'accès au thesaurus en ligne lors de la saisie, par exemple)
- Enfin un faible degré d'ergonomie, qui ne reprend pas en compte les développements de ces dernières années : fenêtres de visualisation, scrolling, video inverse, etc ...

Les fonctionnalités souhaitées demandant beaucoup de développements en sus du logiciel standard, 30 jours d'un ingénieur spécialiste ont été comptés pour mettre au point l'application. Pour des raisons de disponibilité (la réalisation a eu lieu au mois d'août), ID France en a confié la responsabilité à Bayard Service Informatique (BSI), SSCI reliée au groupe Bayard Presse (dont un des journaux, Le Pèlerin, est par ailleurs client pilote du messageur).

L'application a été portée sur MicroVAX 2000, matériel construit par Digital Equipment Corporation :

- 6 Mo de mémoire centrale
- 1 disque de 159 Mo
- Logiciel d'exploitation MicroVMS

Afin de gagner du temps, on a opéré a la présaisie d'un des rares fichiers manuels directement exploitables : le fichier des films, classé par ordre alphabétique des titres. Chaque fiche contient les informations suivantes :

- Titre du film
- Metteur en scène
- Principaux comédiens
- Date
- Numéro de négatif
- Numéro de planche couleur

Seules les fiches des films couleur ont été reprises, le noir et blanc posant des problèmes de manipulation lors de la digitalisation qui seront réglés plus tard. Pour permettre cette présaisie, BSI a fourni , avant d'avoir finalisé sa prestation, le logiciel standard BASIS et un masque de saisie; le travail a pu commencer le 19 août.

INTEGRATION DE LA BASE DE DONNES

La base de données documentaire n'est qu'une des briques de l'édifice mis en place à KIPA. La saisie d'information textuelle s'insère dans une chaîne de traitement des photos, où les opérations se suivent, toujours dans le même ordre : sélection, indexation sommaire, étiquetage, numérisation, compression, reclassement en panodias, indexation fine, rangement.

D'autre part, l'utilisateur disposant à la fois du texte et de l'image, le terminal aura à gérer l'interaction de deux modes opératoires : recherche visuelle et recherche textuelle.

I - LA CHAINE DE TRAITEMENT

1 - Sélection des photos à numériser - Elimination des mauvaises photos - Mise de côté des doublons - Remplacement des caches détériorés.

2 - Regroupement des photos en lots homogènes, correspondant à une unité de premier niveau. Pour les photos fraîches, il s'agit de déterminer leur destination finale, pour les photos d'archives de corriger de petits déclassements locaux.

3 - Saisie au clavier des informations de base concernant un lot : code de classement, numéro de série, autres numéros, date, titre, photographe, nombre de photos.

4 - Après validation de la fiche précédente, la base de données générera automatiquement un numéro de série si celui-ci est manquant, et les numéros individuels des photos. Toutes ces informations seront récupérées par un PC qui commandera une imprimante code à barres. Sur chaque cache de photo sera collée une étiquette reprenant les informations saisies à l'étape précédente, avec le numéro séquentiel sous forme de code à barres. Le titre sera tronqué à 28 caractères. A cette étape seront également imprimées des étiquettes grand format, pour les panodias.

5 - Digitalisation . Chaque image numérisée sera accompagnée dans la banque d'image des informations saisies à l'étape 3 et du numéro de code à barres. L'opérateur pourra éventuellement corriger l'image et devra indiquer si elle est verticale ou horizontale. Celle-ci sera ensuite compressée puis stockée.

6 - Reclassement des photos dans les panodias correspondants

7 - Rappel de la fiche documentaire saisie à l'étape 3, par simple lecture du code à barres d'une des photos du lot. Saisie d'informations complémentaires au niveau du lot puis au niveau de chaque photo, ou plutôt de sous-groupe de photos (voir pp 20-21).

8 - Rangement des photos en rayon ou envoi au labo pour duplication dans le cas des photos fraîches (en cas d'urgence, les duplicatas seront faits entre les étapes 6 et 7).

II - PROFIL DES POSTES

Des opérations de simulation ont permis de fixer à 5 le nombre de personnes nécessaires, en régime de croisière, pour tenir le rythme de 2000 images par jour.

Poste 1 : étapes 1, 2 et 3 : 2 personnes

Les personnels devront surtout avoir le "coup d'oeil", puisqu'il s'agira de passer toutes les photos en revue à l'aide d'un compte-fil (sorte de loupe), au dessus d'une table lumineuse, afin d'écartier ce qui est flou ou de peu d'intérêt. Une expérience de la photo est indispensable; en revanche, ce poste ne demande pas une connaissance approfondie du logiciel documentaire. L'alternance des tâches de sélection et de saisie au clavier devrait permettre de limiter la fatigue visuelle due à la table lumineuse.

Poste 2 : étapes 4 et 5 : 1 personne

Ce poste pourrait être tenu par un assistant de laboratoire; il s'agit surtout d'opérations de manipulation des photos et de correction de couleur. En début de réalisation, certaines tâches, qui par la suite seront automatisées, devront être effectuées manuellement (comme de présenter la diapo sous l'imprimante code à barres par exemple).

Poste 3 : étapes 6, 7 et 8 : 2 personnes

Ce poste sera tenu par deux documentalistes, qui ont reçu une formation à l'application. Ces deux personnes seront responsables de la base de données : elles devront notamment assurer les sauvegardes, les mises à jour et la gestion du vocabulaire contrôlé.

Les postes 1 et 3 travailleront en rythme différé. La simulation a en effet montré qu'un reportage long pénalisait le poste 2 par rapport au poste 1, alors que l'inverse se produisait pour un reportage court. L'équilibre est atteint sur une dizaine de reportages.

A chacun des postes 1 et 3 se trouvera au moins une personne travaillant déjà à l'agence, l'autre personne venant de l'extérieur.

III - LE TERMINAL

Plusieurs logiciels documentaires (Texto, logiciel puss, BRS) ont été interfacés pour permettre de gérer l'interaction

de l'image et du texte sur micro-imageur et vidéodisque. Le problème particulier posé par le Messager est que l'image arrive d'abord réduite au 1/9 puis ensuite plein écran, et que surtout le terminal n'a pas d'accès direct à l'ensemble des données (du moins pour le terminal client), qui lui arrivent sous forme fragmentaire.

Dans un premier temps, le Messager ne sera pas relié à la base de données : la seule information textuelle disponible sur le terminal sera celle qui est associée aux images dans la banque d'images elle-même (90 caractères). Simplement, KIPA aura la possibilité, avant le téléchargement des images du jour, de leur adjoindre de courts textes de promotion.

A terme, le client devra pouvoir facilement :

CHOISIR, parmi : - les photos du jour de plusieurs agences
- les photos d'archives demandées
- les chapitres mis en mémoire

VISIONNER : - par panneaux de 9
- plein écran

MANIPULER, par panneaux de 9 : - tourner les pages
- éliminer
- sélectionner
- réorganiser les panneaux
- former des chapitres

LIRE la documentation associée : - 90 caractères de légende
- un texte de promotion
- la fiche d'un lot
- la fiche d'une photo

COMMANDER les originaux aux agences, par messagerie.

CONCLUSION

Ouvrir une conclusion est une bonne occasion pour passer du style impersonnel et quelque peu guindé d'un rapport de stage à la subjectivité triomphante d'une première personne du singulier.

"Je" profiterai donc de cette dernière page pour dire la fierté que je ressens d'avoir participé à un projet aussi ambitieux que risqué, puisqu'il ne vise rien de moins que de révolutionner la distribution des images fixes en France (et ailleurs).

Car pas de demi-mesures : aucun journal n'acceptera de payer 6000 F par mois si seuls les fonds de Kipa lui sont offerts sur son terminal. L'affaire ne sera rentable que lorsque 4, 10, 80 agences seront montées sur le Messager.

Projet un peu fou, quand on y réfléchit : mettre en place, en quelques mois, un système destiné à gérer des millions d'images numérisées (quand d'autres mettent des années à réaliser un vidéodisque : 54 000 malheureuses images). Un défi que seuls des gens un peu fous comme ceux de TRIBVN pouvaient relever.

Et pourtant, si l'on réfléchit un peu plus, n'est-il pas inéluctable que nous nous engagions dans cette voie-là ? En soi, aucun des éléments du projet Messager n'est très nouveau, c'est le fait de les rassembler qui est novateur.

Mon rôle dans tout cela : la documentation (et un peu plus). Une partie moins brillante peut-être que le reste , car moins nouvelle (il y a bien 20 ans maintenant que l'on sait gérer des bases de données textuelles), mais pourtant indispensable : comment s'y retrouver dans toutes ces photos, mêmes numérisées, sans base de données ?

Je retirerai aussi de ce stage la satisfaction d'avoir vu beaucoup de choses, beaucoup de systèmes liés à l'images et d'avoir pu rencontrer beaucoup de gens, au Musée de la mode, au centre de doc de la RATP, à la vidéothèque de Paris, à l'Institut du monde arabe, à la BPI, au ministère de la Culture, à la BBC, chez Logiciel puss, ID France, EDI7, Gamma, VSD, France-Soir, Burda, ... et bien sûr chez KIPA et TRIBVN.

ANNEXES

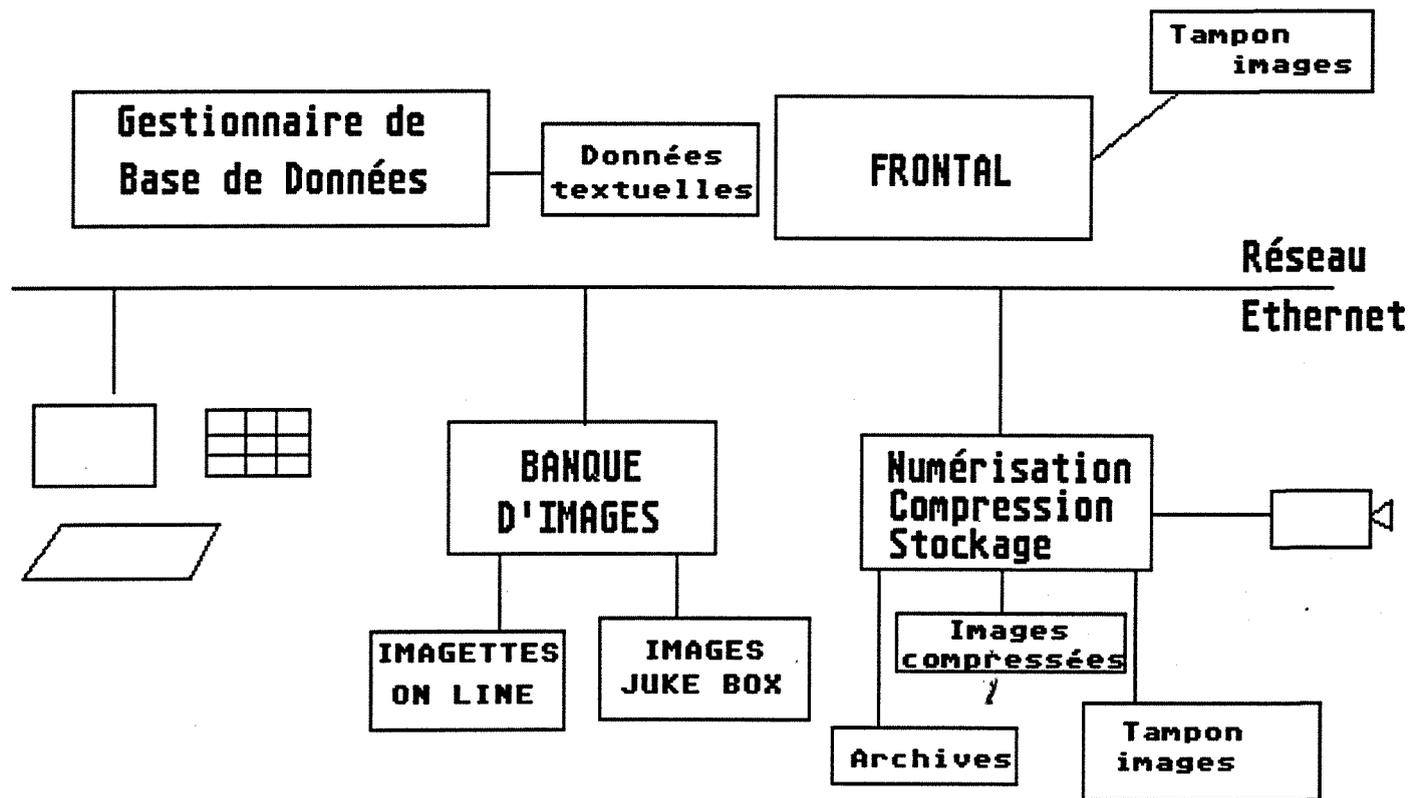
SYNOPTIQUE DU SERVEUR

VERSION

DE

BASE

OPTIONS



AGENCE KIPA : COMPOSITION DES FONDS

ALPHA	Photos de personnalités de la télévision, du spectacle, du show-biz et, d'une façon générale de toute personne très médiatique
CLASSEURS	Reportages sur des émissions de télévisions, des tournages de films, de séries télévisées, des pièces de théâtre, des événements culturels et médiatiques divers.
VIE MONDAINE	La vie du Tout-Paris : soirées de galas, diners mondains, mariages célèbres, etc
MODE	Les défilés de haute-couture et divers événements liés à la mode
ILLUSTRATION GENERALE	Reportages sur les sujets les plus divers (de l'exposition canine à l'emballage du Pont Neuf); c'est un fonds marginal, mais qui tourne quand même.
L.G.I.	Photos de personnalités américaines, ou de passage aux Etats-Unis, et susceptibles d'intéresser le public français; elles proviennent d'une agence américaine qui occupe le même créneau que KIPA
DANSE	Photos des productions chorégraphiques des 20 dernières années : les spectacles eux-même et des portraits des artistes
OPERA	Idem pour le lyrique
BOTTI	Photos des années 50 et 60, sur des sujets divers, mais dans l'axe du positionnement de l'agence : spectacle, télé, show biz
SPITZER	Idem
N.Y.	Photos récupérées auprès de l'agence Globe même type de fonds que LGI, en plus vieux
CINEMA	Photos de production de près de 20 000 films (dont 4000 en couleur), auxquelles il faut rajouter des tirages récupérés auprès de la Fédération française des ciné-clubs.

Codes à utiliser pour le champ : Nature du reportage

1. Plateau télévision
2. Téléfilm
3. Feuilleton, série
4. Télévision étrangère
5. Film
6. Tournage film
7. Théâtre
8. Opéra
9. Danse
10. Concert classique
11. Concert non classique
12. Variétés
13. Cirque
14. Autres spectacles
15. Evènement mondain
16. Rendez-vous studio
17. Rendez-vous domicile
18. Rendez-vous extérieur
19. Divers

CODES A UTILISER POUR LE CHAMP DESCRIPTEUR**APPROCHE DU SUJET**

A1 = Portrait
 A2 = Portrait en situation
 A3 = Pose
 A4 = Dos
 A5 = Ensemble
 A6 = Flou artistique
 A7 = Décor

LOOK

B1 = Costume de ville
 B2 = Décontracté
 B3 = Tenue de soirée
 B4 = Fantaisie
 B5 = Maillot de bain
 B6 = Nu
 B7 = Tenue de sport
 B8 = Costume d'époque
 B9 = Uniforme
 B10 = Coiffure originale
 B11 = Chauve
 B12 = Barbe
 B13 = Moustache

ACCESSOIRES

. C1 = Lunettes
 . C2 = Chapeau
 . C3 = Cigare
 . C4 = Cigarette
 . C5 = Pipe
 . C6 = Téléphone
 . C7 = Micro
 . C8 = Instrument de musique
 . C9 = Piano
 . C10 = Guitare
 . C11 = Arme
 . C12 = Fleurs
 . C13 = Parapluie

DIVERS

. D1 = Enfant
 . D2 = Bébé
 . D3 = Femme enceinte
 . D4 = Voiture
 . D5 = Moto
 . D6 = Vélo
 . D7 = Animal
 . D8 = Chien
 . D9 = Chat

LIEU

. E1 = Rue
 . E2 = Café
 . E3 = Campagne
 . E4 = Jardin
 . E5 = Mer
 . E6 = Neige

ACTION

. F1 = Baiser
 . F2 = Violence
 . F3 = Course
 . F4 = Rire

GENRE (CINEMA)

. G1 = Policier
 . G2 = Fantastique
 . G3 = Guerre
 . G4 = Historique
 . G5 = Western
 . G6 = Musical
 . G7 = Dessin animé
 . G8 = Horreur



17

18



* 9 5 4 0 8 9 F *