



**Enseignement général de défense**  
**Cycle « Techniques d'expression et travail en équipe »**

Synthèse de la conférence sur le thème

**"Environnement de l'information - recherche sur Internet, outils"**

prononcée par Messieurs Eric Balançat et Jean-Pascal Rengnet  
le 9 septembre 2004.

## Fiche documentaire

1. Conférence du cycle TE2 « environnement de l'information – recherche sur Internet, outils » par Messieurs Eric Balançat et Jean-Pascal Rengnet.
2. Recherche sur Internet.doc
3. Commandant (air) Sébastien Faivre (France)
4. 9 septembre 2004
5. Division A
6. Compte rendu de la conférence sur la maîtrise des systèmes informatiques.
7. En préalable à une phase pratique d'application, les conférenciers se sont attachés à décrire l'environnement de l'information que représente Internet et les méthodes facilitant toute recherche sur ce réseau.
8. Mots clefs : *Internet-Recherche-Information*.



# CYCLE DE « TECHNIQUES D'EXPRESSION ET TRAVAIL EN EQUIPE »

COLLEGE INTERARMEES  
DE DEFENSE  
CDT FAIVRE (A3)

Synthèse de la conférence sur le thème

## "Environnement de l'information - recherche sur Internet, outils"

prononcée par Messieurs Eric Balançat et Jean-Pascal Rengnet  
le 9 septembre 2004.

Durant sa scolarité au Collège Interarmées de Défense (CID), tout stagiaire est amené, à un moment ou à un autre, à se tourner vers ce formidable outil de recherche en matière d'informations que représente Internet. Si cette manne quasi-intarissable représente un formidable progrès pour chacun, il n'en reste pas moins vrai que, comme pour toute information, se posent des problèmes en termes de ciblage, d'accès, de validation et de validité des documents ainsi obtenus.

Cette conférence a pour objet d'une part d'éclairer la nébuleuse qui entoure le sigle www (world wide web), plus communément appelé Internet, et d'autre part de fournir au stagiaire du CID quelques méthodes de recherche fort utiles lui permettant à l'avenir de se retrouver dans ce dédale d'informations.

### Environnement de l'information

L'année 1992 voit la naissance du world wide web. Ce réseau permet d'accéder à un véritable système de réseaux parallèles pas toujours visibles (exemple : GOPHER, USENET ou FTP). Cette partie émergée de l'iceberg abrite une masse considérable d'informations<sup>1</sup> auxquelles s'ajoutent nombre de pages non référencées (Web invisible) dans les différents robots et moteurs de recherche.

L'**interlocuteur** se présente sous deux formes distinctes : le site Internet ou le forum de discussion. C'est sous cette dernière qu'apparaissent en général la majorité des sites personnels.

L'**information** à laquelle Internet permet d'accéder peut être de nature :

- ouverte ou blanche ; elle est accessible à tous,
- semi-ouverte ou grise ; elle est accessible mais se caractérise par des difficultés dans la connaissance de son existence ou de son accès,
- ou encore fermée ou noire ; son accès est protégé.

Le réseau Internet n'est pas sans danger et nécessite la mise en œuvre de certaines **précautions** élémentaires (antivirus, « pare-feu »). Il est le reflet de notre société, la rapidité des échanges en plus, et s'accompagne de son cortège de dérives et déboires (sites interdits à caractère pédophile, raciste ..., attaques « virales », infiltration de

---

<sup>1</sup> Si les chiffres disponibles sont rares, imprécis, voire tendancieux, il n'en reste pas moins vrai que le volume d'informations évolue à une vitesse considérable avec un taux de l'ordre d'un million de pages supplémentaires recensées chaque jour.

programmes (« backdoor », « cheval de Troie »)). Dans tous les cas, la meilleure des protections reste la mise en place d'un poste dédié totalement isolé.

### **Méthode de recherche**

La recherche d'informations sur Internet s'articule en 5 phases successives.

Première phase et préliminaire incontournable, l'**analyse** approfondie des besoins et l'élaboration d'un véritable plan de renseignement (qui, quoi, où, pourquoi et comment ?) autorise un gain de temps substantiel par la suite. Cette première phase de ciblage permet notamment d'éliminer nombre d'informations peu pertinentes, de surcroît du fait de l'existence de plus en plus de sites commerciaux qui faussent la donne. La phase de **recherche** proprement dite peut alors commencer au moyen des outils dédiés décrits infra.

Vient ensuite la phase de **recueil** de l'information. Les moyens les plus couramment utilisés sont le « copier-coller », le transfert de fichiers et le rapatriement de données. Il importe ensuite de l'**exploiter**, d'une part en la hiérarchisant selon sa pertinence et d'autre part en la validant, si nécessaire, notamment suivant un processus de recoupement avec d'autres sources.

Enfin, cette information peut être **diffusée** en faisant appel aux technologies et moyens de communication en place au sein de l'entité.

### **Outils de recherche**

On distingue en la matière deux grandes familles, les **moteurs** de recherche (exemple : *Google*) qui, malgré une surabondance de l'information (« bruit »), présentent l'avantage d'assurer une excellente couverture du réseau et les **annuaires thématiques** (exemple : *Yahoo*) qui recensent un certain nombre de sites au travers de fiches descriptives et présentent l'avantage de limiter le « bruit » dans les réponses mais aussi l'inconvénient d'assurer une couverture moindre. Pour chacun de ces outils, des fonctionnalités avancées (de type opérateurs booléens (union, intersection, exclusion ...), mots clefs) permettent d'affiner la recherche.

D'autres outils existent, tels les méta-outils (exemple : le méta-moteur de recherche *Copernic*) dont le principe de fonctionnement est d'adresser la requête à plusieurs moteurs et/ou annuaires simultanément.

### **Conclusion**

« Savoir pour prévoir afin de pouvoir » disait Auguste Comte. Cette phrase illustre parfaitement l'importance fondamentale de la maîtrise de l'information et son caractère capital, en tant que matière première stratégique, notamment dans le cadre de la prévention. A cet égard, Internet constitue une source considérable et quasi-intarissable d'informations qui ne doit en aucun cas être négligée. La connaissance de cet environnement virtuel et des méthodes favorisant son accès est de toute évidence incontournable à l'aube du 21<sup>ème</sup> siècle.