# SPAM: Analyser vous-même une en-tête de mail – Partie 1

09/05/2021


## Bonjour le Monde !

**Dans le premier article de cette série**, j'ai essayé de vous expliquer ce qu'est un **SPAM**, ce que font les fournisseurs de service mail pour le combattre et comment régler vos programmes de mails pour gérer les différents modes de prévisualisation des messages afin de minimiser les risques dus aux SPAMs !

Vous savez déjà faire un bon tri entre SPAM et mail légitime en regardant si vous connaissez l'expéditeur et si l'objet du mail est cohérent avec cet expéditeur, si l'objet du mail n'est pas dans votre langue, etc…

Comme promis, je vais essayer de vous apprendre à analyser un message que vous soupçonnez être un SPAM quand vos méthodes habituelles de détection

n’ont pas permis de décider si c’était vraiment un SPAM ou non !

Pour l’apprentissage, je vous fournis un exemple d’une de mes newsletter qui n’est donc pas un SPAM (qui a dit “un peu quand même” ?) ! Le fichier s’appelle “**En-tete-Newsletter-Li-Pti-fouineu-Gmail.txt**”

**Ce fichier texte peut être téléchargé depuis le bas de cet article (mais aussi depuis le premier article)** et vous pouvez donc l’ouvrir séparément pour comparer avec ce que je vais vous expliquer ici !

**La première chose à faire est d’isoler la partie en-tête du corps du message**; le début du texte est la fameuse en-tête recherchée, le reste est le message lui-même en langage HTML et encodé en MIME (voir plus loin) le plus souvent mais il peut aussi être en texte simple (et donc plus facilement lisible mais nettement moins joliment présenté):

De la 1ère ligne de texte jusqu’au dernier “**From**“, il s’agit de l’en-tête que nous allons analyser.

Dès la ligne contenant “**—-_SiB-b9e2b1a74c63c10b-Part_1**“, il s’agit du début du message lui-même qui n’est pas très compréhensible mais votre navigateur ou programme mail sait quoi faire avec tout ça et vous afficher ma très jolie newsletter !!!

Veuillez noter que j’ai remplacé manuellement le “@” de mes adresses mail par ” at ” pour ne pas que ces adresses soient capturées par des robots automatiques.

**Notre en-tête à analyser est donc**:

```
From — Wed Mar 24 14:19:54 2021
X-Mozilla-Status: 0001
X-Mozilla-Status2: 00000000
X-Mozilla-Keys:
Delivered-To: ecollart at gmail.com
Received: by 2002:ac2:5223:0:0:0:0:0 with SMTP id
i3csp1865229lfl;
Sat, 20 Mar 2021 02:46:43 -0700 (PDT)
X-Google-Smtp-Source:
ABdhPJyunsaf4BTo5uPrj5PRhPaqShjfKKXGeV2yvLG2b91GsEnXWAzRJBLeMh
```

OchKmVUb/FxHoJ
X-Received: by 2002:a5d:6ca6:: with SMTP id a6mr8501914wra.179.1616233603805;
Sat, 20 Mar 2021 02:46:43 -0700 (PDT)
ARC-Seal: i=1; a=rsa-sha256; t=1616233603; cv=none;
d=google.com; s=arc-20160816;
b=lbiP+sZliEGfTV+JsfOIch1FLFU98N+pGFIgqTEsmvLJcq9o7/zkDgRHRq8f
BzFduJ
nl6g/P+k5YtcdhfDNZT1u1ts5XojXn+i2JUaNEz0osTFDgcqe24VjKBieZe+6W
yyBPtd
VsHedPuaOHmCzEllQRHWEexUS49SRYYODtjvcbwtcUNyfSwypVAHzG6TIWXpde
5kmFi2
cDDnNvZSgmu63yiWWtdzmzYGENfZMCU2IIZBveeHJ7CCmutl/Ur2FJsnOxYZ0b
JlyhoB
0DQmYPCgEGHZWDJQL+ZQpRXGyQvjLnYZhPLF1KeqBSh84TA/bxpCOMWh0bz8Dy
DV2k5S
DfYg==
ARC-Message-Signature: i=1; a=rsa-sha256; c=relaxed/relaxed;
d=google.com; s=arc-20160816;
h=from:feedback-id:mime-version:reply-to:list-unsubscribe-post
:list-unsubscribe:precedence:origin-messageid:message-
id:subject
:date:to:dkim-signature;
bh=UU+U3sdvC9zoej27S/yyN6OJLhA6PJXhy1KPAg7JP0U=;
b=Epr9X2ikj4kaNDjYe6SPDwRo3X5+/I89xezlbhPAvJfNKgYLMoyzomePSEJv
+JFfN5
+C8gunqjxVYMtmotj6afGIqBHETfuMA9hSr99yoL2LI7XyhKabO2KpkR9xplje
ogp4Qd
7bkVLhcNG5e1G0v/3DbQYwyViwF++POL1fE2g6RAYtiHYHz66C6Eg/MkstTNdy
YlmHjm
w355WgJAkuhamEhiFdbG7ICBwkqe+K1OBWHKveKymM1JiviM3ORPsbh7BAI0zu
hrTevA
pK2sqOqlJSUFmXTWTrQxrvMKJ14orowkscgt4WboX0F6w4uSaNDd76RyECioa4
7BEe/B
I4MQ==
ARC-Authentication-Results: i=1; mx.google.com;
dkim=pass header.i=@sendinblue.com header.s=mail

header.b=spob+wMu;
spf=pass (google.com: domain of bounces-113727262-ecollart=gmail.com@ae.d.mailin.fr designates 185.41.28.5 as permitted sender) smtp.mailfrom="bounces-113727262-ecollart=gmail.com@ae.d.mailin.fr"
Return-Path: <bounces-113727262-ecollart=gmail.com@ae.d.mailin.fr>
Received: from ae.d.mailin.fr (ae.d.mailin.fr. [185.41.28.5])
by mx.google.com with ESMTPS id 88si7529256wrn.466.2021.03.20.02.46.43
for <ecollart at gmail.com>
(version=TLS1_3 cipher=TLS_AES_256_GCM_SHA384 bits=256/256);
Sat, 20 Mar 2021 02:46:43 -0700 (PDT)
Received-SPF: pass (google.com: domain of bounces-113727262-ecollart=gmail.com@ae.d.mailin.fr designates 185.41.28.5 as permitted sender) client-ip=185.41.28.5;
Authentication-Results: mx.google.com;
dkim=pass header.i=@sendinblue.com header.s=mail header.b=spob+wMu;
spf=pass (google.com: domain of bounces-113727262-ecollart=gmail.com@ae.d.mailin.fr designates 185.41.28.5 as permitted sender) smtp.mailfrom="bounces-113727262-ecollart=gmail.com@ae.d.mailin.fr"
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed; d=sendinblue.com;
q=dns/txt; s=mail;
bh=UU+U3sdvC9zoej27S/yyN6OJLhA6PJXhy1KPAg7JP0U=;
h=from:reply-to:subject:date:mime-version:content-type:list-unsubscribe:x-csa-complaints:list-unsubscribe-post;
b=spob+wMuYyi7rjEc+TUKm+JyEfmk1ZreLl8LVRID4X6Bx08Nz1OeYyvLDT8BuU7FthiCDyiVZv2s
7N/BB+ksl8Aw1xW5LFtDjTNqu03psbpVp3Or/y8to3fQSOHsyualvZ2B2mhXrgyykqIMEH3dXs8M
/GxyPLFnk1BL/0Qyohc=
X-Mailin-EID:
MTEzNzI3MjYyfmVjb2xsYXJ0QGdtYWlsLmNvbX48MjAyMTAzMjAxMDQ2Ljg4MzIyMTk4ODgyQHNtdHAtcmVsYXkubWFpbGluLmZyPn5hZS5kLm1haWxpbi5mcg%3

```
D%3D
To: <ecollart at gmail.com>
Date: Sat, 20 Mar 2021 09:46:43 +0000
Subject:                    =?UTF-8?Q?Li_P’ti_Fouineu_-
_Je_me_suis_fait_virer_de_mon_boulot_=E2=80=93_=C3=89pisode_3?
=
Message-Id:    <cd0731c1-235a-4f51-97bb-1b0635ae74fa@smtp-
relay.sendinblue.com>
Origin-messageId:      <202103201046.88322198882@smtp-
relay.mailin.fr>
Content-Type:   multipart/alternative;   boundary=”–_SiB-
b9e2b1a74c63c10b-Part_1″
Precedence: bulk
X-Newsletter-Email-ID: 64
X-Auto-Response-Suppress: OOF, AutoReply
List-Unsubscribe:
<https://ecollart.xyz/?na=uc&nk=103-f92464eb84&nek=64->
List-Unsubscribe-Post: List-Unsubscribe=One-Click
Reply-To: info at ecollart.xyz
MIME-Version: 1.0
X-sib-id:      hAi3jun64DFqP3HToivmmiBLf5DJzMgCzxa6R-
taaDY8WH8RqUSJPYWF98f-z1Lvf2acUlA78CL5d2Zyo4lbPXeTgnzcKt60VQg-
lXLtherFhpYGRvaSj_EUHOfj4WiUj-
lCu03Abm3IkD6nMbY4a1D1Rrw1Y7SWZaeIa5c3aC4
X-CSA-Complaints: whitelist-complaints@eco.de
Feedback-ID: 185.41.28.5:3210647_-1:3210647:Sendinblue
From: “Li P’ti Fouineu” <info at ecollart.xyz>
```

---

Pour savoir de qui ce message vient, par où il est passé et qui va recevoir la réponse si on répond, **il faut commencer par la dernière ligne de l’en-tête qui est donc**:

**From: “Li P’ti Fouineu” <info at ecollart.xyz>**

Cette ligne indique l’expéditeur probable (et oui, un pirate peut tricher ici)

---

On remonte ensuite:

**Feedback-ID: 185.41.28.5:3210647_-1:3210647:Sendinblue**

Cette ligne vous informe que cette newsletter est envoyée par le service SendinBlue qui collecte feedback et statistiques pour ce message qui a reçu un numéro d'identification (ID) unique pour cela.
Toutes mes newsletters sont envoyées par SendinBlue et tout est donc normal jusque-là. Il ne faut pas hésiter à comparer un message sur lequel on a un doute avec un ou plusieurs précédents venant de la même source.

---

On continue de remonter:

```
MIME-Version: 1.0
X-sib-id:      hAi3jun64DFqP3HToivmmiBLf5DJzMgCzxa6R-
taaDY8WH8RqUSJPYWF98f-z1Lvf2acUlA78CL5d2Zyo4lbPXeTgnzcKt60VQg-
lXLtherFhpYGRvaSj_EUHOfj4WiUj-
lCu03Abm3IkD6nMbY4a1D1Rrw1Y7SWZaeIa5c3aC4
X-CSA-Complaints: whitelist-complaints@eco.de
```

Ces 3 lignes ne nous intéressent pas et concernent la popote interne de la messagerie d'envoi.
MIME est le système utilisé pour encoder les différentes composantes du message comme les pièces jointes qui ne sont pas du texte. Un mail encodé n'utilise QUE des lettres non-accentuées, des chiffres et quelques rares signes spéciaux ! Les pièce jointes qui ne sont pas du texte sont donc encodées selon le format MIME. Cet encodage fait que la taille d'une pièce jointe dans un mail encodé fait en moyenne 1,33x la taille originale de la pièce jointe. C'est pour cette raison que si votre messagerie est limitée à 10MB par message que celui-ci peut être refusé si la pièce jointe que essayez d'envoyer a une taille de 9MB car 9 x 1,33 = 11,97MB !

---

On remonte encore:

**Reply-To: info at ecollart.xyz**

**Cette ligne est très importante** et nous apprend vers qui sera envoyée une réponse éventuelle au message reçu.
**Pour ne pas risquer d'être considéré comme SPAM, les bonnes pratiques demandent que le "Reply-To" soit identique au "From" précédemment noté ou au moins que ce soit une adresse mail du même domaine (qui est ici "ecollart.xyz") !**
Ici, ma newsletter respecte donc les bonnes pratiques et les outils anti-spam ne m'attribueront pas de mauvais score.

---

Remontons toujours plus haut:

```
Precedence: bulk
X-Newsletter-Email-ID: 64
X-Auto-Response-Suppress: OOF, AutoReply
List-Unsubscribe:
<https://ecollart.xyz/?na=uc&nk=103-f92464eb84&nek=64->
List-Unsubscribe-Post: List-Unsubscribe=One-Click
```

Ces lignes sont des informations ajoutées par mon plugin WordPress qui crée automatiquement une nouvelle newsletter lorsque je publie un article; ce plugin s'appelle Newsletter
Les 2 lignes où vous voyez "**Unsubscribe**" sont utilisée pour vous désinscrire de la bonne liste d'envoi (ben oui, je pourrais en faire plusieurs) si vous deviez cliquez sur le lien de bas de page de la newsletter intitulé "Pour modifier ou annuler votre abonnement, cliquez ici."
La ligne "**X-Newsletter-Email-ID: 64**"est l'identification unique (ID) de la newsletter crée par le plugin.
Je ne sais pas à quoi correspondent les autres lignes mais elles sont utilisées par le plugin.

---

Vous êtes toujours là ? Et bien on continue de remonter dans le texte:

```
Content-Type: multipart/alternative; boundary="_SiB-
```

**b9e2b1a74c63c10b-Part_1**“

**Cette ligne indique où commence le corps du message**. Maintenant que vous le savez, vous pourrez utiliser cette information pour une prochaine analyse de message douteux qui sera peut-être plus complexe que mon exemple.

---

Continuons vers le haut:

```
Message-Id:    <cd0731c1-235a-4f51-97bb-1b0635ae74fa@smtp-relay.sendinblue.com>
Origin-messageId:      <202103201046.88322198882@smtp-relay.mailin.fr>
```

Ces 2 lignes fournissent l’identification unique (ID) assignée par chaque serveur de SendinBlue. Mailin.fr fait partie de SendinBlue mais vous ne pouvez savoir cela que si vous êtes un utilisateur de SendinBlue qui, je le rappelle, est le fournisseur de messagerie de mon site web.
Dans les cas complexes, ces identifications sont recherchées sur tout le parcours du message entre SendinBlue qui l’envoie et votre serveur de mail qui va le recevoir.

---

Mais encore:

```
To: <ecollart at gmail.com>
Date: Sat, 20 Mar 2021 09:46:43 +0000
Subject:                    =?UTF-8?Q?Li_P’ti_Fouineu_-_Je_me_suis_fait_virer_de_mon_boulot_=E2=80=93_=C3=89pisode_3?=
```

Là on voit à qui le message est envoyé (**To:**); c’est là que vous verrez votre propre adresse mail si vous analysez un de vos messages reçus.
La date et l’heure de création du message, souvent mais pas toujours à l’heure GMT ou UTC (c’est la même chose, c’est le “+0000”). L’utilisation de l’heure GMT permet de comparer plus facilement les log des serveurs quelque soit le pays où est situé le serveur.
L’objet du message indique l’encodage du texte (UTF-8) mais n’aime ni les

espaces (remplacés ici par “_”) ni les accents, apostrophes et autres  (remplacés ici par leur code UTF-8)

---

Le morceau qui suit est moins intéressant:

```
ARC-Seal: i=1; a=rsa-sha256; t=1616233603; cv=none;
        d=google.com; s=arc-20160816;

b=lbiP+sZliEGfTV+JsfOIch1FLFU98N+pGFIgqTEsmvLJcq9o7/zkDgRHRq8f
BzFduJ

nl6g/P+k5YtcdhfDNZT1u1ts5XojXn+i2JUaNEz0osTFDgcqe24VjKBieZe+6W
yyBPtd

VsHedPuaOHmCzEllQRHWEexUS49SRYYODtjvcbwtcUNyfSwypVAHzG6TIWXpde
5kmFi2

cDDnNvZSgmu63yiWWtdzmzYGENfZMCU2IIZBveeHJ7CCmutl/Ur2FJsnOxYZ0b
JlyhoB

0DQmYPCgEGHZWDJQL+ZQpRXGyQvjLnYZhPLF1KeqBSh84TA/bxpCOMWh0bz8Dy
DV2k5S
         DfYg==
ARC-Message-Signature: i=1; a=rsa-sha256; c=relaxed/relaxed;
d=google.com; s=arc-20160816;
           h=from:feedback-id:mime-version:reply-to:list-
unsubscribe-post
                  :list-unsubscribe:precedence:origin-
messageid:message-id:subject
         :date:to:dkim-signature;
        bh=UU+U3sdvC9zoej27S/yyN6OJLhA6PJXhy1KPAg7JP0U=;

b=Epr9X2ikj4kaNDjYe6SPDwRo3X5+/I89xezlbhPAvJfNKgYLMoyzomePSEJv
+JFfN5

+C8gunqjxVYMtmotj6afGIqBHETfuMA9hSr99yoL2LI7XyhKabO2KpkR9xplje
```

ogp4Qd

7bkVLhcNG5e1G0v/3DbQYwyViwF++POL1fE2g6RAYtiHYHz66C6Eg/MkstTNdy
YlmHjm

w355WgJAkuhamEhiFdbG7ICBwkqe+K1OBWHKveKymM1JiviM3ORPsbh7BAI0zu
hrTevA

pK2sqOqlJSUFmXTWTrQxrvMKJ14orowkscgt4WboX0F6w4uSaNDd76RyECioa4
7BEe/B
I4MQ==
ARC-Authentication-Results: i=1; mx.google.com;
dkim=pass header.i=@sendinblue.com header.s=mail
header.b=spob+wMu;
spf=pass (google.com: domain of bounces-113727262-
ecollart=gmail.com@ae.d.mailin.fr designates 185.41.28.5 as
permitted sender) smtp.mailfrom="bounces-113727262-
ecollart=gmail.com@ae.d.mailin.fr"
Return-Path: <bounces-113727262-
ecollart=gmail.com@ae.d.mailin.fr>
Received: from ae.d.mailin.fr (ae.d.mailin.fr. [185.41.28.5])
by mx.google.com with ESMTPS id
88si7529256wrn.466.2021.03.20.02.46.43
for <ecollart at gmail.com>
(version=TLS1_3 cipher=TLS_AES_256_GCM_SHA384
bits=256/256);
Sat, 20 Mar 2021 02:46:43 -0700 (PDT)
Received-SPF: pass (google.com: domain of bounces-113727262-
ecollart=gmail.com@ae.d.mailin.fr designates 185.41.28.5 as
permitted sender) client-ip=185.41.28.5;
Authentication-Results: mx.google.com;
dkim=pass header.i=@sendinblue.com header.s=mail
header.b=spob+wMu;
spf=pass (google.com: domain of bounces-113727262-
ecollart=gmail.com@ae.d.mailin.fr designates 185.41.28.5 as
permitted sender) smtp.mailfrom="bounces-113727262-
ecollart=gmail.com@ae.d.mailin.fr"

```
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed;
d=sendinblue.com;
                    q=dns/txt;                    s=mail;
bh=UU+U3sdvC9zoej27S/yyN6OJLhA6PJXhy1KPAg7JP0U=;
  h=from:reply-to:subject:date:mime-version:content-type:list-
unsubscribe:x-csa-complaints:list-unsubscribe-post;

b=spob+wMuYyi7rjEc+TUKm+JyEfmk1ZreLl8LVRID4X6Bx08Nz1OeYyvLDT8B
uU7FthiCDyiVZv2s

7N/BB+ksl8Aw1xW5LFtDjTNqu03psbpVp3Or/y8to3fQSOHsyualvZ2B2mhXrg
yykqIMEH3dXs8M
        /GxyPLFnk1BL/0Qyohc=
X-Mailin-EID:
MTEzNzI3MjYyfmVjb2xsYXJ0QGdtYWlsLmNvbX48MjAyMTAzMjAxMDQ2Ljg4Mz
IyMTk4ODgyQHNtdHAtcmVsYXkubWFpbGluLmZyPn5hZS5kLm1haWxpbi5mcg%3
D%3D
```

**Tout ce fatras concerne l'identification, la validation de la réputation et l'autorisation** des serveurs de SendinBlue à envoyer des messages en masse (certains sites ont des milliers de membres) et de Google (dans ce cas-ci) à recevoir des mails pour ne pas être considéré comme spammeurs !

Ce sont les paramètres **DKIM** et **SPF** que j'ai dû ajouter manuellement dans le serveur DNS maître pour le nom de domaine de mon site. Ça, c'est un truc de geek.

Il y a aussi les paramètres **ARC** qui sont ajoutés par les systèmes antivirus et anti-spam (ici de Google)

---

Qu'est-ce qu'il nous reste:

```
Delivered-To: ecollart at gmail.com
Received: by 2002:ac2:5223:0:0:0:0:0 with SMTP id
i3csp1865229lfl;
        Sat, 20 Mar 2021 02:46:43 -0700 (PDT)
X-Google-Smtp-Source:
```

```
ABdhPJyunsaf4BTo5uPrj5PRhPaqShjfKKXGeV2yvLG2b91GsEnXWAzRJBLeMh
OchKmVUb/FxHoJ
X-Received: by 2002:a5d:6ca6:: with SMTP id
a6mr8501914wra.179.1616233603805;
        Sat, 20 Mar 2021 02:46:43 -0700 (PDT)
```

N'oubliez pas que ceci se lit du bas vers le haut.
Le "**X-Received**" indique quel serveur a reçu votre message en entrée chez Google; Google n'est pas indiqué sur cette ligne mais on peut vérifier en utilisant l'adresse IPv6 (2002:a5d:6ca6::) dans un outil de recherche spécialisé (non, je ne vais pas vous infliger ça).
Il peut arriver que vous trouviez plusieurs "**X-Received**" ce qui pourrait être une indication que votre message passe par un serveur intermédiaire et il y a donc un risque de capture/modification de votre message. **Dans un tel cas, il est intéressant d'essayer de savoir s'il s'agit bien d'un serveur de l'expéditeur ou du destinataire et pas un truc anormal !** Les outils anti-spam font ce genre de vérifications.
Le "X-Google-Smtp-Source" confirme que c'est bien chez Google que c'est arrivé (normal puisque le destinataire est "ecollart at gmail.com" et que Gmail appartient à Google).
Le "Received: by 2002:ac2:5223:0:0:0:0:0" indique quel serveur qui gère ma boîte mail chez Google a reçu le message et quand. PDT est le nom de la "timezone" ou le fuseau horaire = Pacific Daylight Time, c'est donc aux États-Unis où l'heure est GMT-7 (le "-0700").
Le "Delivered-To" confirme le destinataire final qui doit en principe être identique au "**To:**" vu plus haut. **Si ce n'est pas le cas, il pourrait aussi y avoir problème**. Cette ligne est la dernière de l'en-tête d'un message.

Voilà pour ce qui est de l'étude de l'en-tête d'une newsletter légitime qui n'est pas un SPAM et où tout semble OK.

Le prochain article de la série essayera de montrer l'analyse d'un vrai SPAM; comme déjà dit dans le 1er article, les impatients peuvent télécharger l'en-tête de ce SPAM depuis la page du premier article, le fichier s'appelle "**En-tete-SPAM-Colruyt-recu-sur-Gmail.txt**"

# Li P’ti Fouineu vous salue bien !

Ouvrir En-tête Newsletter Li P’ti Fouineu Gmail dans un nouvel onglet ou le télécharger:

Télécharger