lxml

formaction attr allowing javascript in Cleaner() if safe_attrs_only is set

Bug #1888153 reported by Kevin Chung on 2020-07-19

This bug affects 1 person 256

Affects	Status	Importance	Assigned to	Milestone
lxml	Fix Released	Medium	Unassigned	lxml 4.6.3

Bug Description

```
Python : sys.version_info(major=3, minor=8, micro=3, releaselevel='final',
serial=0)
lxml.etree : (4, 5, 2, 0)
libxml used : (2, 9, 10)
libxml compiled : (2, 9, 10)
libxslt used : (1, 1, 34)
libxslt compiled : (1, 1, 34)

The following script creates a form with a button with formaction which
still allows XSS through when clicking the button.

```
from lxml.html.clean import Cleaner
cleaner = Cleaner(
    forms=False,
    safe_attrs_only=False,
)
cleaner.clean_html("""<form id="test"></form><button form="test"
formaction="javascript:alert(1)">X</button>""")
```

However, this same kind of idea doesn't apply to action on the form which
is somewhat equivalent:

```
In [1]: cleaner.clean_html("""<form id="test"></form><button form="test"
formaction="javascript:alert(1)">X</button>""")
Out[1]: '<div><form id="test"></form><button form="test" formaction=
"javascript:alert(1)">X</button></div>'

In [2]: cleaner.clean_html("""<form id="test" action="javascript:alert(1)
"></form><button form="test" type="submit">X</button>""")
Out[2]: '<div><form id="test" action=""></form><button form="test"
type="submit">X</button></div>'
```

safe_attrs_only is an unsafe setting to disable but it seems to respect
the javascript setting so I would argue that formaction should be added to
the list of attributes that are removed by the javascript setting.
```

CVE References

2021-28957

Kevin Chung (kchung) wrote on 2020-07-19: #1

I believe the issue is that lxml isn't treating a button as a "link" and not running formaction through `doc.rewrite_links()`

https://github.com/lxml/lxml/blob/34aa8896f99f93a43f3c61fc66beb459ce163acd/src/lxml/html/clean.py#L300

From there iterlinks() is called which doesn't see a button as a link.

I'm not sure if you would want to modify iterlinks() or add a special case for button in the clean function.

Kevin Chung (kchung) wrote on 2020-07-19: #2

A similar issue exists with:

```
<input type="submit" formaction="javascript:alert(1)">
```

scoder (scoder) wrote on 2020-07-29: #3

Thanks for the report. This seems worth fixing. If "formaction" is not currently looked at for links, then it should be.

PR welcome.

Changed in lxml:
importance:Undecided → Medium
status:New → Confirmed

Kevin Chung (kchung) wrote on 2021-03-20: #4

PR created: https://github.com/lxml/lxml/pull/316
Also requested a CVE from Mitre.

Kevin Chung (kchung) wrote on 2021-03-21: #5

This has been allocated CVE-2021-28957.

scoder (scoder) on 2021-03-21

Changed in lxml:
milestone:none → 4.6.3

Report a bug

This report contains **Public Security** information

Everyone can see this security related information.

You are not directly subscribed to this bug's notifications.

Edit bug mail

Other bug subscribers

Subscribe someone else

Notified of all changes

Kevin Chung

May be notified

Esha Wang
Gary Zhao
Henning Janßen
KeithSloan
Ravikant
scoder

status:Confirmed → Fix Committed

scoder (scoder) on 2021-03-21

information type:Private Security → Public Security

scoder (scoder) on 2021-03-21

Changed in lxml:
status:Fix Committed → Fix Released

See full activity log

To post a comment you must log in.