

The invention claimed is:

1. A computer implemented method of identifying an element within a target web page as a changed web element, comprising the step of:
 - 5 identifying within a pre-region of the target web page a pre-region longest matching segment of the pre-region that matches a pre-region segment of a pre-region of a source web page, where the source web page is the target web page before the element was changed;
 - identifying within a post-region of the target web page a post-region longest
 - 10 matching segment of the post-region that matches a post-region segment of the post-region of the source web page;
 - determining the pre-region longest matching segment is within a pre-defined distance of the post-region longest matching segment;
 - identifying the element of the target web page as the changed web element that is
 - 15 between the pre-region longest matching segment and the post-region longest matching segment in response to the pre-region longest matching segment being determined to be within a pre-defined distance of the post-region longest matching segment.
- 20 2. The method of claim 1, comprising the steps of:
 - before the step of determining, determining the pre-region longest matching segment is not within a pre-defined distance of the post-region longest matching segment;

in response to determining the pre-region longest matching segment is not within a pre-defined distance of the post-region longest matching segment,

reducing a size of the pre-region of the source web page by excluding the pre-region longest matching segment from the pre-region of the source web page;

reducing a size of the pre-region of the target web page by excluding the pre-region longest matching segment from the pre-region of the target web page;

reducing the size of the post-region of the source web page by excluding the post-region longest matching segment from the post-region of the source web page;

reducing the size of the post-region of the target web page by excluding the post-region longest matching segment from the post-region of the target web page;

repeating at least the steps of:

identifying within the pre-region of the target web page the pre-region longest matching segment; and,

identifying within a post-region of the target web page the post-region longest matching segment.

3. The method of claim 1, wherein the step of determining comprises the step of proceeding, starting with the pre-region longest matching segment not within a pre-defined distance of the post-region longest matching segment, to reduce a size of the pre-region of the source web page and of the pre-region of the target web page by
5 excluding the pre-region longest matching segment from pre-region of the source web page and the pre-region of the target web page and to reduce a size of the post-region of the source web page and of the post-region of the target web page by excluding the post-region longest matching segment from the post-region of the source web page and the post-region of the target web page, with the reduced pre-
10 regions and the reduced post-regions acting as an initial value for a next iteration, until the pre-region longest matching segment is within a pre-defined distance of the post-region longest matching segment.
4. The method of claim 1, wherein the pre-region of the source web page is an area of
15 the source web page before the element in the source web page, the post-region of the source web page is an area of the source web page after the element in the source web page, the pre-region of the target web page is an area of the target web page before an original location of the element, the post-region of the target web page is an area of the source web page after the original location of the element.
- 20 5. The method of claim 1, comprising the step of position encoding the source web page and position encoding the target web page, wherein the pre-region of the source web page is an area of the source web page before an element position encoded index of the element in the source web page, the post-region of the source

web page is an area of the source web page after the element position encoded index, the pre-region of the target web page is an area of the target web page before the element position encoded index, the post-region of the target web page is an area of the target web page after the element position encoded index.

5

6. The method of claim 1, comprising the step of type encoding the source web page and the target web page, and wherein

the pre-region longest matching segment is a pre-region longest matching segment of encoded values and

10

the post-region longest matching segment is a longest matching segment of encoded values, and

the step of identifying within a pre-region of the target web page a pre-region longest matching segment is further defined in that the encoded values of the pre-region of the target web page are compared to the encoded values of the pre-region of the source web page to identify the pre-region longest matching segment of encoded values of the pre-region that matches a pre-region segment of encoded values of the pre-region of the source web page; and,

15

the step of identifying within a post-region of the target web page a post-region longest matching segment is further defined in that the encoded values of the post-region of the target web page are compared to the encoded values of the post-region of the source web page to identify a post-region longest matching segment of encoded values of the post-region of the target web

20

page that matches a post-region segment of encoded values of the post-region of the source web page.

7. The method of claim 6, wherein the step of type encoding is further defined in that opening tags are assigned a first encoding value, closing tags are assigned a second encoded value, and tags which do not have a closing tag by default are assigned a third encoded value, and wherein the step of type encoding results in a sequence of encoded values for the source web page corresponding to a sequence of tags within the source web page, and a sequence of encoded values for the target web page corresponding to a sequence of tags within the target web page.

8. The method of claim 7, wherein the step of type encoding is further defined in that the target web page comprises HTML code and the source web page comprises HTML code.

9. The method of claim 6, comprising the steps of tokenizing the source web page by creating an ordered list of tags of the source web page, and tokenizing the target web page by creating an ordered list of tags of the target web page.

10. The method of claim 1, comprising the steps of, at least before the step of determining:
identifying a repeated sequence range within the source web page containing repeated sequences of tags,

reducing a size of the pre-region or a size of the post-region to within the repeated sequence range.

11. The method of claim 10, comprising the steps of:

identifying an instance range of an instance of a repeated sequence within the repeated sequence range;

reducing a size of the pre-region or a size of the post-region within the instance range.

12. The method of claim 1, comprising a step of parsing the source web page and the

target web page to pair unpaired tags in each of the source web page and the target web page.

13. The method of claim 1, wherein the step of identifying the element of the target web page as the changed web element comprises the step of identifying a location of the

element of the target web page within the target web page.

14. A method of identifying a sought element within a target web page that has been changed, comprising the step of:

comparing a one or more locators known to be associated with the sought

element to a one or more locators of a one or more elements of the target web page to identify a one or more locators of the one or more locators of the one or more elements of the target web page that uniquely identify an element within the target web page;

identifying the web element of the one or more elements of the target web page that is associated with a locator that matches a locator of the one or more locators known to be associated with the sought element and that uniquely identifies the sought element within the target web page.

5

15. The method of claim 14, where the step of identifying is further defined in that the locator comprises an attribute of an html tag.

16. The method of claim 14, wherein the step identifying is further defined in that the
10 locator comprises an Xpath of an html tag.

17. The method of claim 14, wherein the step identifying is further defined in that the locator comprises a CSS selector associated with an html tag.

15 18. The method of claim 14, wherein the step of identifying is further defined in that the locator comprises an image associated with an html tag.

19. The method of claim 18, wherein in the step of identifying is further defined in that
match is found where a perceptual hash value of the image is within a pre-defined
20 threshold of at least one perceptual hash value of an image known to be associated with the sought element.

20. A method of identifying a sought element within a target web page as a changed web element, comprising the step of:

comparing a one or more locators known to be associated with the sought element to a one or more locators of a one or more elements of the target web page to identify a one or more locators of the one or more locators of the one or more elements of the target web page that uniquely identify an element within the target web page;

determining that the one or more elements of the target web page are not associated with a locator that uniquely identifies the sought element within the target web page and matches a locator of the one or more locators known to be associated with the sought element;

in response to said determining,

identifying within a pre-region of the target web page a pre-region longest matching segment of the pre-region that matches a pre-region segment of a pre-region of a source web page, where the source web page is the target web page before the sought element was changed;

identifying within a post-region of the target web page a post-region longest matching segment of the post-region that matches a post-region segment of the post-region of the source web page;

determining the pre-region longest matching segment is within a pre-defined distance of the post-region longest matching segment;

identifying an element of the target web page as the changed web element that is
between the pre-region longest matching segment and the post-region
longest matching segment in response to the pre-region longest matching
segment being determined to be within a pre-defined distance of the post-
5 region longest matching segment.

21. The method of claim 20, wherein the step of determining the pre-region longest
matching segment is within a pre-defined distance of the post-region longest
matching segment, comprises the step of proceeding, starting with the pre-region
10 longest matching segment is not within a pre-defined distance of the post-region
longest matching segment, to reduce a size of the pre-region of the source web page
and of the pre-region of the target web page by excluding the pre-region longest
matching segment from the pre-region of the source web page and the pre-region of
the target web page and to reduce a size of the post-region of the source web page
15 and of the post-region of the target web page by excluding the post-region longest
matching segment from the post-region of the source web page and the post-region
of the target web page, with the reduced pre-regions and the reduced post-regions
acting as an initial value for a next iteration, until the pre-region longest matching
segment is within a pre-defined distance of the post-region longest matching
20 segment.

22. A web element rediscovery system, comprising:

a memory;

a processor in communication with the memory;

a pre-region identifying function executable by the processor to identify within a pre-region of a target web page a pre-region longest matching segment of the pre-region that matches a pre-region segment of a pre-region of a source web page, where the source web page is the target web page before a sought element was changed;

a post-region identifying function executable by the processor to identify within a post-region of the target web page a post-region longest matching segment of the post-region that matches a post-region segment of the post-region of the source web page;

a determining function executable by the processor to determining the pre-region longest matching segment is within a pre-defined distance of the post-region longest matching segment;

a web element identifying function executable by the processor to identify a web element of the target web page as a changed web element that is between the pre-region longest matching segment and the post-region longest matching segment in response to the pre-region longest matching segment being determined to be within a pre-defined distance of the post-region longest matching segment.

23. The system of claim 22, wherein determining function comprises a reduction function of proceeding, starting with the pre-region longest matching segment is not within a pre-defined distance of the post-region longest matching segment, to reduce

a size of the pre-region of the source web page and of the pre-region of the target web page by excluding the pre-region longest matching segment from the pre-region of the source web page and the pre-region of the target web page and to reduce a size of the post-region of the source web page and of the post-region of the target web page by excluding the post-region longest matching segment from the post-regions, with the reduced pre-regions and the reduced post-regions acting as an initial value for a next iteration, until the pre-region longest matching segment is within a pre-defined distance of the post-region longest matching segment.

24. The system of claim 22, wherein the pre-region of the source web page is an area of the source web page before the sought element in the source web page, the post-region of the source web page is an area of the source web page after the sought element in the source web page, the pre-region of the target web page is an area of the target web page before an original location of the sought element, the post-region of the target web page is an area of the source web page after the original location of the sought element.

25. The system of claim 22, comprising an position encoding function stored on the memory and executable by the processor to position encode the source web page and position encoding the target web page, wherein the pre-region of the source web page is an area of the source web page before an element position encoded index of the sought element in the source web page, the post-region of the source web page is an area of the source web page after the element position encoded index, the pre-region of the target web page is an area of the target web page before the element

position encoded index, the post-region of the target web page is an area of the target web page after the element position encoded index.

26. The system of claim 22, comprising a type encoding function stored on the memory
5 and executable by the processor to type encode the source web page and type
encode the target web page, and wherein
- the pre-region longest matching segment is a pre-region longest matching
segment of encoded values and
- the post-region longest matching segment is a longest matching segment of
10 encoded values, and
- the pre-region identifying function is configured to compare the encoded values
of the pre-region of the target web page to the encoded values of the pre-
region of the source web page to identify the pre-region longest matching
segment of encoded values of the pre-region that matches a pre-region
15 segment of encoded values of the pre-region of the source web page; and,
- the post-region identifying function is configured to compare the encoded values
of the post-region of the target web page to the encoded values of the post-
region of the source web page to identify a post-region longest matching
segment of encoded values of the post-region of the target web page that
20 matches a post-region segment of encoded values of the post-region of the
source web page.

27. The system of claim 26, wherein the type encoding function is further configured to

assign opening tags a first encoding value, closing tags a second encoded value,
and tags which do not have a closing tag by default a third encoded value;
produce a sequence of encoded values for the source web page corresponding to
a sequence of tags within the source web page; and,
5 produce a sequence of encoded values for the target web page corresponding to
a sequence of tags within the target web page.

28. The system of claim 27, wherein the type encoding function is further configured to
encode HTML tags from the target web page comprising HTML code and from the
10 source web page comprising HTML code.

29. The system of claim 26, comprising a tokenization function configured to tokenize
the source web page by creating an ordered list of encoded tags of the source web
page, and to tokenize the target web page by creating an ordered list of encoded tags
15 of the target web page.

30. The system of claim 22, comprising:
a repeated sequence range identifying executable by the processor to identifying a
repeated sequence range within the source web page containing repeated
20 sequences of tags; and,
a region reduction function executable by the processor to reduce a size of the
pre-region or a size the post-region to within the repeated sequence range.

31. The system of claim 30, comprising:

- an instance range identification function executable by the processor to identify
an instance range of an instance of a repeated sequence within the repeated
sequence range; and,
a size reduction function executable by the processor to reduce a size of each of
5 the pre-regions or a size of each of the post-regions within the instance
range.
32. The system of claim 22, comprising a parsing function executable by the processor
to parse the source web page and the target web page to pair unpaired tags in each of
10 the source web page and the target web page.
33. The system of claim 22, wherein the web element identifying function is further
configured to identifying a location of the web element of the target web page within
the target web page.
15
34. The system of claim 22, wherein the pre-region identifying function, the post-region
identifying function, the determining function, and the web element identifying
function are stored in the memory.
- 20 35. A web element locator computer system, comprising:
a memory;
a processor in communication with the memory;

a comparing function executable by the processor to compare a one or more
locators known to be associated with a sought element to a one or more
locators of a one or more elements of a target web page to identify a one or
more locators of the one or more locators of the one or more elements of
5 the target web page that uniquely identify an element within the target web
page; and

an identifying function executable by the processor to identify the web element
of the one or more elements of the target web page that is associated with a
locator that matches a locator of the one or more locators known to be
10 associated with the sought element and that uniquely identifies the sought
element within the target web page.

36. The system of claim 35, wherein the identifying function is further defined in that
the locator comprises an attribute of an html tag.

15

37. The system of claim 35, wherein the identifying function is further defined in that
the locator comprises an Xpath of an html tag.

38. The system of claim 35, wherein the identifying function is further defined in that
20 the locator comprises a CSS selector associated with an html tag.

39. The system of claim 35, wherein the identifying function is further defined in that
the locator comprises an image associated with an html tag.

40. The system of claim 39, wherein the identifying function is further defined in that
match is found where a perceptual hash value of the image is within a pre-defined
threshold of at least one perceptual hash value of an image known to be associated
5 with the sought element.
41. A web element rediscovery computer system, comprising:
- a memory;
 - a processor in communication with the memory;
 - 10 a comparing function executable by the processor to compare a one or more
locators known to be associated with a sought element to a one or more
locators of a one or more elements of a target web page to identify a one or
more locators that uniquely identify an element within the target web page;
and
 - 15 a first determining function executable by the processor to determine that the
one or more elements of the target web page are not associated with a
locator that uniquely identifies the sought element within the target web page
and
matches a locator of the one or more locators known to be associated with
20 the sought element;
 - a pre-region identifying function executable by the processor to identify within a
pre-region of the target web page a pre-region longest matching segment of
the pre-region that matches a pre-region segment of a pre-region of a source

web page, where the source web page is the target web page before the sought element was changed;

a post-region identifying function executable by the processor to identify within a post-region of the target web page a post-region longest matching segment of the post-region that matches a post-region segment of the post-region of the source web page;

a determining second function executable by the processor to determining the pre-region longest matching segment is within a pre-defined distance of the post-region longest matching segment;

a web element identifying function executable by the processor to identify a web element of the target web page as a changed web element that is between the pre-region longest matching segment and the post-region longest matching segment in response to the pre-region longest matching segment being determined to be within a pre-defined distance of the post-region longest matching segment.

42. The system of claim 41, wherein determining function comprises a reduction function executable by the processor to proceed, starting with the pre-region longest matching segment is not within a pre-defined distance of the post-region longest matching segment, to reduce a size of the pre-region of the source web page and of the pre-region of the target web page by excluding the pre-region longest matching segment from the pre-region of the source web page and the pre-region of the target web page and to reduce a size of the post-region of the source web page and of the

post-region of the target web page by excluding the post-region longest matching segment from the post-region of the source web page and the post-region of the target web page, with the reduced pre-regions and the reduced post-regions acting as an initial value for a next iteration, until the pre-region longest matching segment is
5 within a pre-defined distance of the post-region longest matching segment.

43. The system of claim 41, wherein the pre-region of the source web page is an area of the source web page before the sought element in the source web page, the post-region of the source web page is an area of the source web page after the sought
10 element in the source web page, the pre-region of the target web page is an area of the target web page before an original location of the sought element, the post-region of the target web page is an area of the source web page after the original location of the sought element.

15 44. The system of claim 41, comprising an position encoding function stored on the memory and executable by the processor to position encode the source web page and position encoding the target web page, wherein the pre-region of the source web page is an area of the source web page before an element position encoded index of the sought element in the source web page, the post-region of the source web page is
20 an area of the source web page after the element position encoded index, the pre-region of the target web page is an area of the target web page before the element position encoded index, the post-region of the target web page is an area of the target web page after the element position encoded index.