**Encoding**

**What is Encoding?**

Is the process of changing a sequence of characters into a centralised, specialised format

* Commonly used for transferring data
* Bi-directional

Why is this needed?

* To provide a standard format of data
  + Normalisation
* Commonly needed for apps
* Can be used for Unicode as well as ASCII

Base64

Hex

URL

**Base64**

An encoding schemie used for the HTTP protocol

Inside of HTTP it is used to transfer data safely

Invented as a part of the MIME content transfer encoding

**How it works**

* Base64 will divide the bytes of the entire set of data into groups of 6 bits
* Every 6 bits will go into a group
* Every group will be provided a specific character, based on the table below

Graphical user interface, application, Word

Description automatically generated

Because this is represented in 6-bit intervals, Base64 encoded data will be about 33% larger than the original data

Graphical user interface, application, Word

Description automatically generated

It’s important to understand that these must be groups of 6 bits

* If this is not, it will pad the end of the bits with 0s
  + When this happens, there’ll be an ‘=’ at the end of the Base64 string

**Hex** – Hexadecimal

Used to interact with machines at a lower level as well as represent data

* Not exactly an ‘encoding’ method

**How does it work?**

Hex can be identified in a number of ways

* #11 (usually graphic)
* H1
* &11 (usually programming)
* 0x11 (usually memory)

Hex uses a base-16 system

We use a 10-base counting system (10 fingers)

No formula

* Need to look up hex value in a table
* ASCII tables will provide the hex value for a specific character

Example: the capital letter ‘A’ is 0x41 and the lower case ‘a’ is 0x61

**URL**

Not quite ‘encoding’ again

**A URL is required to have only ASCII passed through it** – can’t have kanji in it

Something as simple as a space cannot be interpreted without being encoded by the browser

Converting a special character into its digit form so it can represent it on the backend in a more formatted way

* So that it doesn’t have to worry about a character it doesn’t understand

There is a set value on the ASCII table for all ASCII text

* Look up the ASCII table and you will see another column called DEC (decimal)
  + This column shows the digit representation of the ASCII text
  + The value from the ASCII table will be represented by a %
    - Any encoded character will be prepended with a ‘%’

**Example:**

https//redteamnation.com/thisisatest+test

* + sign must be converted
* Browser converts it to 43
* Prepend ‘%’

Conversion: https//redteamnation.com/thisisatest**%43**test