**EDI**—Electronic Document or Data Interchange—is a rather generic term. Under this broad term, there are 2-3 major **standards** as well as less common ones which specify the details:

* ANSI > [ASC X12](http://www.x12.org/) is the predominant standard in North America
* UN/EDIFACT is the predominant standard outside North America
* GS1 EDI used in global supply chain

It seems that all of the EDI we deal with is of the **X12** standard. The remainder of the document assumes X12.

Beyond standards, there can be different **versions** of the standard. (I have yet to find documentation on what these versions are however.)

An EDI message consists of **segments**. These are separated by a **segment terminator** (defined in the initial interchange segment). Since new lines / line breaks (\r\n) are a common delimiter, it’s easy to think of segments as lines of data. Each segment begins with a **segment identifier** in the form of a 2-3 character alpha-numeric code.

Each segment is divided into **data elements** by means of a **data element delimiter**. This seems to normally be an “**\***” but could in theory be any character. Data elements can potentially be further divided by a **repetition delimiter** (to effectively create arrays of values) and/or a **composite delimiter**. Data elements have an id, and the specific data element may appear in different segments. So for example, data element 188 is a weight unit code which may appear as the 5th element in a B2 segment, but 188 could also appear in other segments. The possible codes for 188 are the same regardless of what segment (though partners may only accept a subset of available codes).

The EDI standard defines three “**envelopes**”: Interchange, Group, and Transaction. Each of these are delimited by starting and ending segments (**ISA**/**IEA**, **GS**/**GE**, and **ST**/**SE** respectively). (These each contain a control number which must match.)

In theory, an EDI string could have multiple interchanges. An **interchange** represents one overall exchange with a “trading partner” (the term typically used to refer to the companies engaged in the communication).

An interchange may have multiple groups. A **group** is a set of one or more transactions of the same type (204s, 214s, et cetera).

Finally, a **transaction** contains the data for one message / “document”. The first element in the ST segment (after the segment identifier) is the **transaction / document type**. The following types are the ones we deal with primarily:

* **204** - Motor Carrier Load Tender
* **210** - Motor Carrier Freight Details and Invoice (billing)
* **214** - Transportation Carrier Shipment Status Message
* *920 - Loss or Damage Claim - General Commodities*
* **990** - Accept or Decline the 204 Shipment
* 997 - Generic EDI received acknowledgment

(see <https://en.wikipedia.org/wiki/X12_Document_List> for more)

Though in theory an EDI file could have multiple items in each envelope, in our experience it seems typical for each file to contain only one transaction (which still must be wrapped in an interchange and group). Exceptions to this are that we send multiple 214 and 210 transactions in a single EDI. (214s and 210s are separate, so no more than one group is sent in a single EDI.)

Envelope segments are reasonably well defined and reasonably consistent between trading partners (with the caveat that they may be using different *versions* of EDI). However, the contents of a transaction are less consistent. There are many common segments for a 204 message for instance. But different trading partners may include or exclude some, and the values they place in these segments may vary. This is where documentation for the specific implementation is important.

I’ll further explain the complexity of EDI by comparing it to the better known CSV file. With a CSV, each line is effectively separated into columns. These columns are the same for every line. So for instance in a hypothetical CSV, the first column might always be a date, the second the shipper’s name, etc. With EDI, imagine that the columns are different *for each line* (technically, each segment type). Furthermore, perhaps a third of the values in EDI are codes which alter the meaning of other “columns” (values). A good example of this is the name segment. What name is this? The name segment is rather generic. The meaning of the name is determined by a code. EDI codes are usually 2-3 characters and are more often than not as non-self-evident as “X1” or “93”. In other words, just because I have a name segment doesn’t mean I know the meaning of the values in every “column”. This effectively means that there are multiple types of name segments. This is just an example of one segment; most segments rely on one or more codes like this. And again, some partners use and/or require certain codes while refusing others. All of this greatly multiplies the potential variations of an EDI message.