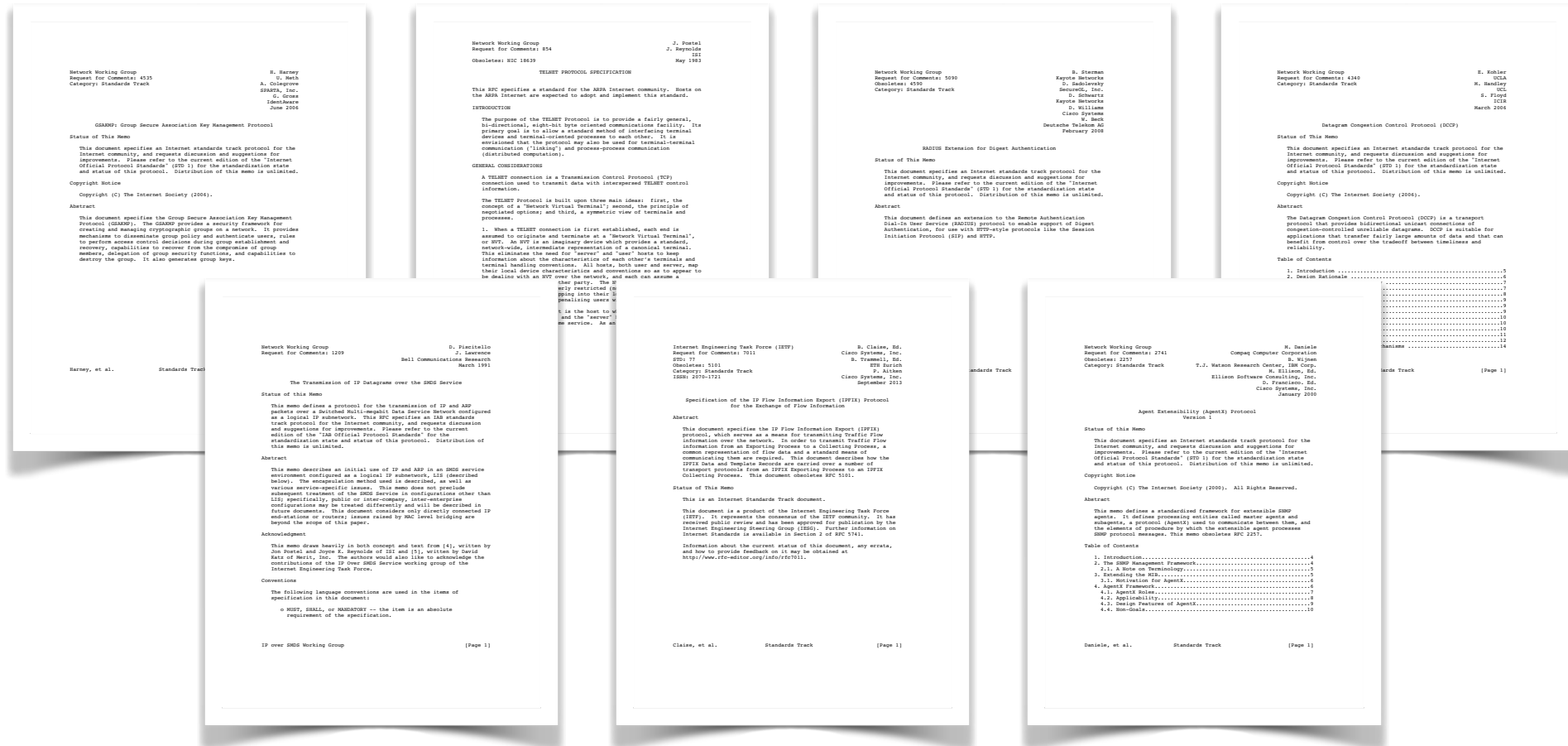# Machine-readable documents and their tools
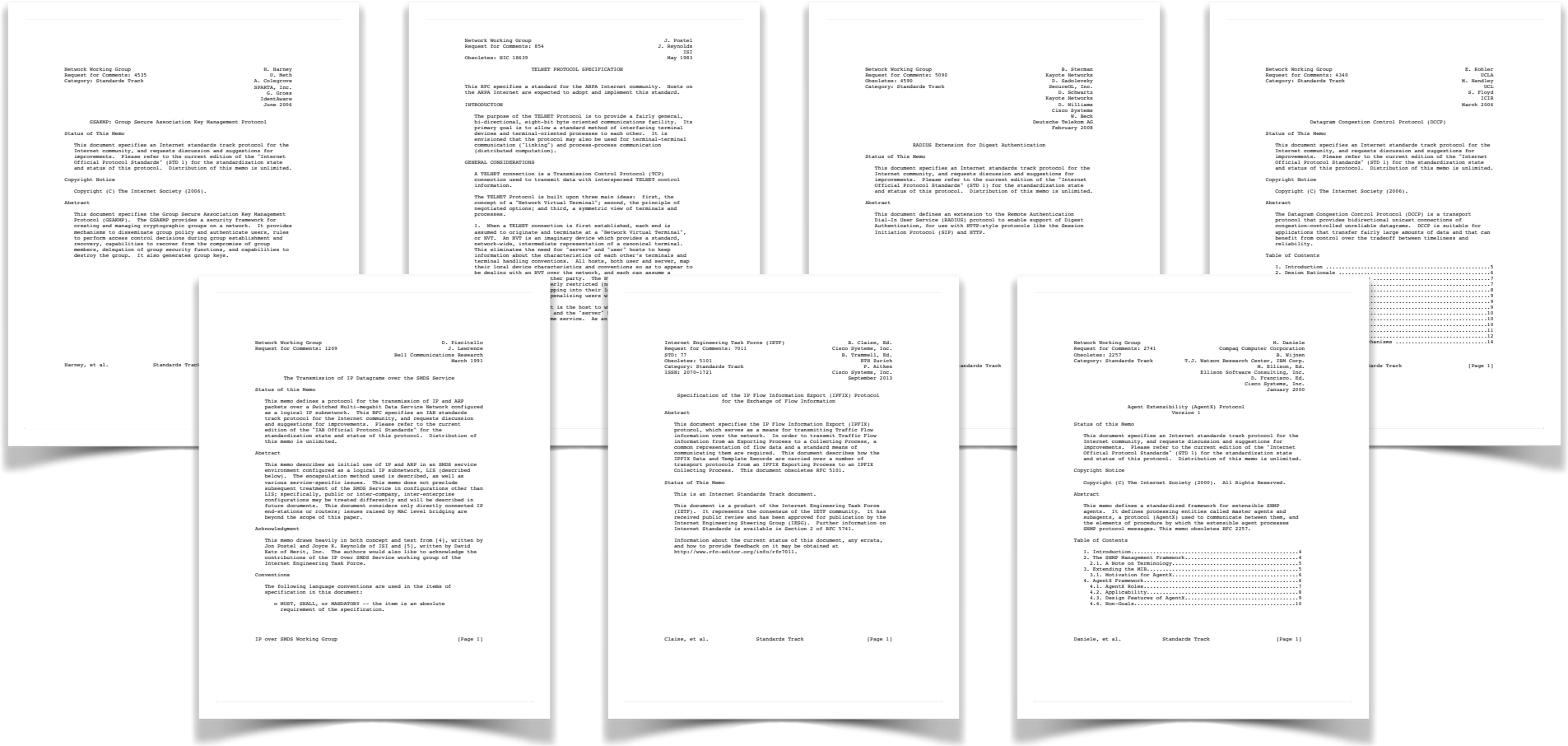
Stephen McQuistin
Marc Petit-Huguenin

IETF 112 Hackathon

Machine-readable specifications and their tools

Machine-readable specifications and their tools

Computerate Specifying

Machine-readable specifications and their tools

Computerate Specifying

ABNF

YANG

TLS presentation language

Augmented Packet Header Diagrams

QUIC packet notation

4

Machine-readable specifications and their tools

NLP

Structured language parsers

Computerate Specifying

ABNF

YANG

TLS presentation language

Augmented Packet Header Diagrams

QUIC packet notation

Machine-readable specifications and their tools
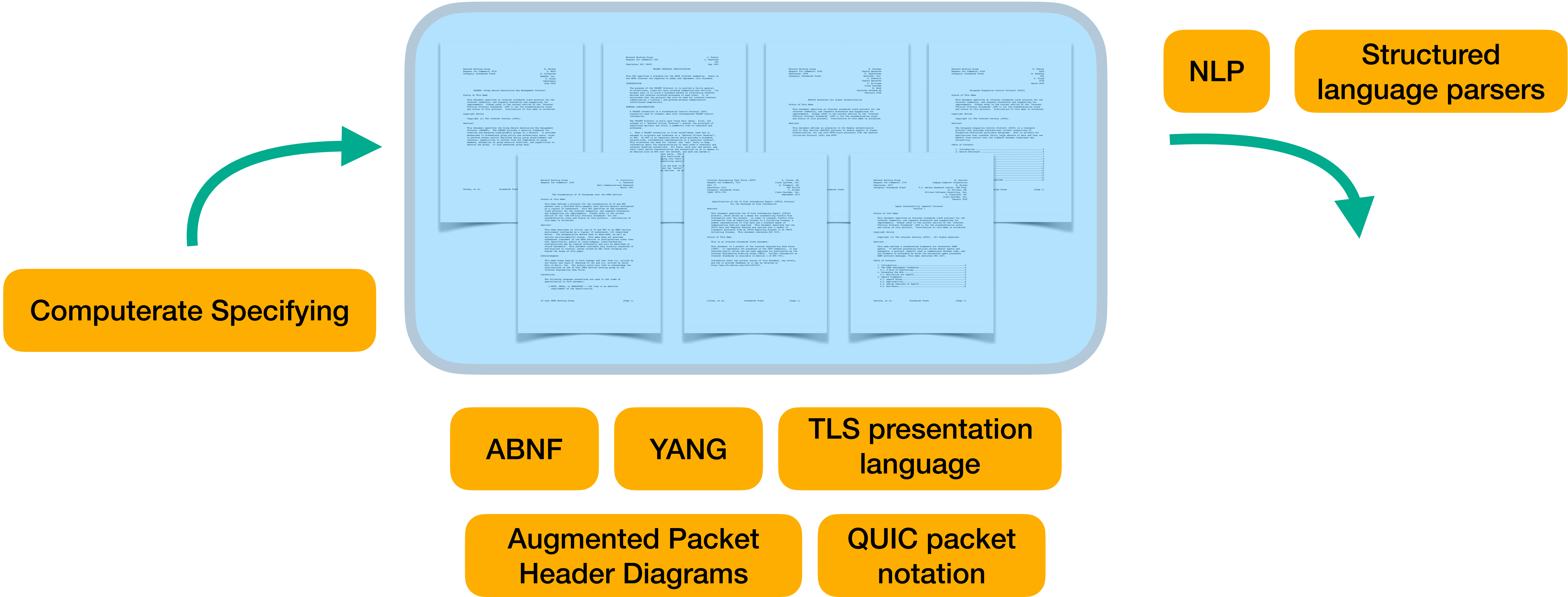
Computerate Specifying

NLP
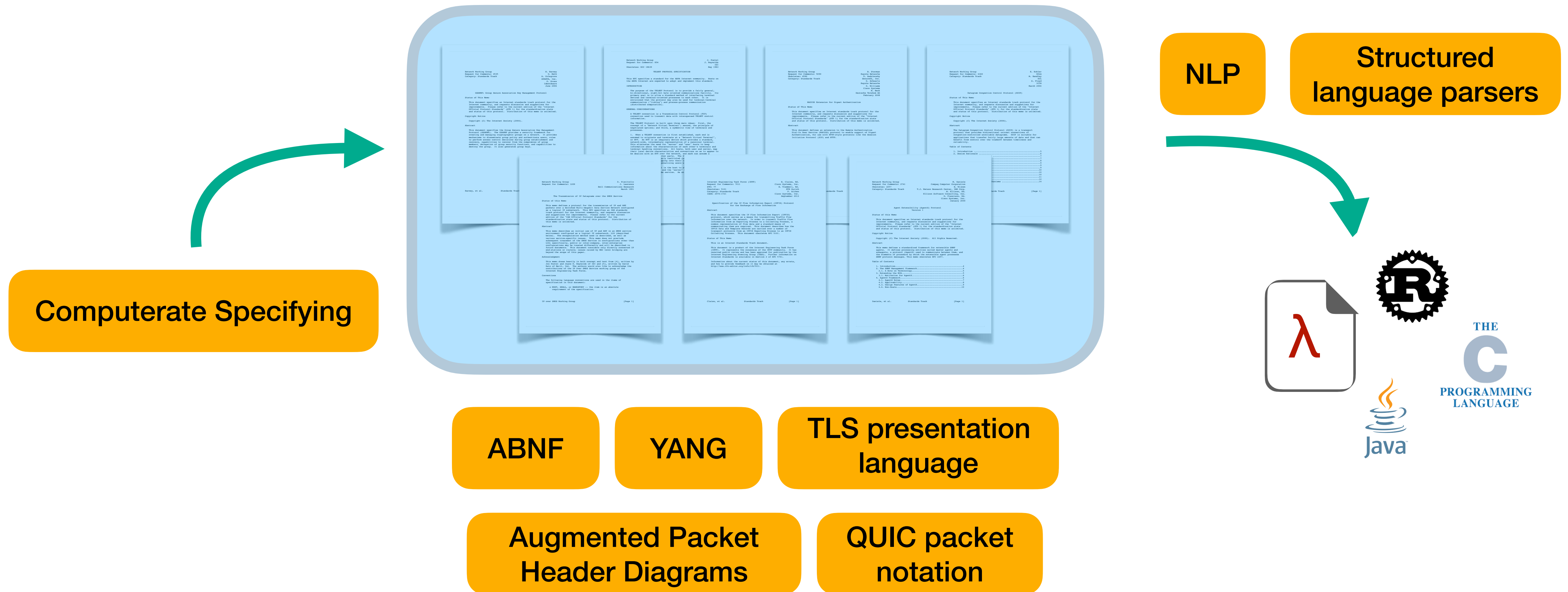
Structured language parsers

ABNF

YANG

TLS presentation language

Augmented Packet Header Diagrams

QUIC packet notation

Machine-readable specifications and their tools

Computerate Specifying

NLP

Structured language parsers

ABNF

YANG

TLS presentation language

Augmented Packet Header Diagrams

QUIC packet notation

Augmented Packet Header Diagrams

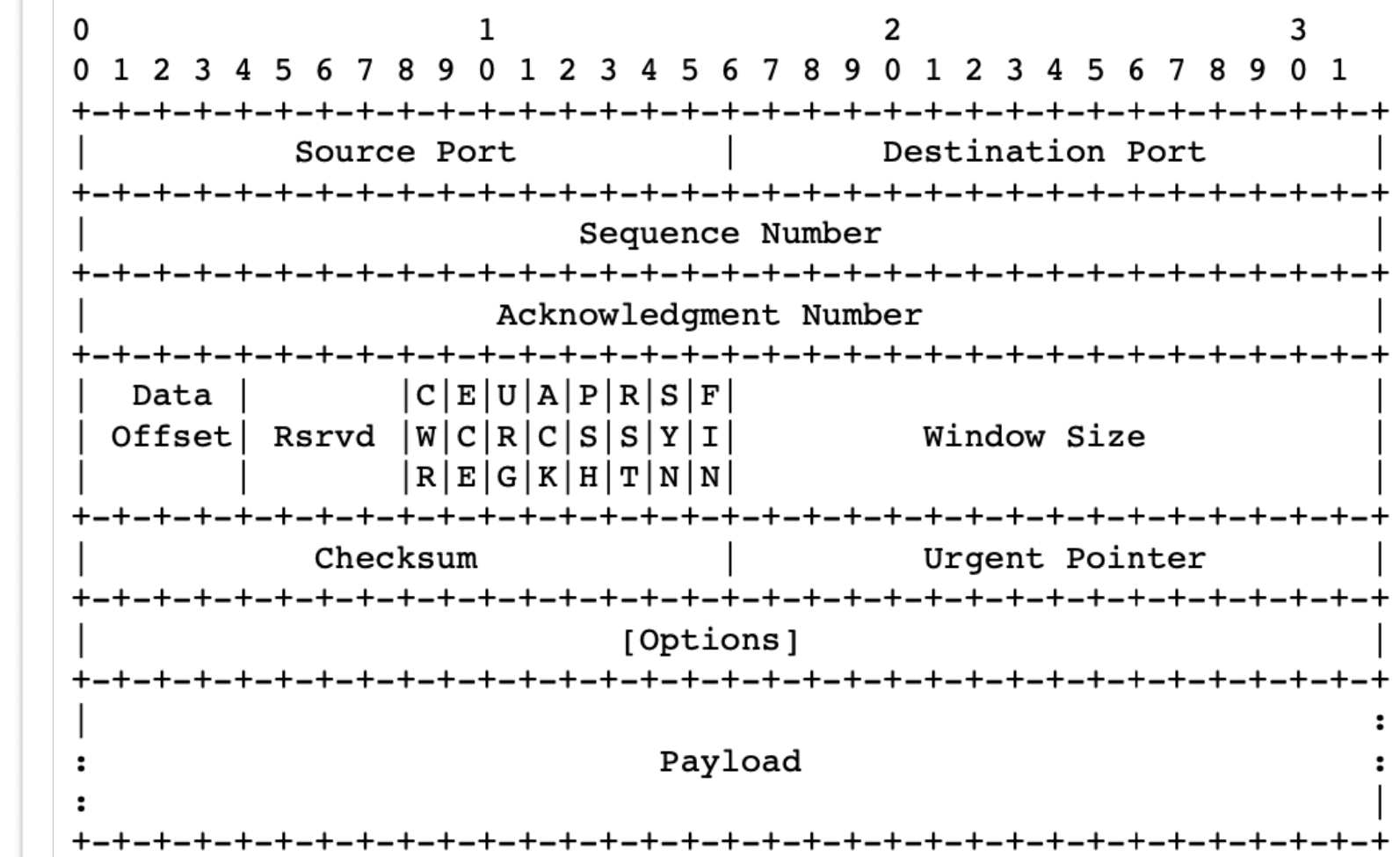- The format of packet header diagrams can be regularised with minimal change

- The format remains extremely close to that in common use, easing adoption

- It balances structure and uniformity, needed for machine parsing, with the flexibility needed for practical use

- Prototype tooling that supports this input format

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Source Port          |       Destination Port        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sequence Number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Acknowledgment Number                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Data |       |C|E|U|A|P|R|S|F|                               |
| Offset| Rsrvd |W|C|R|C|S|S|Y|I|          Window Size           |
|       |       |R|E|G|K|H|T|N|N|                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Checksum            |         Urgent Pointer        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            [Options]                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                              :
:                            Payload                            :
:                                                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

where:

Source Port: 16 bits.  The source port number.

Destination Port: 16 bits.  The destination port number.

Sequence Number: 32 bits.  The sequence number of the first data
   octet in this segment (except when the SYN flag is set).  If SYN
   is set the sequence number is the initial sequence number (ISN)
   and the first data octet is ISN+1.

Acknowledgment Number: 32 bits.  If the ACK control bit is set, this
   field contains the value of the next sequence number the sender of
   the segment is expecting to receive.  Once a connection is
   established, this is always sent.

Data Offset (DOffset): 4 bits.  The number of 32 bit words in the TCP
   Header.  This indicates where the data begins.  The TCP header
```
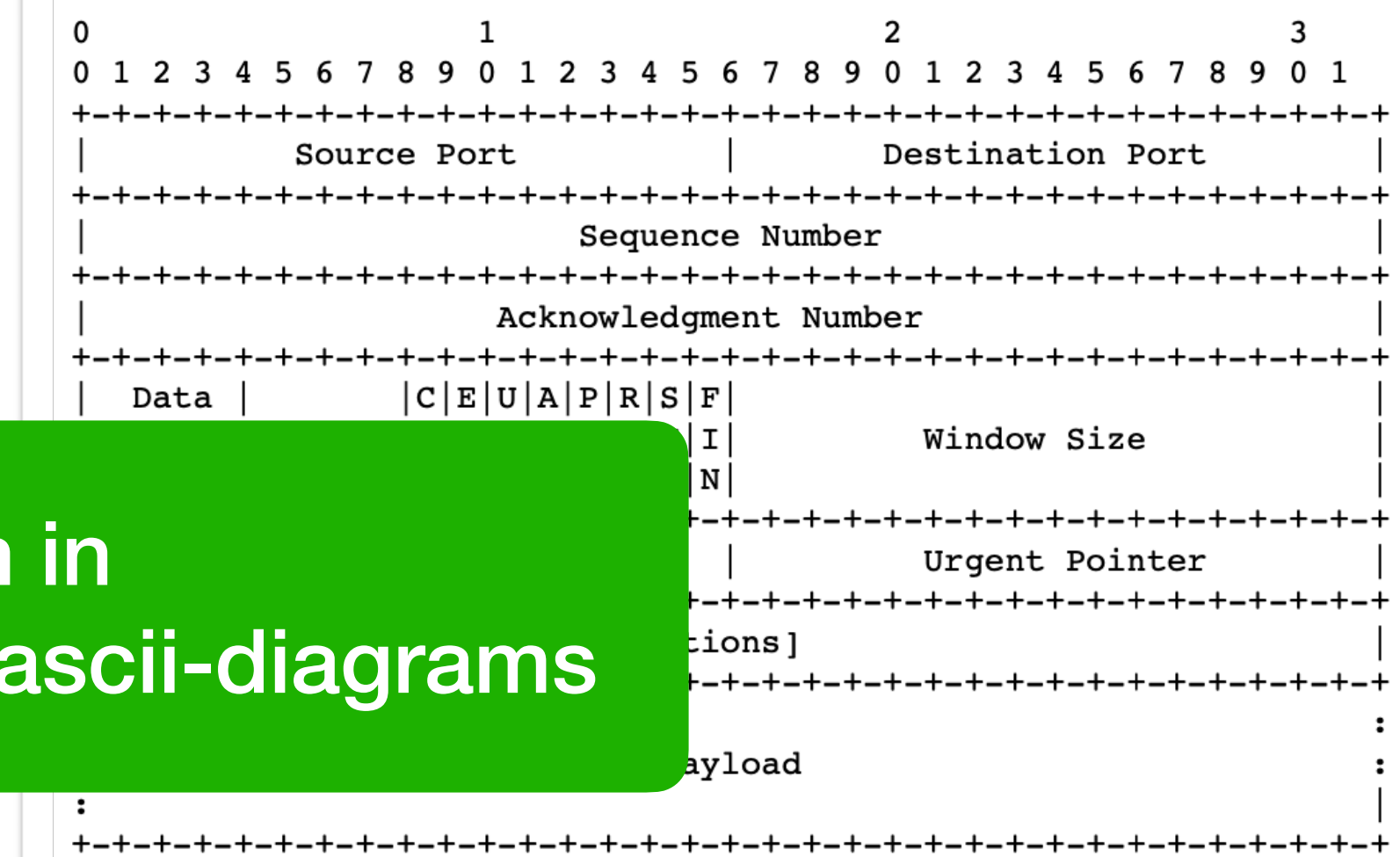
8

Augmented Packet Header Diagrams

- The format of packet header diagrams can be regularised with minimal change

- The format remains ex⬚⬚⬚ in common use, easin⬚⬚⬚

- It balances structure and uniformity, needed for machine parsing, with the flexibility needed for practical use

- Prototype tooling that supports this input format


```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Source Port          |       Destination Port        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sequence Number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Acknowledgment Number                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data  |       |C|E|U|A|P|R|S|F|                               |
|       |       | | |I|           Window Size                  |
|       |       | | |N|                                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               |        Urgent Pointer         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         ⬚⬚⬚ions]                           :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                           ⬚⬚ayload                            :
:                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

where:

Source Port: 16 bits.  The source port number.

Destination Port: 16 bits.  The destination port number.

Sequence Number: 32 bits.  The sequence number of the first data
   octet in this segment (except when the SYN flag is set).  If SYN
   is set the sequence number is the initial sequence number (ISN)
   and the first data octet is ISN+1.

Acknowledgment Number: 32 bits.  If the ACK control bit is set, this
   field contains the value of the next sequence number the sender of
   the segment is expecting to receive.  Once a connection is
   established, this is always sent.

Data Offset (DOffset): 4 bits.  The number of 32 bit words in the TCP
   Header.  This indicates where the data begins.  The TCP header
```
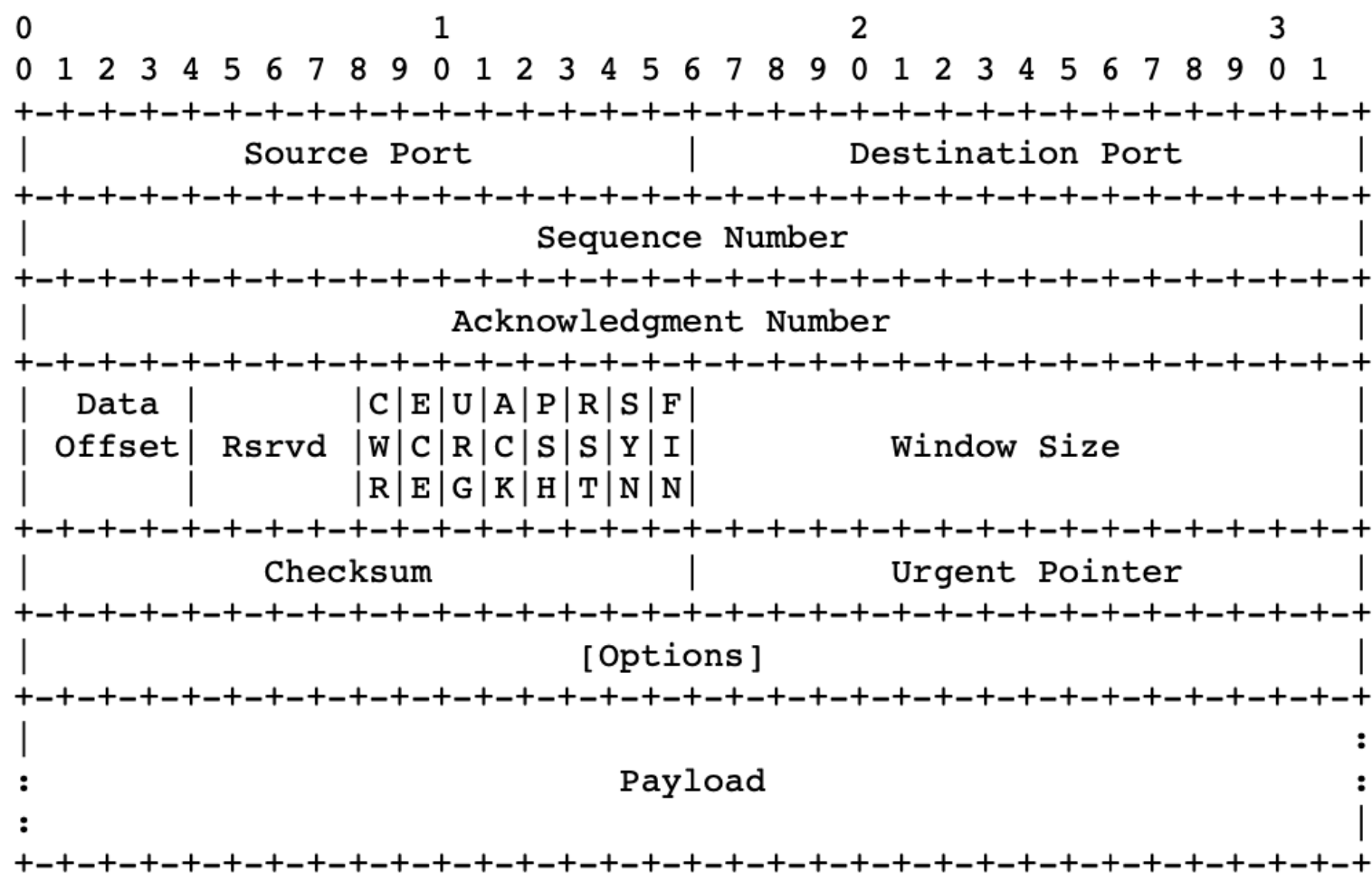
More information in
draft-mcquistin-augmented-ascii-diagrams

Progress this week

- Good discussions about the particular language we use, and how this could be improved for *human* readability

- How do we mark structured text so that it isn't edited beyond the allowed phrases?

- Started work to improve upon our prototype tooling, adding flexibility and robustness

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Source Port          |       Destination Port        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sequence Number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Acknowledgment Number                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Data |       |C|E|U|A|P|R|S|F|                               |
| Offset| Rsrvd |W|C|R|C|S|S|Y|I|          Window Size          |
|       |       |R|E|G|K|H|T|N|N|                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Checksum            |         Urgent Pointer        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           [Options]                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               :
:                           Payload                             :
:                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

where:

Source Port: 16 bits.  The source port number.

Destination Port: 16 bits.  The destination port number.

Sequence Number: 32 bits.  The sequence number of the first data
   octet in this segment (except when the SYN flag is set).  If SYN
   is set the sequence number is the initial sequence number (ISN)
   and the first data octet is ISN+1.

Acknowledgment Number: 32 bits.  If the ACK control bit is set, this
   field contains the value of the next sequence number the sender of
   the segment is expecting to receive.  Once a connection is
   established, this is always sent.

Data Offset (DOffset): 4 bits.  The number of 32 bit words in the TCP
```