

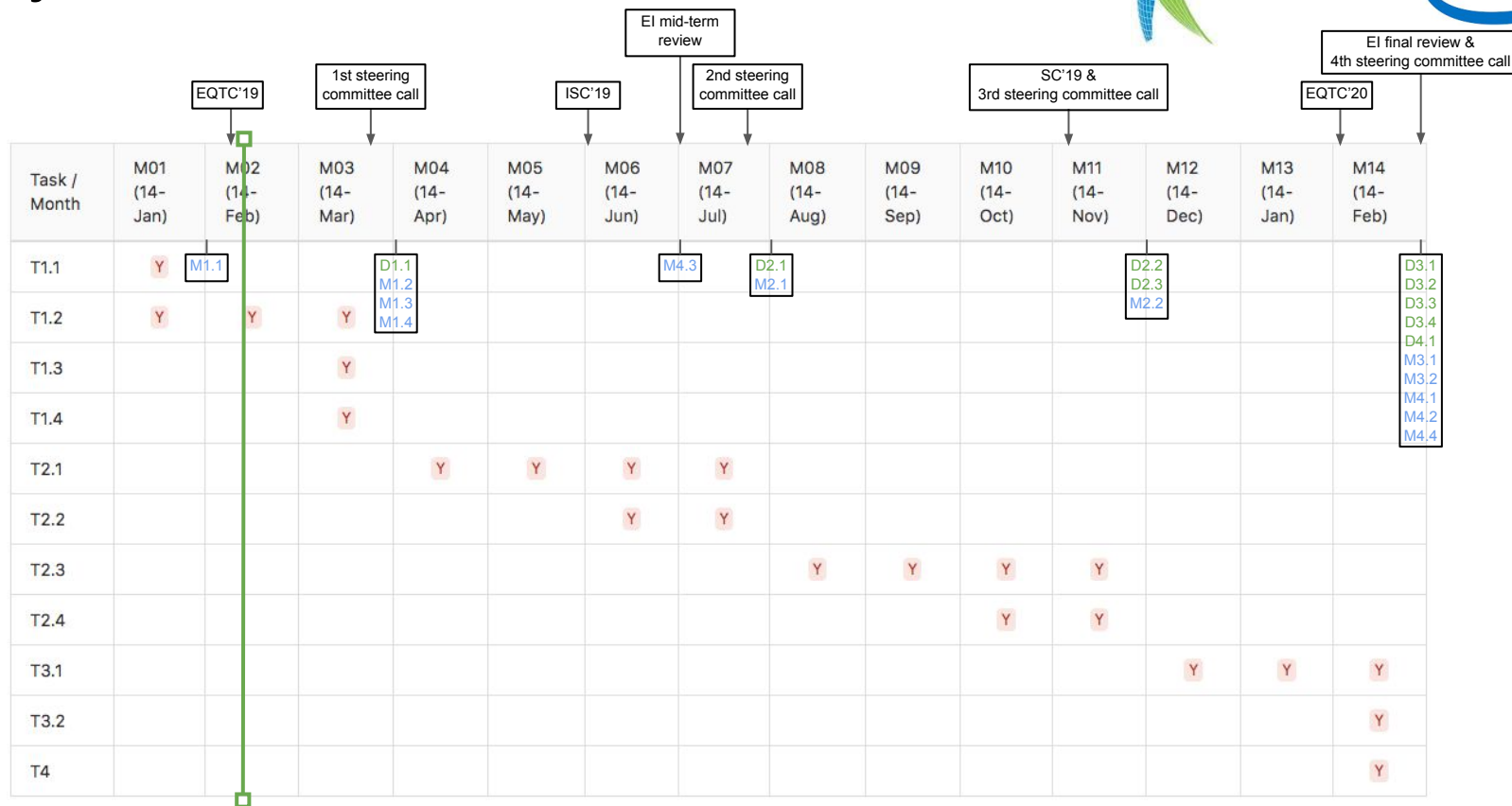


# Quantum Natural Language Processing

All-hands meeting  
25-Feb-2019

ICHEC, Dublin

# Project timeline



**Status of prior 3 weeks** 🟢🟡🔴

- Quantum nearest neighbour algorithm 🟡
  - Surveyed four versions
  - Implementation challenges in two versions
    - Oracle operations required
    - Extension to multi-dimensional vector space is not straightforward
  - Third & fourth versions for binary vectors; extensible for meaning space vectors (real)
  - Test implementations of third/fourth version
- Preparation of representative corpus 🟢
  - 10x-100x words; tagging; meaning space analysis
- qHiPSTER on Kay 🔴
  - qHiPSTER\_Installation\_and\_Usage\_Issues.md
  - Working for smaller problem sizes with AVX512
- ISC 2019 🟢
  - Draft of project poster submitted on 20-Feb
  - Liaised with LRZ and Intel DE for tutorial

**Plans for next 3 weeks**

- Quantum nearest neighbour algorithm
  - Preliminary implementation for binary vectors
  - Formalise mapping for meaning space vectors
    - Pre-computation required for state initialisation
    - Gates/operators to be implemented in qHiPSTER
    - Quantum circuit to be implemented
- CSC sentence similarity algorithm
  - Define mapping NLP operations to Dirac notations, Q operations
- ISC 2019
  - Presentation/presence at Intel Booth
    - ICHEC to work with FB
- SC 2019
  - Proposal for tutorial session (16-Apr)
    - ICHEC will circulate details mid-March
  - Paper submission (02-Apr, 10-Apr)
    - TBD

**High-level dashboard**

- Project members
  - Intel -- FB, JK, BQ
  - ICHEC -- MD, VK, LOR, PW
- Deliverables/Milestones
  - M1.1 (M01; 14-Feb)
  - D1.1 (M03; 15-Apr)
  - M1.2, M1.3, M1.4 (M03)

**Issues / Bottlenecks**

- qHiPSTER scalability using BigMPI
  - MPI communication error for larger problem sizes; Details shared in qHiPSTER\_Installation\_and\_Usage\_Issues.md in repository

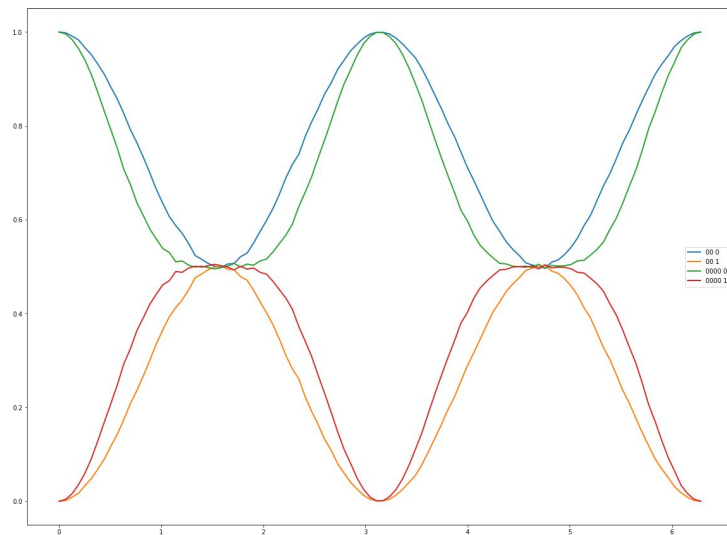
**Action needed from ILE staff**

- Feedback on qHiPSTER scalability issues with BigMPI
- Press release approval

# Data encoding strategies



- Analogue vector encoding:  $n$ -length floating point vector  $\rightarrow \log_2(n)$  qubits
  - Encoding data requires: solution of NL LSQ system; MKL support for this?
  - Potential gates required: Controlled  $R_y$ : available in qHiPSTER;
  - Entanglement-based comparison results



# Data encoding strategies



- Vectors same  $\rightarrow 1$ , vectors orthog  $\rightarrow 0.5$
- Issues: quick development and evaluation needed; Python bindings for qHiPSTER?

