Language Technology:
Research and Development

R&D Projects – From Proposal to Implementation

Sara Stymne

Uppsala University
Department of Linguistics and Philology
sara.stymne@lingfil.uu.se

R&D Projects

Research and development is often organized into projects

- ▶ Time-limited
- ▶ One-time effort
- ▶ Specific goals
- ▶ Separate budget
- ▶ Separate organization

Projects vary in scope and size

- ▶ Term paper (1 person, 240 hours)
- ▶ EU FP project (15–20 sites, 6–10 MEUR)

Life Cycle of a (Funded) Project

Pre-grant activities:

- ▶ Explore research opportunities
- ▶ Write and submit research proposal
- ▶ Sign research contract

Post-grant activities:

- ▶ Start up: mobilize project resources
- ▶ Manage research activities
- ▶ Close down: report project outcome

Research Funding in Sweden – Government

Direct grants to universities (16 BSEK)

▶ Basic funding for research and graduate education

Research councils and agencies (9.5 BSEK)

▶ Swedish Research Council (VR)

▶ Environment, Agricultural Sciences and Spatial Planning (Formas)

▶ Health, Working Life and Welfare (FORTE)

▶ Innovation Systems (VINNOVA)

Research Funding in Sweden – Other

Public research foundations (4 BSEK)

▶ Bank of Sweden Tercentenary Foundation (RJ)

▶ Swedish Foundation for Strategic Research (SSF)

▶ Knowledge Foundation (KK)

Other Swedish non-profit research foundations (3.5 BSEK)

▶ Knut and Alice Wallenberg Foundation

▶ Swedish Cancer Society

Industrial and business research (100 BSEK, 1 BSEK to universities)

Research Funding in the European Union

Horizon 2020 (2014–2020) (80 BEUR)

1. Excellent science
   - European Research Council (13 BEUR)
   - Future and Emerging Technologies (2.5 BEUR)
   - Marie Curie (6 BEUR)
   - European Research Infrastructures (2.5 BEUR)

2. Industrial leadership (20 BEUR)
   - Industrial technologies, risk finance, innovation in SMEs

3. Societal challenges (28 BEUR)
   - Health, food security, energy, transport, climate, security, . . .

Research Funding in the European Union

Horizon Europe (2021–2027) (100 BEUR)

1. Excellent science
   - European Research Council
   - Marie Curie
   - Research Infrastructures

2. Global challenges and European Industrial Competitiveness

3. Innovative Europe

4. Widening participation and strengthening the European Research Areas

Planning a Project

Research question
- ► What are you going to find out?

Previous work
- ► What do we know already?

Approach
- ► How are you going to find out?

Time plan
- ► When are you going to do what?

Research Questions

A research question is a clear, focused, relevant, and interesting question around which we center our research.

- ▶ Clear: Stated concisely using precise terminology
- ▶ Focused: Not too broad (nor too limited)
- ▶ Relevant: Has a bearing on the research topic
- ▶ Interesting: Provides substantial new information if answered

Above all, questions should be researchable.

Research Questions – Good or Bad?

1. Does global warming affect parsing accuracy?
2. Do multiword expressions affect parsing accuracy?
3. How do multiword expressions affect parsing accuracy?
4. How do light verbs affect parsing accuracy?
5. Are light verbs harder to parse than other verbs?
6. What can we do to improve parsing accuracy for light verbs?
7. Can valency info improve parsing accuracy for light verbs?
8. What is the F_1 of the Berkeley parser on light verbs in PTB?
9. How do you install the Berkeley parser on your laptop?

Research Questions – Your own

1. Try to formulate a tentative research question for your project
   (not necessarily the final one)
2. Discuss with some peers

Previous Work

- ▶ Why?
  - ▶ Scientific research should result in new knowledge
  - ▶ We make progress by building on previous results

    "If I have seen further it is by standing on the shoulders of giants." (Newton)

- ▶ How?
  - ▶ Find literature using a focused search (internet, library)
  - ▶ Manage the literature in a database (references, annotations)
  - ▶ Use the literature in your own work (context, motivation)

- ▶ Tips and tricks:
  - ▶ Start with handbook or survey articles if available
  - ▶ Use the snowball method (references of references)
  - ▶ Use citation statistics (with caution)

Useful Resources and Tools

- The ACL Anthology (https://aclanthology.info)
  - Repository of (currently) over 58,000 scientific papers
  - Searchable using general or specialized search engines
  - Full text articles (PDF) and bibliographic references (BibTeX)
- University library (http://ub.uu.se)
  - Databases (Web of Science, ScienceDirect, Google Scholar)
  - Journals and books (printed and electronic)
- Reference management software
  - BibTeX (used with LaTeX) – de facto standard in LT research
  - (EndNote (widely used with MS Word) – basic version free)

Approach

- ▶ Theory:
  - ▶ Theoretical framework (concepts, definitions)
  - ▶ Refinement of research questions
- ▶ Method:
  - ▶ How can we answer the research question?
  - ▶ What theoretical results do we need (and how to prove them)?
  - ▶ What empirical data do we need (and how to get them)?
  - ▶ How do we analyze the results?
- ▶ Approach has to fit research questions

Approach – An Example

► Research question:
  ► Are light verbs harder to parse than other verbs?
► Theory:
  ► Parsing framework
  ► Definitions (light verbs, other verbs)
► Method:
  ► Data selection (sampling, annotation)
  ► Evaluation metrics for verb-specific accuracy
  ► Experimental setup (systems, data splits, tuning)
  ► Hypothesis testing (statistical tests)
  ► Error analysis (quantitative, qualitative)

Designing Experiments

- ▶ Identify variables:
  - ▶ Independent variable – manipulated by researcher
  - ▶ Dependent variable – measured by researcher
  - ▶ Control variable – kept constant by researcher
- ▶ Select data:
  - ▶ Avoid bias in data selection
  - ▶ Distinguish training, development and test data
- ▶ Design measurements and analysis:
  - ▶ Use appropriate metrics
  - ▶ Use a reasonable baseline
  - ▶ Repeat measurements if needed
  - ▶ Use appropriate statistical tests
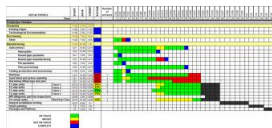 - ▶ Check for alternative explanations

Time Plan

▶ Devise a project plan:
  1. Identify tasks and subtasks
  2. Identify dependencies between tasks
  3. Order tasks and make time estimates
  4. Set up milestones and contingency plans

▶ Words of wisdom:
  1. Keep it simple!
  2. Keep deadlines deadly!
  3. Multiply all time estimates by three!

Time Plan

▶ Devise a project plan:
1. Identify tasks and subtasks
2. Identify dependencies between tasks
3. Order tasks and make time estimates
4. Set up milestones and contingency plans

▶ Words of wisdom:
1. Keep it simple!
2. Keep deadlines deadly!
3. Multiply all time estimates by three!

▶ Your time plan can be quite basic

Writing a Project Proposal

▶ Scientific part:
1. Introduction (research questions, motivation)
2. Background (previous work, current issues)
3. Project description (theory, method, time plan)
4. Expected results (significance)

▶ Administrative part:
1. Organization and management
2. Deliverables and milestones
3. Budget
4. Participants' qualifications (CV, publications)

VR Guidelines (Research Plan)

► Purpose and goals
  ► Present the overall purpose and specific goals of the project.
► State of the art
  ► Summarize previous research with key references.
► Significance and scientific novelty
  ► Describe short-term and long-term significance of the project.
► Preliminary and previous results
  ► Describe pilot studies that support the feasibility of the project.
► Project description
  ► Give a summary of the project describing its theory, methods, time plan, and implementation.

VR Guidelines (Research Plan)

- ▶ Purpose and goals
  - ▶ Present the overall purpose and specific goals of the project.
- ▶ State of the art
  - ▶ Summarize previous research with key references.
- ▶ Significance and scientific novelty
  - ▶ Describe short-term and long-term significance of the project.
- ▶ Preliminary and previous results
  - ▶ Describe pilot studies that support the feasibility of the project.
- ▶ Project description
  - ▶ Give a summary of the project describing its theory, methods, time plan, and implementation.

Hints for Writing a Project Proposal

- ▶ Content:
  - ▶ State research questions clearly and concisely from the start
  - ▶ Use background to motivate research questions
  - ▶ Be as specific as possible about theory and method
  - ▶ Avoid unnecessary details – convey the big picture
  - ▶ Make sure to follow the guidelines closely
- ▶ Form:
  - ▶ Use exact terminology (but avoid obscure technical jargon)
  - ▶ Use correct grammar and spelling (but keep it simple)
  - ▶ Use concrete examples to exemplify abstract concepts
  - ▶ Use graphical illustrations when appropriate
  - ▶ Respect page limits with reasonable margins and font sizes

Research Proposal Evaluations

▶ Reserach proposals are typically evaluated by broad panels (aided by experts)
▶ Typical panels for language technology projects:
  ▶ Linguistics and Languages
  ▶ Computer Science
▶ Important that proposals are accesible to non-experts
▶ Need to balance:
  ▶ General introduction and motivation of problem
  ▶ (Technical) details

Implementing the Project

▶ Start up: mobilize project resources
  ▶ Hire researchers and other personnel
  ▶ Acquire equipment, software, data, literature
▶ Manage research activities
  ▶ Implement project plan
  ▶ Revise plans if necessary
▶ Close down: report project outcome
  ▶ Dissemination of results (publications)
  ▶ Report to funding agency

Your Project Proposals

- Maximum 3 pages plus references (and not much shorter)
  - Strict limit!
- Structure (following the VR Guidlines):
  - Purpose and goals (max 0.5 page)
  - State of the art (max 1 page)
  - Project description (min 1 page)
    - Theory and method
    - Time plan and implementation
  - References
- Use the LaTeX template available!

Your Projects

- ▶ Time to start thinking seriously about a project
  - ▶ Proposals due **October 2** (3 pages)
  - ▶ Presentations October 7 (8 minutes with slides, plus questions+discussion)
- ▶ Contact your group leaders if you need advice