Working Across Boundaries
Public and Private Domains, Part 3
A follow-up survey

Flavio Bonifacio, Metis Ricerche srl, Turin - e-mail: flavio.bonifacio@metis-ricerche.it

Previous history of SI.MUL.STORY project

2014 – Toronto Poster – Part 1: *Tender application for "Preliminary, technical, feasibility studies for industrial research and experimental development. The fund will cover 40% of expenditures. Our project has overcome the first evaluation level.*

2016 – Bergen Poster – Part 2: *The project has been accepted as eligible for funding and was realized: a trial version of SI.MUL.STORY (a Nesstar based multimedia system for analytic storytelling) has been created.*

2017 – Lausanne Poster – Part 3: *Following a period of market actions this poster presents the results of a follow-up survey that we realized in Turin in November in order to understand why the proposal to build a unified repository of mixed type of socioeconomic surveys and administrative data is so few appealing (the story will continue until the happy ending).*

The research

We concentrated the analysis on the real attractiveness of data repositories for a selected population. We adopted a mixed strategy for data gathering: first, we organized a seminar on the topic *Archiving, disseminating and data reusing, where we are? A backward glance in order to go forward;* simultaneously we launched a *cawi* survey on the same topic in order to understand behaviors and attitudes of potential target population extracted from the seminar participants and other persons working in academic, administrative and private field.

Seminar returns

Need of data repositories has been well recognized in Turin almost since 1975. The quantitative data work has produced interesting new data base on specific topics. For example

- WHIP (*Work Histories Italian Panel*), a database of individual work histories, based on administrative archives
- TLS (Turin Longitudinal Study), a database containing census data, information from population register and from death certificates.
- CSI Piemonte collects data from all the aspect of the Piemonte region social and economic life since the late seventies
- Outside of Piedmont UNIDATA BICOCCA DATA ARCHIVES located at the University of Milan inherits the long activity cumulated at the ADPSS archives since the late seventies and early eighties
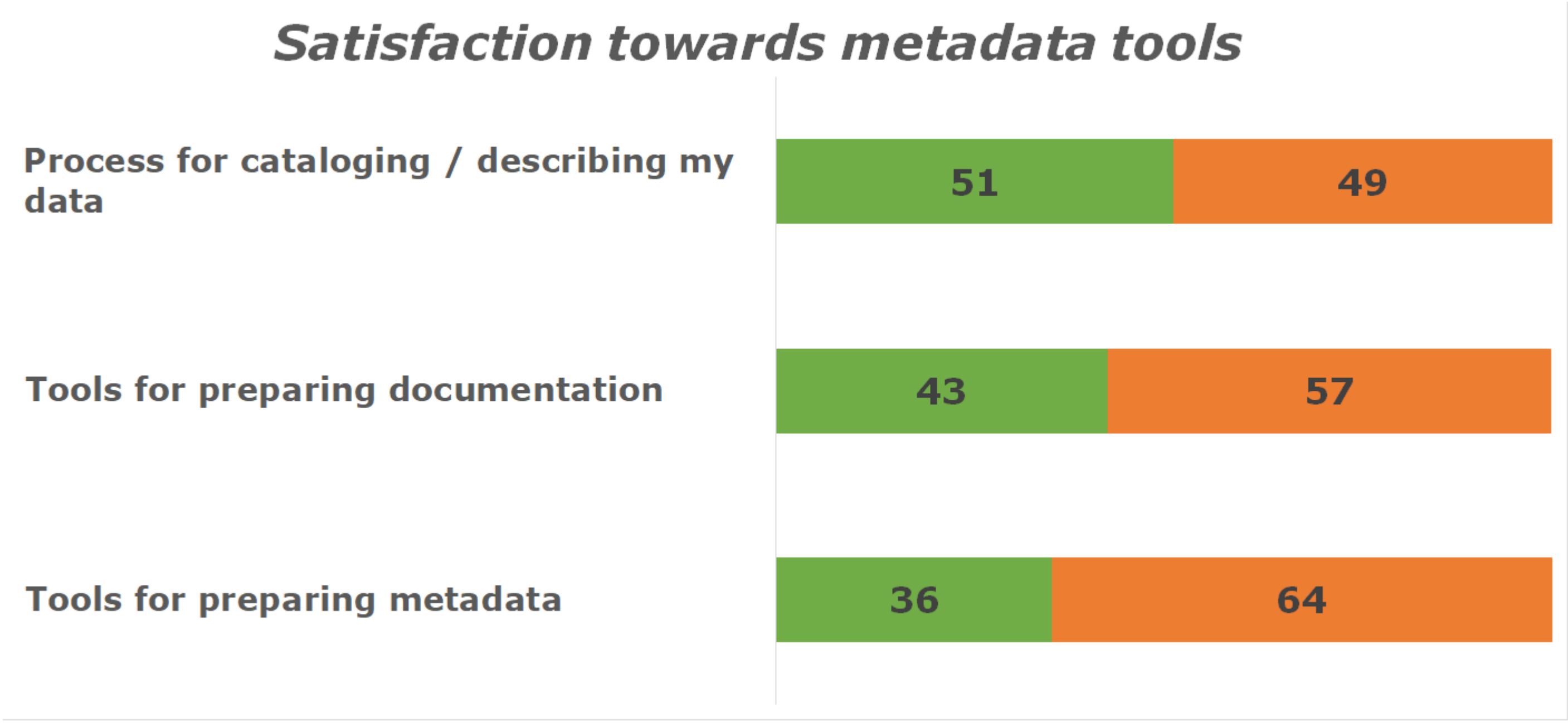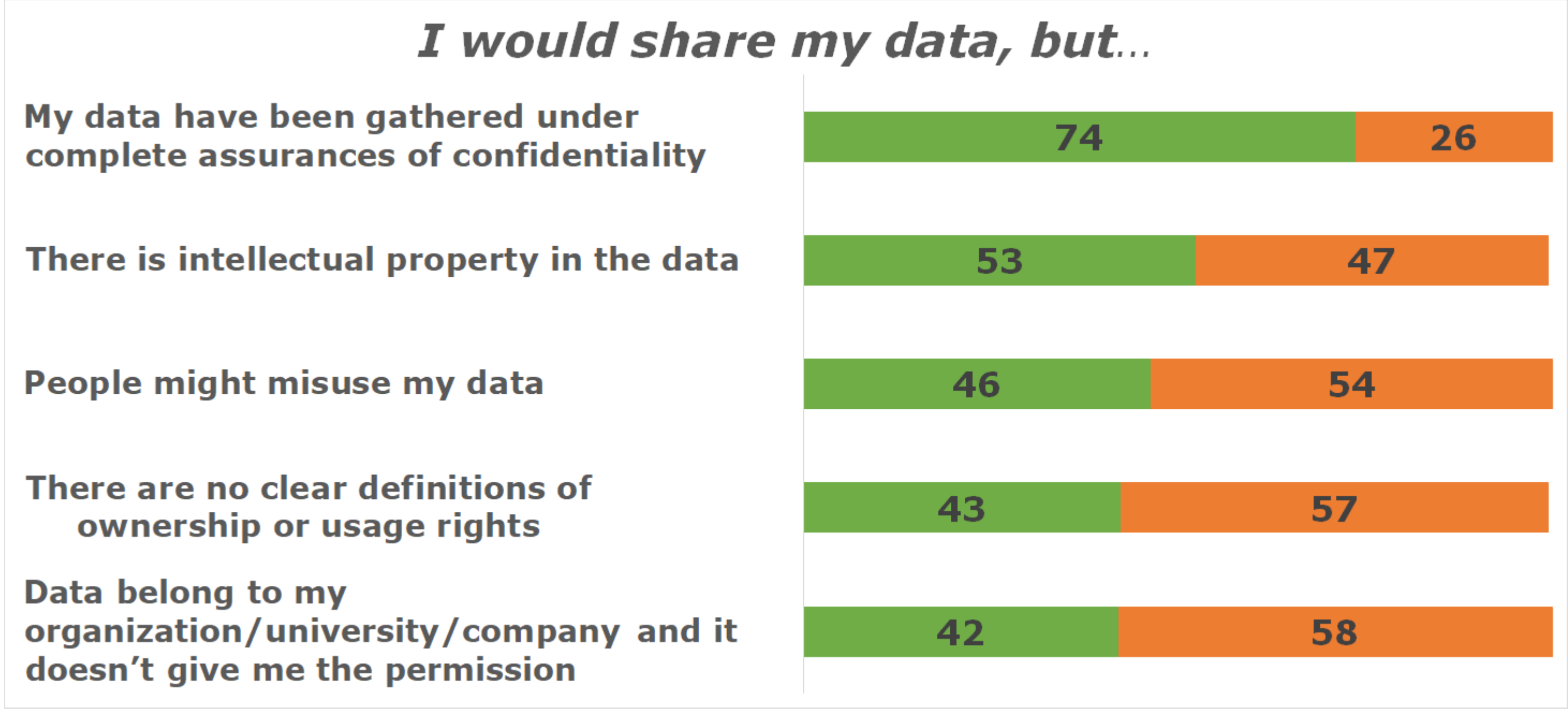
Remarks: *data exist, somehow and somewhere data are conserved and maintained. Data exist but scattered through different producer. What it seems to be missing in our Region is a well organized ecodatasystem addressing them in a uniform, centralized, free and structured way.*
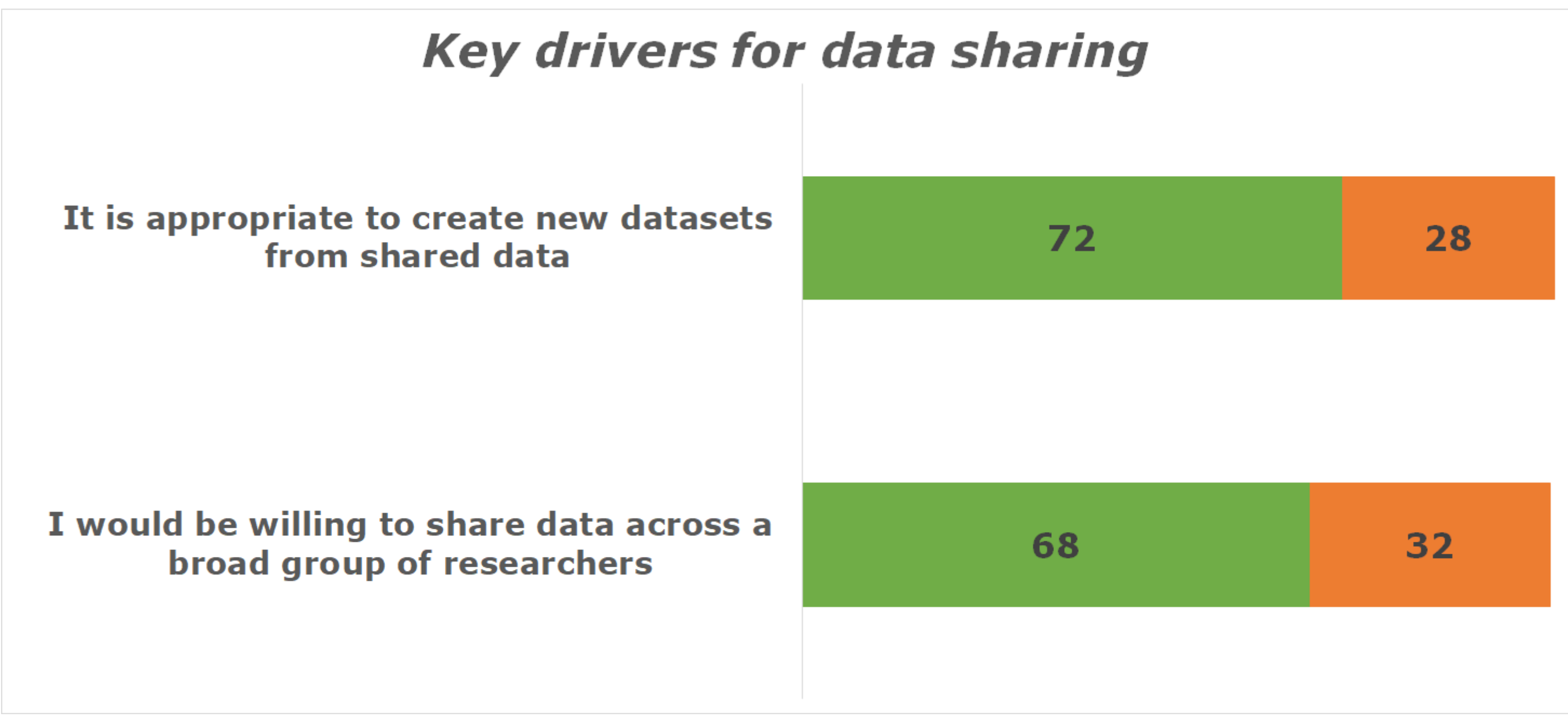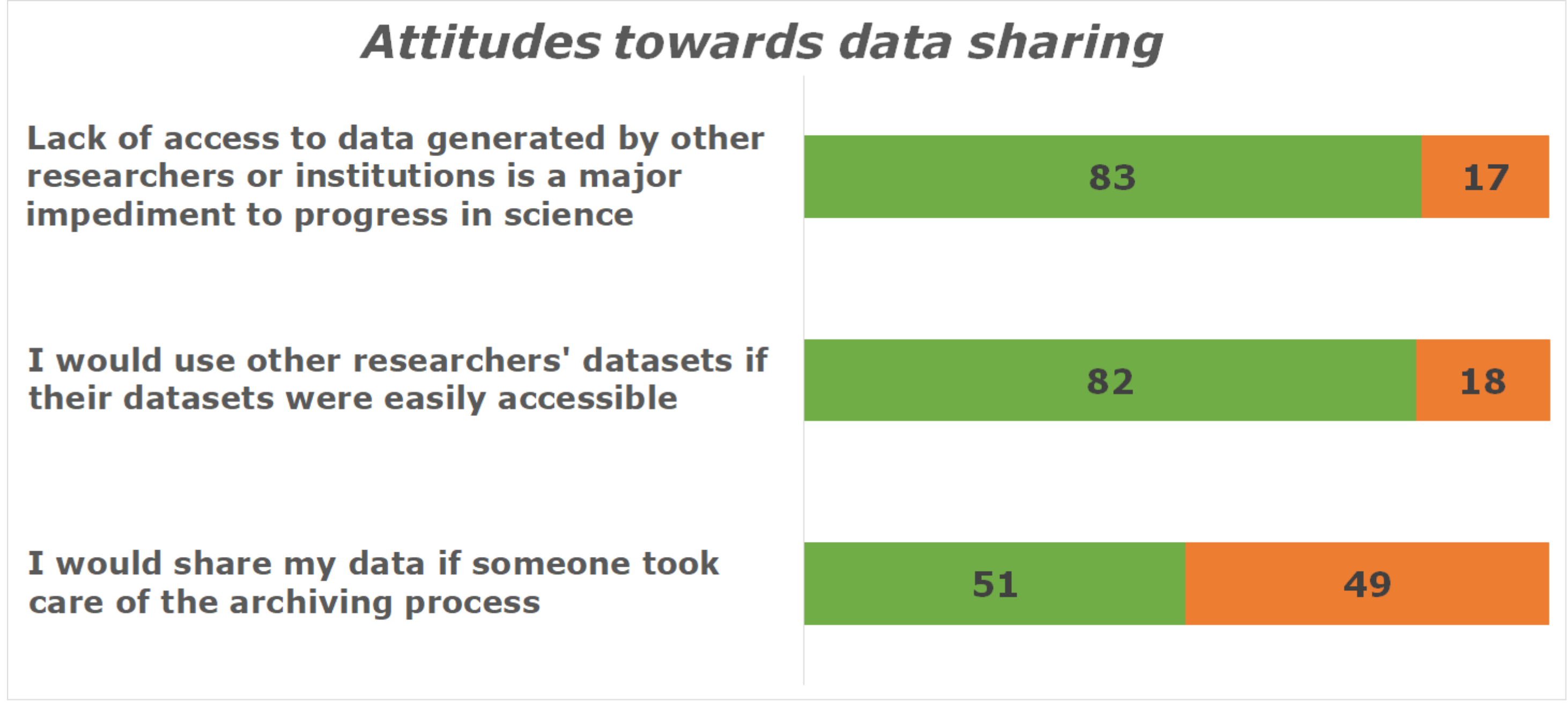
Survey results

The survey questions have been taken from the e-mail exchange that took place on the web among Iassist members answering the question: *I would share my data, but…*in October. The mail focus was on the completion of the above reported sentence excerpt. Graphics report (in green) the aggregate answers *Agree strongly* and *Agree somewhat* for each Item.

For the majority the fundamental obstacle for sharing data is that…*after all data are reserved and private property.*

The problem of documentation is recognized, but … *proper working tools are scarce.*

I would share my data, but…

	Green	Orange
My data have been gathered under complete assurances of confidentiality	74	26
There is intellectual property in the data	53	47
People might misuse my data	46	54
There are no clear definitions of ownership or usage rights	43	57
Data belong to my organization/university/company and it doesn't give me the permission	42	58

Satisfaction towards metadata tools

	Green	Orange
Process for cataloging / describing my data	51	49
Tools for preparing documentation	43	57
Tools for preparing metadata	36	64

People know that data sharing is a fundamental issue but…*sharing my data would be feasible if someone other would archive the data.*

Team working and data file preparation from shared data are *key drivers to stimulate favorable attitudes for data sharing.*

Attitudes towards data sharing

	Green	Orange
Lack of access to data generated by other researchers or institutions is a major impediment to progress in science	83	17
I would use other researchers' datasets if their datasets were easily accessible	82	18
I would share my data if someone took care of the archiving process	51	49

Key drivers for data sharing

	Green	Orange
It is appropriate to create new datasets from shared data	72	28
I would be willing to share data across a broad group of researchers	68	32

If I would be a dealer

If I would be a system data archive dealer, Piemonte Region and Turin would be for me a sort of wonderland:

FACTS	DEALER OPPORTUNITY
Data exist in structured although scattered archives	Collect the scattered data in some place
Sharing data is recognized to be absolutely necessary	Promote data sharing
Some specialized person working on archive is stated to be necessary too	Build a well trained archivists working team
People say that specialized tools are required.	Provide them whit the proper working tools

Open the store, begin to sell… et voila les jeux sont faites. But the real world is different. What is getting wrong?I would say just the vision of interviewed persons. Because …after all data are reserved and private property. Often this point of view is shared also by the Institutions too. *How to change this view?*

A virtuous circle

The survey shows us two main streets: team working and working to build data files on shared data. The process would be simplified because:

- team working forces people to share their data
- data files creation on shared data creates new data ready for reuse
- new data generates new research stream for new working teams

Finally

SI.MUL.STORY will be successful as a sort of data hub only by means of a market campaign showing in practice the advantages of data sharing. Fighting the most widespread solipsistic vision and looking for a key institution that may support this action.