START PAGE

MARIE SKLODOWSKA-CURIE ACTIONS

Individual Fellowships (IF)
Call: H2020-MSCA-IF-2015

PART B

“OSEGA”

This proposal is to be evaluated as:

[Standard EF]

TABLE OF CONTENTS

0 List of Participants

Participants	Legal Entity Short Name	Academic	Non-academic	Country	Dept. / Division / Laboratory	Supervisor	Role of Partner Organisation
Beneficiary							
- NAME							
Partner Organisation							
- NAME							

Data for non-academic beneficiaries

Name	Location of research premises (city / country)	Type of R&D activities	No. of fulltime employees	No. of employees in R&D	Website	Annual turnover (approx. in Euro)	Enterprise status (Yes/No)	SME status (Yes/No)

Note that:

- Any inter-relationship between different participating institutions or individuals (e.g. family ties, shared premises or facilities, joint ownership, financial interest, overlapping staff or directors, etc.) must be declared and justified in this part of the proposal;
- The information in the table for non-academic beneficiaries must be based on current data, not projections;
- The data provided relating to the capacity of the participating institutions will be subject to verification during the Grant Agreement preparation phase.

1 Excellence

Please note that the principles of the European Charter for Researchers and Code of Conduct for the Recruitment of Researchers promoting open recruitment and attractive working conditions are expected to be endorsed and applied by all beneficiaries in the Marie Sklodowska-Curie actions.

1.1 Quality, innovative aspects and credibility of the research (including inter/multidisciplinary aspects)

You should develop your proposal according to the following lines:

- Introduction, state-of-the-art, objectives and overview of the action
- Research methodology and approach: highlight the type of research and innovation activities proposed
- Originality and innovative aspects of the research programme: explain the contribution that the project is expected to make to advancements within the project field. Describe any novel concepts, approaches or methods that will be employed.

Explain how the high-quality, novel research is the most likely to open up the best career possibilities for the Experienced Researcher and new collaboration opportunities for the host organisation(s).

As a solid mathematical framework to model strategic decision making, game theory has proved useful in many real-world applications from economics and political science to logic, computer science and psychology. Security resource allocations and scheduling problems comprise yet another application area of critical concern, that has recently been shown to greatly benefit from game-theoretic approaches. Since 2007, the so-called ARMOR software[1] is used at the Los Angeles International Airport (LAX) to effectively determine checkpoints on the roadways leading to the airport, and to canine patrol routes within terminals. Similarly, such programs as IRIS,[2] PROTECT,[3] and TRUSTS[4] are respectively being deployed at the US Federal Air Marshals, the US coast guard patrolling, and the Los Angeles Metro system's fare inspection strategy. These methods, while being remarkably effective in their corresponding application arenas, usually rely on a pre-defined model of the environment. However, such information may in general not be available in many real-world scenarios. A key objective forming the basis of this grant proposal, is thus to design efficient and theoretically sound, data-driven methods that can actively interact with the environment to *learn* a fair model through repeated games. As discussed in the sequel, this may be achieved in an online fashion or through an exploration phase prior to the algorithm's final launch.

From a game-theoretic perspective, a security problem is viewed as a two-player game that captures the interaction between a defender (e.g., border patrols, metro inspectors, network administrators) and an attacker (e.g., terrorists/drug smugglers, illegal metro users, malicious cyber attackers). The action of the defender (attacker) is defined as selecting a subset of targets to protect (attack). For each defender/attacker action pair, utilities are defined as the players' gain or loss, and the players' objectives are to maximise their corresponding pay-offs. From the defender's perspective, this corresponds to efficiently allocating a limited number of resources to secure some predefined targets from the attacker. Solutions to such games rely on randomised strategies, making the defender's scheme highly unpredictable for the attacker, thus giving rise to a significant advantage over the original mechanisms that are based on deterministic human schedulers. In the case of games that are fully competitive between the two players (i.e. the so-called zero-sum games), these methods are provably robust in that they provide guaranteed performance against *any* possible attacker. In this case, such guarantees hold, even if the defender's strategy is completely revealed to the attacker. The extension of this guarantee to a more general (non zero-sum) game is provided by Stackelberg equilibrium, a notion that generalises the famous Nash equilibrium.[5]

[1] **pita2008deployed**.

[2] **tsai2009iris**.

[3] **shieh2012protect**.

[4] **yin2012trusts**.

[5] **korzhyk2011stackelberg**.

Related Work. Some of the main issues forming the primary focus of research in security games have been scalability, or devising strategies that take advantage of the attacker's potentially limited rationality or bounded memory.[6] Another important research goal that has been extensively addressed is to devise methods that are robust with respect to uncertainty about the environment.[7] However, little has been done to generalise the framework to a more realistic setting where the player's objective includes to actively learn the unknown environment. Achieving this goal is indeed crucial, since algorithms that make use of environmental knowledge are arguably more reliable than those merely designed to be robust against this lack of information. With this motivation, some interesting advancements have recently been made through links with optimisation and machine learning methods. These methods focus mostly on the case where the attacker's preferences are not fully known and are thus to be learned; the learning objective is achieved through a repeated a game.[8] propose analyses in terms of the number of required queries to learn the optimal defender's strategy.[9] take a Bayesian approach where, given a prior distribution, planning techniques based on Partially Observable Markov Decision Processes (POMDPs) are used to update the posterior over the adversary's preferences. The main theoretical drawback of this planning method is in that the algorithm is based on Upper Confidence Trees (UCT), which, as shown by,[10] are provably sub-optimal. Recently an extended analysis is given by[11] for the case of multiple attackers, where at each round of the game, a single attacker is chosen adversarially from a fixed, finite, set of known attackers. The latter work shows strong connections with adversarial bandit theory.

Main Goal. The purpose of this proposal is extend the effort to bring machine learning techniques to apply security games in a broad range of real world situations. Our goal is to have a theoretically sound approach by designing efficient algorithms for which we can provide finite sample analysis. Stochastic assumptions will be made when dealing with noise in the model and adversarial assumption when dealing with the adversary to make our approach both realistic and robust. One difference that we want to explore is that, contrary to the previously mentioned approaches, where the uncertainty is about the attackers' utilities, we will explore the case where the uncertainty is on the utilities of the defender. This for instance happens when we can not assess for sure the precise return of a given action (checkpoint might not stop deterministically the attacks and the probability of success needs to be determined, here learned). We can also look to different formulations of the games that corresponds to real world possibilities or requirements: we might be required to learn defence strategies that are not necessarily the best in expectation but instead also guarantee not to possess large variances in their performance. Here we plan to make connection with risk averse learning algorithm. Another possibility is that in some situation we do not want the learning process to happen during the use of the program but before hand. Then we can assume that we use of a pre launch exploration phase where we try to learned as precisely as possible the model given some budget constraint or some targeted performance guarantees. Extending the previous approaches to complex problem that involves some combinatorial structure is also important.

Objective 1 Pure exploration in Stackelberg games As explained above, it is of interest to address the case where the defender does not have the complete knowledge of the efficiency of its actions but instead actually need to learn while playing or during a preparatory phase. Here we address the question when the defender is given a preparatory phase during which he can explore his own utility and assess them though experiments. This for instance mean that he can run tests of the security in a variety of predetermined attack scenario and therefore probe his own probability of defence. The objective of this approach is to determine the best strategy during a given exploration phase and is therefore closely related to the general theory of optimisation and has been study in the discrete context of multi arm bandit as pure exploration problems.[12] This initial work has been extended in a flurry variant setting where one

[6] **tambe2012game**.

[7] **Nguyen14RO**; **aghassi2006robust**.

[8] **blum2014learning**; **letchford2009learning**.

[9] **Marecki12PR**; **qian2014online**.

[10] **munos2014bandits**.

[11] **Balcan15CR**.

[12] **Audibert10BA**.

tries to find the best(s) arms. Victor has a nice expertise in that and has participated to the extension and application of such a framework in more and more complex problem (cite my work?) and is working on extension to combinatorial bandits that would improve upon the seminal work by Chen. Taking into account the particular structure of the problem will be necessary when dealing with Stackelberg equilibrium in security games. There the function to optimise is even more complex. One first step is to relax the problem as shown in Krause et al finding the best response to a given adversary. This is known to be is NP hard problem but can be solve almost optimally be a greedy algorithm thank to a sub modularity property of the problem. This gives rise to a first objective which would be to learning optimise stochastic submodular function under a pure exploration setting. Note that I worked on similar subject with learning in submodular functions.

connections with risk averse (Cite the work of Amir Sani) maybe a separate section for this. talk about the classical cumulative regret setting also!

Objective 2 Learning more complex adversarially chosen attacker in Stakleberg The idea would be to extend the work of Balcan using more complex bandit algorithms. They use a version with k known attackers. We can assume that k is extremely large but there is some structure that permits us to use for instance combinatorial bandits.

Objective 3 Repeated Network Security Games The security issue naturally has application in graph problem that model the network of roads/ connection between computers that agents might need to secure. Therefore there has been study that apply game theory to this problems. For instance it has been used to monitor road barrage in mumbai (connection) The goal is there to put some check point on a road to stop some terrorist. Its a one shot game where you try to minimise the probability of the player to pass. Utilities are not really defined and complex here You just want to maximise the probability of catching the attacker. We are interested in a version of this game that is repeated . Everyday the same problem arises. We would minimise the cumulative regret. Therefore the defender can be adaptive and if the attacker is not smart and repeat always the same plan we will catch him often (not totally a worse case scenario). This can be seen actually has a specific problem of adversarial combinatorial bandits where the attacker is limited to a very specific structure of losses which are path in a graph. We can expect to use the specificity of the graph by using some result from spectral graph theory. Maybe also we can use this theory to solve some issue with the scalability of the algorithm.

1.2 Clarity and quality of transfer of knowledge/training for the development of the researcher in light of the research objectives

Outline how a two way transfer of knowledge will occur between the researcher and the host institution, in view of their future development and past experience: (please see Section 5.2 of this Guide):

- *Explain how the Experienced Researcher will gain new knowledge during the fellowship at the hosting organisation(s)*
- *Outline the previously acquired knowledge and skills that the researcher will transfer to the host organisation*

The overall trianing objective is to significantly develop Dr Gabillon's scientific, organisational, communication and technology transfer skills. This will enable him to continue building his portfolio of outstanding research to attain a position of independence and gain recognition in the international research community.

The proposed project is primarily a research project, and the main training objectives are to enhance the fellow's scientific skills. Dr Gabillon is already an expert in the modern theory of bandits, including best arm identification, and reinforcement learning. Therefore this project's main training objective for Dr Gabillon will be to develop his skills and knowledge in advanced statistical methods(?) and game theory. [**TODO:** What research knowledge will be learned and from who (including places you might visit)?]

Lancaster University is world-leading in industrially-inspired statistics. Learn from STOR-i and DSI. Work with SMEs from Infolab. Etc.

In addition, Dr Gabillon will be given the opportunity to:

1. Receive training on preparing funding applications by co-authoring proposals for UK and EU funding agencies with Prof. Leslie and others.
2. Gain further experience of developing industry/academic partnerships by working with Profs. Leslie and Eckley and other staff in STOR-i in technology transfer activities.
3. Attend staff training workshops designed specifically for early-career researchers, including [**TODO:** XXX].
4. Develop public communication skills by presenting research results to varied audiences.
5. Participate in the organisation of workshops in Lancaster and at the Royal Statistical Society.
6. Opportunity (but not obligation) to participate in teaching and research supervision at undergraduate and graduate level. The fellow will benefit from peer observation, mentoring, and constructive criticism.
7. Gain experience of research planning and decision-making.

Throughout the fellowship, Dr Gabillon will adhere to the "European Charter for Researchers", and the training objectives will be managed through a Personal Career Development Plan that Prof. Leslie and Dr Gabillon will write together. This plan will be revised regularly throughout the fellowship to ensure that all objectives are met. In addition, Dr Gabillon will have regular meetings with the host supervisor to discuss his research and to receive advice.

Lancaster is the leading UK institution in bandit theory, with expertise in index policies (Glazebrook, Kirkbride, Jacko), Thompson sampling and contextual bandits (Grunewalder, Leslie) and application in medical trials (Vilar). Gabillon brings expertise from another aspect of online learning and decision-making with expertise in the design and analysis of algorithmic approaches to learning, especially with combinatorial bandit problems. This will complete the portfolio of bandit research at Lancaster. Dr Gabillon's expertise in best-arm identification will be of great interest to the Medical and Pharmaceutical Statistics research group, who are exploring the use of such methods in clinical trial designs, and his expertise in combinatorial bandits complements current research of the Supervisor.

1.3 Quality of the supervision and the hosting arrangements

Required sub-heading:

Qualifications and experience of the supervisor(s)

Information regarding the supervisor(s) must include the level of experience on the research topic proposed and document its track record of work, including the main international collaborations. Information provided should include participation in projects, publications, patents and any other relevant results. To avoid duplication, the role and profile of the supervisor(s) should only be listed in the "Capacity of the Participating Organisations" tables (see section 6 below).

Hosting arrangements[13]

The text must show that the Experienced Researcher should be well integrated within the hosting organisation(s) in order that all parties gain the maximum knowledge and skills from the fellowship. The nature and the quality of the research group/environment as a whole should be outlined, together with the measures taken to integrate the researcher in the different areas of expertise, disciplines, and international networking opportunities that the host could offer.

For GF both phases should be described - for the outgoing phase, specify the practical arrangements in place to host a researcher coming from another country, and for the incoming phase specify the measures planned for the successful (re-)integration of the researcher.

Describe briefly how the host will contribute to the advancement of their career. In that context the following section of the European Charter for Researchers refers specifically to career development:

[13] The hosting arrangements refer to the integration of the Researcher to his new environment in the premises of the Host. It does not refer to the infrastructure of the Host as described in Criterion Implementation.

Qualifications and experience of the supervisor(s)

Prof. Leslie leads the Statistical Learning research group in the Department of Mathematics and Statistics, Lancaster University. He is a world-leading researcher in statistical learning, Bayesian inference, decision-making and game theory, with 19 refereed articles in top journals of several different research fields, and collaborators from France, USA and Australia. His research on contextual bandit algorithms[14] is used by many of the world's largest companies to balance exploration and exploitation in real-time website optimisation. He is expert in the mathematics of learning in games,[15] stochastic approximation,[16] and the mathematics of statistically-inspired reinforcement learning.[17] Prof. Leslie is the holder of a Google Faculty Award which funds a student to investigate multiple-action selection in bandits. Prior to his relocation to Lancaster, he was a senior lecturer in the statistics group of the School of Mathematics, University of Bristol. He continues to be co-director of the £1.5m EPSRC-funded cross-disciplinary decision-making research group at the University of Bristol, and was on the management team of the £5.5m ALADDIN project, a large strategic partnership between BAE Systems and EPSRC, involving researchers from Imperial College, Southampton, Oxford, Bristol and BAE Systems.

Prof. Leslie's mentoring approach is one of 'guided freedom' in which the mentee takes responsibility for their own research, while regular discussions ensure that dead ends are avoided and promising openings are exploited. In the 10 years since taking up a Faculty position, he has supervised 17 PhD students, 2 post-doctoral fellows, numerous MSc and undergraduate dissertations, and an undergraduate secondment from ENS Lyon.

Hosting arrangements

Dr Gabillon will be embedded within the statistical learning group which is lead by Prof. Leslie. This is a team of 5 academic staff and around 5 PhD students within the Department of Mathematics and Statistics. The Researcher will participate in weekly group meetings and benefit from advice from the senior scientists in the group, including the Supervisor, on research direction and management, personal development, workshop organisation, teaching, and other aspects of academic life. The group also has extremely strong links with both the Data Science Institute (XXX) and the STOR-i Centre for Doctoral Training (YYY). These exciting initiative will provide multiple further opportunities to develop informal mentoring relationsjips in addition to the formal process which takes place for all staff at Lancaster University.

1.4 Capacity of the researcher to reach and re-enforce a position of professional maturity in research

/em Applicants should demonstrate how their proposed research and personal experience can contribute to their professional development as an independent/mature researcher.

Please keep in mind that the fellowships will be awarded to the most talented researchers as shown by the proposed research and their track record (Curriculum Vitae, section 4), in relation to their level of experience.

[**TODO:** Victor to have a first stab]

2 Impact

2.1 Enhancing research- and innovation-related skills and working conditions to realise the potential of individuals and to provide new career perspectives

IExplain the expected impact of the planned research and training, and new competences acquired during the fellowship on the capacity to increase career prospects for the Experienced Researcher after this fellowship finishes.

Demonstrate also to what extent competences acquired during the fellowship, including any secondments will increase the impact of the researcher?s future activity on European society, including the science base and/or the economy

[14] **MayEtAl2012**.

[15] **LeslieCollins03**; **LeslieCollins05**; **LeslieCollins06**; **ChapmanEtAl2013**; **PerkinsLeslie2014**.

[16] **LeslieCollins03**; **PerkinsLeslie2012**; **PerkinsLeslie2014**.

[17] **LeslieCollins05**; **LarsenEtAl2010**.

2.2 Effectiveness of the proposed measures for communication and results dissemination

The new knowledge generated by the action should be used wherever possible to advance research, to foster innovation, and to promote the research profession to the public. Therefore develop following three points.

- Communication and public engagement strategy of the action
- Dissemination of the research results
- Exploitation of results and intellectual property rights

Concrete plans for the above must be included in the Gantt Chart (see point 3.1). The following sections of the European Charter for Researchers refer specifically to public engagement and dissemination: ????? ?

Public engagement Researchers should ensure that their research activities are made known to society at large in such a way that they can be understood by non-specialists, thereby improving the public's understanding of science. Direct engagement with the public will help researchers to better understand public interest in priorities for science and technology and also the public's concerns.

Dissemination, exploitation of results All researchers should ensure, in compliance with their contractual arrangements, that the results of their research are disseminated and exploited, e.g. communicated, transferred into other research settings or, if appropriate, commercialised. Senior researchers, in particular, are expected to take a lead in ensuring that research is fruitful and that results are either exploited commercially or made accessible to the public (or both) whenever the opportunity arises.

3 Implementation

3.1 Overall coherence and effectiveness of the work plan, including appropriateness of the allocation of tasks and resources

Describe the different work packages. The proposal should be designed in such a way to achieve the desired impact. A Gantt Chart should be included in the text listing the following:

- Work Packages titles (for EF there should be at least 1 WP);
- List of major deliverables;[18][19]
- List of major milestones;[20]
- Secondments if applicable.

The schedule should be in terms of number of months elapsed from the start of the project.

3.2 Appropriateness of the management structure and procedures, including quality management and risk management

Develop your proposal according to the following lines:

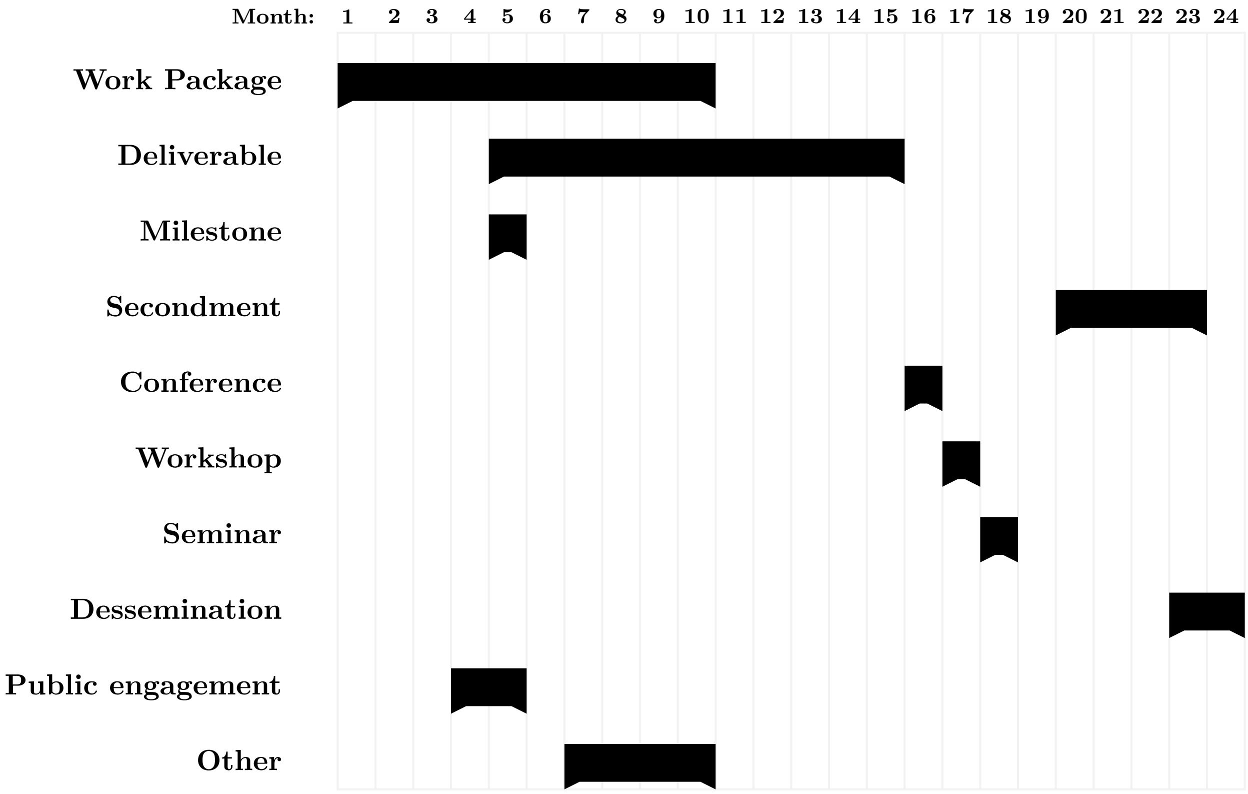
- Project organisation and management structure, including the financial management strategy, as well as the progress monitoring mechanisms put in place;
- Risks that might endanger reaching project objectives and the contingency plans to be put in place should risk occur.

[18] A deliverable is a distinct output of the action, meaningful in terms of the action?s overall objectives and may be a report, a document, a technical diagram, a software, etc.

[19] Deliverable numbers ordered according to delivery dates. Please use the numbering convention <WP number>.<number of deliverable within that WP>. For example, deliverable 4.2 would be the second deliverable from work package 4.

[20] Milestones are control points in the action that help to chart progress. Milestones may correspond to the completion of a key deliverable, allowing the next phase of the work to begin. They may also be needed at intermediary points so that, if problems have arisen, corrective measures can be taken. A milestone may be a critical decision point in the action where, for example, the researcher must decide which of several technologies to adopt for further development.

Gantt chart Reflecting work package, secondments, training events and dissemination / public engagement activities



3.3 Appropriateness of the institutional environment (infrastructure)

- Give a description of the main tasks and commitments of the beneficiary and partners (if applicable).
- Describe the infrastructure, logistics, facilities offered in as far they are necessary for the good implementation of the action.

3.4 Competences, experience and complementarity of the participating organisations and institutional commitment

The active contribution of the beneficiary to the research and training activities should be described. For GF also the role of partner organisations in Third Countries for the outgoing phase should appear. Additionally a letter of commitment shall also be provided in Section 7 (included within the PDF file of part B, but outside the page limit) for the partner organisations in Third Countries. NB: Each participant is described in Section 5. This specific information should not be repeated here.

4 CV of the Experienced Researcher

This section should be limited to maximum 5 pages and should include the standard academic and research record. Any research career gaps and/or unconventional paths should be clearly explained so that this can be fairly assessed by the independent evaluators. The Experienced Researchers must provide a list of achievements reflecting their track, and this may include, if applicable:

1. Publications in major international peer-reviewed multi-disciplinary scientific journals and/or in the leading international peer-reviewed journals, peer-reviewed conference proceedings and/or monographs of their respective research fields, indicating also the number of citations (excluding self-citations) they have attracted.

2. Granted patent(s).

3. Research monographs, chapters in collective volumes and any translations thereof.

4. Invited presentations to peer-reviewed, internationally established conferences and/or international advanced schools.

5. Research expeditions that the Experienced Researcher has led.

6. Organisation of International conferences in the field of the applicant (membership in the steering and/or programme committee).

7. Examples of participation in industrial innovation.

8. Prizes and Awards.

9. Funding received so far

10. Supervising, mentoring activities

5 Capacities of the Participating Organisations

All organisations (whether beneficiary or partner organisation) must complete the appropriate table below, which will give input on the profile of the organisation as a whole. Complete one table of maximum one page for the beneficiary and half a page per partner organisation (min font size: 9). The experts will be instructed to disregard content above this limit.

Beneficiary X	
General Description	
Role and Commitment of key persons (supervisor)	(Including names, title, qualifications of the supervisor)
Key Research Facilities, Infrastructure and Equipment	(Demonstrate that the team has sufficient facilities and infrastructure to host and/or offer a suitable environment for training and transfer of knowledge to recruited Experienced Researcher)
Independent research premises?	
Previous Involvement in Research and Training Programmes	
Current involvement in Research and Training Programmes	(Detail the EU and/or national research and training actions in which the partner is currently participating)
Relevant Publications and/or research/innovation products	(Max 5)

Partner Organisation Y	
General Description	
Key Persons and Expertise (supervisor)	
Key Research facilities, infrastructure and equipment	
Previous and Current Involvement in Research and Training Programmes	
Relevant Publications and/or research/innovation product	(Max 3)

ENDPAGE

MARIE SKLODOWSKA-CURIE ACTIONS

Individual Fellowships (IF)
Call: H2020-MSCA-IF-2014

PART B

“OSEGA”

This proposal is to be evaluated as:

[Standard EF]