

In che modo fare del bene può peggiorare accidentalmente le cose, e come fare per evitarlo

Articolo originale <https://80000hours.org/articles/accidental-harm>

Consigliamo alle persone di lavorare su problemi importanti ma trascurati e di cercare di aumentare il contributo che sono in grado di dare.

Questi passaggi rendono più facile produrre un grande impatto, ma aumentano anche le possibilità di peggiorare le cose: più il problema è importante, più è difficile rimandarlo; più un'area è trascurata, più si ha un impatto sulla sua traiettoria; e più influenza si ha, più è rilevante se si sbaglia.

Questo vale anche se non si sta facendo nulla di direttamente dannoso e si cerca di essere prudenti: è facile peggiorare le cose per caso, anzi, peggiorarle di molto.

In alcune aree della vita, i lati negativi sono relativamente limitati. Se cercate di scrivere un grande romanzo e nessuno vuole pubblicarlo, la cosa peggiore che avete fatto è perdere un po' di tempo.

Ma dimostreremo che quando si tratta di fare del bene, soprattutto in "ambiti fragili", ci sono molti modi per ostacolare l'insieme, e quindi i lati negativi non sono limitati allo stesso modo. Il potenziale di impatto negativo può essere pari o superiore a quello positivo.

Quindi, se si vuole cercare di produrre un impatto, e soprattutto se si vuole essere ambiziosi, è molto importante considerare con attenzione come si potrebbe accidentalmente peggiorare la situazione.

Questo non significa attenersi alle "cose sicure" su cui tutti sono d'accordo. Per produrre un impatto significativo, è necessario scommettere contro la saggezza convenzionale, e qualsiasi novità susciterà un certo numero di controversie. Il problema è come fare queste scommesse con abilità senza fare danni lungo il percorso.

Non ci piace particolarmente scrivere di questo argomento, che è certamente demoralizzante. Per ironia della sorte, ci aspettiamo che le persone prudenti - quelle che meno hanno bisogno di questo consiglio - siano quelle che più probabilmente lo prenderanno a cuore. Ma pensiamo che sia importante discuterne se vogliamo lavorare seriamente sui problemi più urgenti del mondo.

In questo articolo elencheremo sei modi in cui le persone possono involontariamente ostacolare la loro causa. Forse conoscete già la maggior parte di questi rischi, ma spesso vediamo persone che ne trascurano uno o due quando sono alle prime armi in un settore ad alto rischio - compresi noi quando abbiamo fondato 80,000 Hours.

Purtroppo, non abbiamo un modo per eliminare completamente questi rischi. La realtà è che bilanciare questi rischi con i potenziali vantaggi dei nuovi progetti richiede una serie di valutazioni.

Per questo motivo, concludiamo illustrando sette modi per ridurre le possibilità di peggiorare le cose, anche nelle aree col rischio maggiore. La versione più semplice consiste nell'eliminare tutte le opzioni che potrebbero avere grossi svantaggi. Ma se non potete farlo: considerate molte prospettive; non siate ottimizzatori ingenui o unilaterali; abbiate un po' di umiltà nei confronti delle vostre opinioni; costruite la vostra esperienza, migliorate la vostra capacità di giudizio e adattate la vostra esperienza alla difficoltà del progetto; evitate le decisioni difficili da revocare, come una crescita troppo rapida.

Perché sapere quando si ha un impatto negativo (controfattuale) è più difficile di quanto sembri a prima vista

Quello che ci interessa in questo articolo è la possibilità di lasciare il mondo peggiore di quello che sarebbe stato, dato quello che si sarebbe verificato se non si fosse agito.

Un modo per farlo è quello di fare una grande scommessa mal concepita che ha più probabilità di fare danni che benefici.

Purtroppo questo può accadere anche se gli effetti più diretti del vostro lavoro sono chiaramente positivi. Per esempio, potreste fare qualcosa di utile, ma nel frattempo ostacolare qualcuno che è ancora più qualificato.

Immaginate uno studente di medicina al primo anno (poco avveduto) che si imbatte in un pedone gravemente ferito in una strada trafficata, annuncia di conoscere la procedura di primo soccorso e la esegue sulla persona ferita.

Ha buone intenzioni e sembra che sappia il fatto suo. Ma immaginiamo che un passante che stava per chiamare un'ambulanza si sia astenuto perché lo studente è arrivato e ha preso il comando della situazione. In questo caso, il controfattuale potrebbe essere stato un'assistenza medica migliore da parte di un medico esperto, rendendo il suo apparente aiuto un'illusione. Lo studente di medicina avrebbe dovuto invece chiamare un'ambulanza.

Probabilmente sono poche le persone che si ostinano a compiere azioni evidentemente dannose in questo modo, ma peggiorare le cose controfattualmente in questo modo è probabilmente abbastanza comune e difficile da individuare.

Naturalmente, in questa situazione dovremmo anche pensare all'impatto di tenere libera l'ambulanza per intervenire in scenari ancora più gravi. Il che significa che misurare l'impatto reale può diventare rapidamente complicato.

Ambiti delicati e dove i rischi sono maggiori

Una delle nostre maggiori preoccupazioni in merito a questo articolo è la possibilità di scoraggiare accidentalmente le persone dall'avviare progetti di alto valore. Vogliamo quindi sottolineare che non pensiamo che i rischi di cui parleremo siano ugualmente pressanti in ogni ambito.

Ad esempio, riteniamo che la salute globale sia un problema relativamente sicuro su cui lavorare. Si tratta di un settore enorme, generalmente riconosciuto come legittimo, che gode di una reputazione consolidata e in cui il successo è spesso misurabile. (Anche se, naturalmente, è ancora possibile fallire e causare accidentalmente danni).

La riduzione del rischio di estinzione è generalmente un problema più rischioso su cui lavorare e sia la politica che la strategia dell'IA trasformativa potrebbero essere le più rischiose che suggeriamo. La legittimità di questo problema è contestata, la sua reputazione non è ancora consolidata, il successo è difficile da misurare e un obiettivo chiave è il coordinamento di gruppi con interessi contrastanti.

Più in generale, ci aspettiamo che le conseguenze indesiderate si concentrino nelle seguenti situazioni:

- Settori non consolidati che non hanno ancora una reputazione o una direzione
- Campi con cicli di feedback scarsi, in cui non si può dire se si sta avendo successo
- Campi senza un consenso di esperti, dove la natura di un problema è difficile da misurare e la gente lo giudica in base alle persone coinvolte

- Campi che implicano un conflitto con un altro gruppo che potrebbe usare i vostri errori per mettere in cattiva luce l'intero campo.
- Campi in cui la scoperta o la diffusione di informazioni sensibili può causare danni
- Campi in cui le cose stanno già andando inaspettatamente bene, dal momento che è più probabile che gli shock accidentali peggiorino le cose piuttosto che migliorarle.

Molte di queste situazioni saranno discusse in dettaglio più avanti in questo articolo.

Il resto dell'articolo si concentrerà sui rischi specifici a cui dovete prestare attenzione quando lavorate in questi ambiti fragili.

Modi per provocare un impatto negativo non intenzionale - dal più al meno evidente

Vedremo ora sei modi in cui le persone causano involontariamente danni. Tenete presente che questo post non tratta i casi in cui qualcuno ha preso la migliore decisione possibile, ma ha avuto un impatto negativo per pura sfortuna: in questo caso non c'è niente da fare. Si può anche riuscire a risolvere un problema, ma poi si scopre che la soluzione del problema era in realtà negativa, ma in questo caso ci occuperemo anche della selezione dei problemi (leggi di più).

1. Si affronta un progetto impegnativo e si commette un errore per mancanza di competenze o per scarsa capacità di giudizio.

Il mondo è estremamente complicato e la maggior parte dei progetti ha molti effetti imprevisti, che possono facilmente essere negativi. Quanto peggiore è la vostra capacità di giudizio, tanto più è probabile che ciò accada.

All'inizio della nostra storia abbiamo commesso una serie di classici errori da neo-fondatori. Abbiamo assunto troppo in fretta, abbiamo dato priorità al marketing e alle attività di sensibilizzazione prima di raggiungere il giusto rapporto tra il prodotto e il mercato (PMF: product-market fit) e ci siamo concentrati troppo poco sulla strategia a lungo termine. Fortunatamente, la maggior parte di questi errori ci ha solo rallentato, invece di creare problemi permanenti per l'eldorado nell'ambito dell'altruismo efficace.

Questa è la categoria più ovvia e evidente di visibile impatto negativo: si fa qualcosa che peggiora il problema quando qualcuno con maggiori competenze lo avrebbe previsto in anticipo.

Rientrano in questa categoria i medici che non si lavano le mani, le persone che realizzano programmi sociali dannosi (il nostro esempio di riferimento è Scared Straight) e gli studiosi che pubblicano risultati errati perché utilizzano metodi statistici sbagliati.

Una forma di errore comune tra i principianti è la mancanza di una valutazione strategica. Ad esempio, si potrebbe richiamare l'attenzione su informazioni più delicate di quanto si pensasse.

Questa è una trappola particolarmente pericolosa per chi è alle prime armi con la riduzione del rischio di estinzione. Immaginate di sapere che una nuova tecnologia potrebbe causare una catastrofe se usata impropriamente come arma. Il primo istinto potrebbe essere quello di sensibilizzare l'opinione pubblica per spingere i politici a sviluppare contromisure o una strategia di non proliferazione. Anche se questo può essere utile in alcune circostanze, aumentare l'evidenza di una minaccia può ritorcersi contro di essa facendo sì che diventi oggetto di attenzione da parte di malintenzionati (un esempio di pericolo informativo). Fare attenzione alle informazioni sensibili sembra ovvio, ma se si è nuovi in un settore spesso non è evidente quali siano le informazioni più sensibili.

Un'altra negligenza comune è quella di non rendersi conto di quanto possa essere dannoso il conflitto interpersonale e di quanto sia difficile evitarlo. I conflitti interpersonali possono danneggiare un intero settore riducendo la fiducia e la solidarietà, il che impedisce il coordinamento e rende il reclutamento molto più difficile. Nessuno vuole unirsi a un'élite in cui tutti litigano tra loro.

Un esempio dalla storia: Nel 1847 il dottor Ignaz Semmelweis si rese conto che la pulizia delle mani dei medici poteva salvare la vita dei pazienti. I suoi colleghi furono inizialmente disposti ad assecondare il suo capriccio e i tassi di infezione crollarono nel suo reparto. Ma dopo una serie di errori di comunicazione e di conflitti politici all'interno del sistema ospedaliero, Semmelweis venne considerato un pazzo e fu degradato. La pratica del lavaggio delle mani fu abbandonata e migliaia di pazienti morirono di infezione nei decenni successivi, fino a quando i ricercatori non dimostrarono che aveva ragione. Se avesse dato priorità a una comunicazione chiara delle sue idee e a migliori rapporti con i colleghi, forse l'establishment non avrebbe capito così tragicamente in ritardo che le sue idee erano corrette.

Il rischio di commettere un errore di valutazione è una buona ragione (quando possibile) per non precipitarsi a risolvere un problema complesso senza aver ricevuto la formazione, il tutoraggio, la supervisione o i consigli necessari, e per inserirsi in una comunità di colleghi che potrebbero accorgersene prima che si commetta un errore grave. La storia di Semmelweis sottolinea anche l'importanza di essere bravi a comunicare le proprie idee, non solo a svilupparle.

La maledizione dell'unilateralismo

Un modo particolarmente facile per commettere un errore che provoca un impatto negativo sostanziale è quello di agire unilateralmente in contesti in cui anche una sola persona che compie erroneamente una determinata azione potrebbe comportare costi diffusi per il proprio gruppo o per il mondo intero. Nick Bostrom ha spiegato che, se in questo contesto le persone agiscono solo in base al loro giudizio personale, le azioni rischiose verranno intraprese troppo spesso. Ha chiamato questo fenomeno "maledizione dell'unilateralista", illustrata con il seguente esempio:

Un gruppo di scienziati che lavora allo sviluppo di un vaccino contro l'HIV ha accidentalmente creato una variante dell'HIV trasmissibile per via aerea. Devono decidere se pubblicare la loro scoperta, sapendo che potrebbe essere usata per creare un'arma biologica devastante, ma anche che potrebbe aiutare coloro che sperano di sviluppare difese contro tali armi. La maggior parte dei membri del gruppo ritiene che la pubblicazione sia troppo rischiosa, ma uno non è d'accordo. Menziona la scoperta a una conferenza e ben presto i dettagli vengono resi noti a tutti.

Più in generale, supponiamo che ci sia un gruppo di dieci persone che cercano di stimare il valore atteso di un'iniziativa potenzialmente rischiosa che, se adottata, avrebbe un impatto negativo. Anche se, in media, la stima del gruppo è corretta, ci saranno alcune persone la cui stima del valore è troppo alta e altre la cui stima è troppo bassa. Se ognuno agisce solo in base al proprio giudizio, l'avvio o meno dell'iniziativa sarà determinato esclusivamente dal fatto che il membro più ottimista dell'intero gruppo, cioè quello che ha sovrastimato maggiormente il valore dell'iniziativa, pensa che sarà positiva. Questa è la ricetta per portare avanti molti progetti sbagliati.

Fortunatamente, la maledizione può essere eliminata se si tiene conto del giudizio del resto del gruppo e si evita di intraprendere azioni unilaterali quando la maggior parte di loro non è d'accordo.

2.Reputazione danneggiata

Tutti sanno che un rischio insito nel fallimento è quello di rovinare la propria reputazione. Purtroppo, però, a volte capita che i vostri errori si ripercuotano su tutto il vostro settore. Ciò significa che un errore può danneggiare anche altre persone nella vostra comunità, nel vostro settore di ricerca o nella vostra professione.

Un esempio eclatante di questo fenomeno è il crollo di FTX, il cui amministratore delegato, Sam Bankman-Fried, ha dichiarato di voler fare dell'altruismo efficace guadagnando il più possibile per donare il più possibile. È stato poi accusato di frode nel dicembre 2022.

Non solo la sua reputazione è a pezzi, ma è stata un'enorme battuta d'arresto per la reputazione di tutti i settori a cui era associato e che diceva di voler aiutare.

In generale, più grande è il fallimento, più grandi sono i danni.

Ecco perché è giusto che i progetti più grandi siano sottoposti a un controllo sempre maggiore da parte delle persone parte del progetto, e si preoccupino ancora di più di evitare cattiva pubblicità.

Questo dimostra anche che la ricerca di copertura mediatica può facilmente diventare un'arma a doppio taglio, soprattutto se si diventa famosi. È improbabile che un progetto di basso profilo riceva molta attenzione in caso di fallimento.

Inoltre, un fallimento "normale" sarà molto meno interessante di un progetto che esplode in modo controverso o interessante.

Purtroppo, questo è uno dei motivi per cui può essere costoso rendere la propria vita anti-conformista o insolita in molti modi che non sono correlati al progetto. Se fallite e si scopre, per esempio, che eravate appassionati di My Little Pony (per fare un esempio inventato), indipendentemente dal fatto che questo abbia avuto o meno a che fare con il successo del vostro progetto, la storia diventa molto più interessante, quindi attirerà più attenzione e danneggerà maggiormente la reputazione della causa che stavate cercando di sostenere.

Esistono anche molti modi più sottili per danneggiare la reputazione del vostro settore. Non sono così gravi, ma possono essere più facili da provocare, quindi vale la pena tenerli a mente.

Per esempio, immaginate di essere entusiasti di un modo totalmente nuovo di migliorare il mondo e di voler avviare una raccolta fondi per farlo decollare, quindi andate a parlare con tutti i donatori milionari che sembrano interessati. Sfortunatamente, non avete riflettuto sulle obiezioni alle vostre idee e quindi vi mostrate ingenui. Questi donatori decidono di non finanziarvi e sono anche meno propensi ad accettare incontri con chiunque altro voglia fare qualcosa di simile in futuro.

Non ci vuole molta immaginazione per pensare ad altre possibilità.

Forse il modo più sottile per danneggiare la reputazione di una piccola élite è la volontà di svolgere un lavoro molto visibile ma mediocre o poco incisivo, che dà solo un piccolo contributo diretto al problema in questione.

Purtroppo, quando esprimiamo un giudizio su qualcosa, spesso richiamiamo alla mente tutto ciò che associamo ad esso e poi ci chiediamo "quanto è valido un esempio standard?". È così che un curriculum in cui si dice che si sono pubblicati cinque articoli in riviste accademiche di alto livello può essere più impressionante di un curriculum in cui si dice che si sono pubblicati quegli articoli e se ne sono scritti dieci che non sono stati pubblicati da nessuna parte.

Annacquare un settore con lavori mediocri può anche ostacolare la sua crescita a lungo termine. La domanda è: I ricercatori che lavorano su un argomento trascurato dovrebbero pubblicare contributi di ricerca mediocri o poco incisivi oltre a quelli migliori? Condividere tutte le loro intuizioni con il mondo potrebbe far avanzare la scienza nel breve periodo. Tuttavia, i settori in cui la ricerca migliore è diluita da molti contributi marginali possono sviluppare una reputazione poco brillante, scoraggiando i laureati più promettenti.

Quando un campo è appena agli inizi, qualsiasi pubblicazione potrebbe finire per essere considerata come l'introduzione principale a quell'area tematica di una persona nuova a quel dibattito. I ricercatori farebbero meglio a mettere a punto solo le loro idee migliori, quelle che sarebbero interessanti per gli studenti promettenti che non hanno mai visto un lavoro in quell'area. Questi rischi per la reputazione sono un motivo per fare del proprio meglio e per considerare come il proprio lavoro influisce sulla percezione che gli altri hanno del vostro settore, ma ovviamente non vogliamo enfatizzarli in eccesso. Naturalmente, l'ideale non è pubblicare la propria singola idea migliore e poi andare in pensione per evitare di infangare la propria reputazione con qualcosa di inferiore.

Anche permettere che il perfezionismo e le preoccupazioni per la reputazione ostacolino i potenziali contributi positivi può essere un errore. In generale, l'impatto sullo status dell'élite è solo un fattore da considerare tra i tanti quando si compiono scelte di carriera.

Purtroppo, è difficile capire quando i rischi per la reputazione superano i benefici della condivisione di nuove idee o della promozione di un messaggio importante. Alcuni punti chiave da considerare: (i) è molto meno preoccupante quando si lavora in un campo già ampio e consolidato (ii) se il lavoro può essere sottoposto a revisione paritaria da parte di una rivista autorevole o comunque avere l'approvazione di un esperto, allora probabilmente va bene (iii) in caso di dubbio, considerare la possibilità di condividere il lavoro privatamente con un gruppo specifico prima di pubblicarlo.

3. Risorse deviate

Quasi tutti i progetti alla fine cercano finanziamenti e persone da assumere. Molti cercano anche di attirare l'attenzione della gente.

Sfortunatamente, al mondo ci sono solo un certo numero di soldi, persone e attenzione, non sono risorse infinite.

Se si assume qualcuno, questa persona non lavorerà altrove. Se accettate una donazione, quel denaro non andrà a qualcun altro. E se qualcuno sta leggendo questo articolo, non ne sta leggendo un altro.

Questo significa che un progetto che fa direttamente del bene può comunque essere controproducente se quelle risorse sarebbero andate naturalmente in un posto ancora migliore.

Questo rischio diventa più grande se i) siete un venditore insolitamente bravo, ii) esagerate il vostro impatto, iii) i donatori e i dipendenti non sono in grado di dire quali sono i progetti migliori e iv) attingete a risorse che probabilmente sarebbero utilizzate bene in vostra assenza (piuttosto che apportare risorse "nuove" che altrimenti non si sarebbero concentrate sul fare del bene).

Quando si effettua una valutazione dell'impatto del proprio progetto, è importante cercare di confrontare approssimativamente il proprio impatto con quello che sarebbe stato possibile ottenere se le risorse utilizzate fossero state destinate a un altro progetto d'impatto a cui avrebbero potuto plausibilmente essere destinate. Per esempio, i nostri lettori che si occupano di sviluppo internazionale spesso usano GiveDirectly come "base" che potrebbe assorbire molti fondi.

4. Blocco delle scelte subottimali

Quando '80.000 hours' era una novità, abbiamo promosso l'idea di "guadagnare per donare" per gli enti di beneficenza efficace, soprattutto promuovendo di lavorare nella finanza. Pensavamo che fosse una buona opzione, anche se non eravamo convinti che fosse la migliore. Ciononostante, abbiamo colto l'occasione per farci un po' di pubblicità. Questo ha fatto sì che noi, essendo la più ampia comunità dell'altruismo efficace, venivamo strettamente associati al "guadagnare per donare". Ancora oggi (nonostante le numerose dichiarazioni in cui si afferma che la maggior parte delle persone non dovrebbe guadagnare per donare), molti pensano che sia il nostro percorso più consigliato.

Come abbiamo imparato, un altro modo in cui si può causare un danno è quello di impostare il settore su una traiettoria peggiore di quella che avrebbe preso altrimenti, e poi ci si trova in difficoltà per uscirne. Questo è più probabile nelle prime fasi, quando la vostra strategia non è ancora definita, le persone non hanno un'opinione su di voi e l'ambito è piccolo, quindi le singole azioni possono cambiare significativamente la direzione del campo nel suo complesso. Non è una preoccupazione molto grande in campi più grandi e consolidati.

Le decisioni che si prendono quando si è appena agli inizi e si conosce poco possono rimanere circoscritte o addirittura andare fuori controllo. Questo per una serie di motivi:

- La maggior parte delle persone che vi conoscono vi presterà solo una minima attenzione, quindi è difficile cambiare la loro prima impressione;
- Una volta che i media hanno scritto di voi, le persone continueranno a leggere quegli articoli, dando forma alle percezioni future. In particolare, i giornalisti spesso attingono dal lavoro dei giornalisti precedenti.
- I termini sono difficili da cambiare: oggi faremmo fatica ad abbandonare il termine "altruismo efficace" anche se decidessimo che non ci piace;
- Una volta definito ciò che si crede, si tenderà ad attrarre persone che sono d'accordo con quella visione, rafforzandola ulteriormente;
- Le persone hanno molta difficoltà a cambiare i colleghi, le strutture di gestione o ad abbandonare la loro strategia, quindi le scelte sbagliate spesso continuano anche quando si sa che sono problematiche.

Questi effetti possono essere difficili da notare perché non abbiamo mai la possibilità di vedere come sarebbero andate le scelte alternative.

5. Crowding Out

Se si annuncia che si lavorerà su un particolare problema o si esprimerà un particolare

approccio, si può scoraggiare altre persone dal fare la stessa cosa, perché penseranno che si è già pensato a tutto. Per esempio, una volta che 80,000 Hours ha dichiarato che avrebbe condotto una ricerca su come fare del bene con la propria carriera, abbiamo impedito a chiunque altro nella comunità dell'altruismo efficace di avviare un progetto simile.

Questo è diventato un problema più vasto per noi, dato che ci siamo concentrati maggiormente sull'approfondimento della comprensione dei nostri percorsi prioritari e delle principali priorità globali. C'è la possibilità che la nostra esistenza possa scoraggiare le persone dal fare ricerca su carriere incentrate su altri problemi, come la salute e lo sviluppo globali.

Non vi stiamo dicendo che dovrete iniziare un nuovo progetto solo se siete certi di riuscirci. L'altro lato di questo problema è che è negativo quando persone qualificate identificano con successo una lacuna, ma non prendono l'iniziativa perché pensano che sia già stata gestita da altri, oppure aspettano che arrivi qualcuno migliore.

Stiamo solo facendo notare che annunciare l'inizio di un progetto non è privo di costi per il resto del vostro gruppo, quindi vale la pena di pensarci con dovuto rigore prima di andare avanti e potenzialmente scoraggiare gli altri. Chiedete consiglio a persone di cui vi fidate e che siano oneste nel valutare se il vostro progetto è ragionevole. Chiedete in giro se qualcun altro nel vostro gruppo ha progetti simili; forse dovrete unire i progetti, collaborare o coordinarvi su quale progetto portare avanti.

Infine, cercate di essere onesti con voi stessi sulla probabilità di portare avanti i vostri piani. Una delle forme più evitabili (e costose) di "crowding out" è quella che si verifica quando si annuncia un progetto ma non lo si porta mai a termine. Per esempio, abbiamo sentito diverse persone dire di non voler avviare un gruppo altruista locale perché ne esiste già uno, ma poi il gruppo esistente viene presto trascurato o diventa del tutto inattivo.

6. Altri problemi di coordinamento

Abbiamo scritto un articolo sui benefici sostanziali che possono derivare da grandi gruppi che cooperano in modo efficace. Ma è anche vero che chi non riesce a coordinarsi bene con altri gruppi può causare danni significativi.

I gruppi più grandi sono più difficili da coordinare rispetto a quelli più piccoli. Sia che si tratti di ricerca, di attivismo o di rapporti con gli esterni, entrare a far parte di un gruppo obbliga i vostri colleghi a investire tempo per assicurarsi che voi e loro siate in sintonia. Inoltre, gran parte del coordinamento si basa sulla fiducia, ed è difficile mantenere la fiducia in un gruppo più grande o mutevole in cui non si hanno relazioni consolidate. L'aggiunta di persone a un'area ha un impatto positivo diretto, ma crea anche un costo aggiuntivo sotto forma di maggiore coordinamento del gruppo. Questo rende l'asticella della crescita di una causa (soprattutto se piccola) più alta di quanto sembri a prima vista.

Come si possono mitigare questi rischi?

Quanto sopra può sembrare un vangelo della disperazione. Ci sono così tanti modi per peggiorare accidentalmente le cose.

Purtroppo, non possiamo darvi un modo per evitarli tutti. Dovrete usare il vostro giudizio e soppesare i potenziali vantaggi dei progetti rispetto a questi rischi. Detto questo, pensiamo che ci siano dei passi da fare oltre a tenere a mente questi potenziali aspetti negativi, cercando di anticiparli e di evitarli dove possibile. Abbiamo già menzionato alcuni di questi passi in precedenza, ma qui ne approfondiremo altri.

1. Idealmente, eliminare le azioni che potrebbero avere un forte impatto negativo

Se state confrontando diverse opzioni, e una di esse potrebbe avere un forte impatto negativo, mentre le altre forse no, la linea d'azione più semplice e sicura è quella di eliminare quella con grandi rischi di ribasso, per poi scegliere quella con il maggior potenziale tra quelle che rimangono.

Questo è anche il consiglio che diamo nel nostro articolo sull'ambizione: limitare i lati negativi, poi cercare i lati positivi.

Per "forte impatto negativo" non intendiamo solo che potreste fallire voi personalmente. Piuttosto, intendiamo dire che potreste rovinare la vostra area di ricerca (o la sua reputazione) o che potreste

lasciare le cose molto peggio di come le avete trovate.

Perché di solito ha senso eliminare queste opzioni? Le prossime due sezioni offrono alcune giustificazioni. In breve: Se pensate che un'azione possa avere sia costi che benefici elevati, ma che i benefici siano superiori ai costi, ciò dipende dalla correttezza dei dettagli della vostra stima... ma probabilmente la vostra stima non è corretta. È meglio intraprendere azioni che sembrano valide da una più ampia varietà di prospettive. Allo stesso modo, queste azioni sono spesso controverse (quindi non epistemicamente umili e forse unilaterali), comportano rischi di reputazione per l'intero campo e spesso contribuiscono alla rottura del coordinamento.

Naturalmente potreste scoprire che la maggior parte delle opzioni ha un potenziale di grandi effetti negativi o, come spesso accade, le linee d'azione con i maggiori vantaggi hanno anche i maggiori svantaggi.

In questo caso, dovrete ragionare con maggiore attenzione: le sezioni seguenti vi aiuteranno a farlo.

2. Non essere ottimizzatori ingenui

La vostra comprensione della situazione è sicuramente incompleta. Probabilmente vi mancano informazioni e considerazioni cruciali.

E questo ha un'importanza profonda. Il vostro "modello" della situazione probabilmente vi darà una risposta molto incerta su ciò che è meglio ("incognite note"). Ma poi c'è anche la possibilità che il vostro modello sia sbagliato e che stiate pensando alla situazione in modo del tutto errato, e questo potrebbe avvenire in modi che non avete nemmeno considerato ("incognite sconosciute"). Inoltre, non si sa nemmeno come ragionare su questo tipo di situazioni: non esiste un'unica teoria accettata inequivocabilmente su come prendere decisioni in condizioni di incertezza.

Quindi, se scegliete semplicemente un obiettivo che vi sembra buono e lo perseguite aggressivamente, è inevitabile che ci siano altri risultati importanti che state ignorando.

Quanto più aggressivamente si persegue il proprio obiettivo, tanto più è probabile che involontariamente si rovinino questi altri risultati e si faccia del male.

Questo vale soprattutto se gli altri risultati sono più difficili da misurare rispetto all'obiettivo principale, come di solito accade. Come si è visto nelle sezioni precedenti, ci sono molti modi indiretti e difficili da tracciare per peggiorare le cose, come influenzare la reputazione o il coordinamento; è una tentazione quella di scambiare questi risultati con quelli "concreti", come donare più denaro, far crescere l'organizzazione o ottenere una promozione.

Questo è legato alla "legge di Goodhart": quando una misura diventa un obiettivo, cessa di essere una buona misura. Questo perché probabilmente ci sono casi limite di misura che sono più facili da raggiungere rispetto all'obiettivo reale che ci interessa. (Questo è anche il motivo per cui il problema dell'allineamento dell'IA è importante).

Ad esempio, gli ospedali britannici impiegavano troppo tempo per ricoverare i pazienti, così è stata istituita una penalità per i tempi di attesa superiori alle quattro ore. In risposta, alcuni ospedali hanno chiesto alle ambulanze di guidare più lentamente, perché i lunghi viaggi accorciavano i tempi di attesa dell'ospedale.

Allo stesso modo, fare del bene è un obiettivo molto complesso e nebuloso. Quindi, se si sceglie un solo approccio e lo si persegue in modo aggressivo, si è tentati di prendere scorciatoie, o semplicemente di trascurare altri fattori importanti.

Alcuni esempi di un ottimizzatore ingenuo:

- Cercare di guadagnare il più possibile, in modo da poter donare il più possibile (ma ignorando le questioni relative alla reputazione, al carattere, alle norme cooperative; così come altri percorsi per fare del bene, come la diffusione di idee importanti e la costruzione di un capitale di carriera)
- Ignorare tutte le considerazioni tranne quella che si ritiene più importante (ad esempio, che potrebbero esserci molte generazioni future).
- Un problema per gli imprenditori è quello di essere talmente convinti della missione della propria organizzazione da prendere scorciatoie legali ed etiche pur di portarla al successo.

Come evitare di essere un ottimizzatore ingenuo?

Non esiste una soluzione ampiamente accettata al problema di come ragionare di fronte a una profonda incertezza del modello, ma ecco alcune idee che per noi hanno senso:

- Considerare più modelli della situazione: molte interpretazioni di ciò che conta, molti esiti potenziali e molte prospettive. Questo dovrebbe includere ciò che la saggezza convenzionale e altri esperti direbbero.
- Dovreste cercare attivamente le migliori argomentazioni contro il vostro approccio. In questo modo avrete le migliori possibilità di individuare le considerazioni mancanti.
- Seguite le linee d'azione che sembrano valide secondo molti modelli, o che sono molto valide da un punto di vista e approssimativamente neutre dagli altri. Al contrario, evitate le linee d'azione che sembrano folli secondo alcune prospettive ragionevoli.
- Siate ambiziosi ma non aggressivi.

Per saperne di più: L'altruismo efficace si basa sulla massimizzazione, e la massimizzazione è pericolosa.

3. Avere un certo grado di umiltà

Non solo probabilmente vi sbagliate su ciò che è meglio, ma probabilmente anche gli altri non sono d'accordo con voi.

Nel tentativo di avere un impatto positivo, c'è un compromesso molto importante da fare tra ciò che vi sembra meglio a voi e ciò che gli altri ritengono migliore.

Se si rinvia troppo al giudizio altrui, non si farà mai nulla di innovativo o di nuovo. Inoltre, a volte il "buon senso" può portarci a fare del male: per esempio, la maggior parte delle persone non sembra pensare che l'allevamento intensivo degli animali sia moralmente sbagliato, ma noi pensiamo che lo sia.

Ma, se si segue semplicemente il proprio punto di vista, c'è una buona probabilità di sbagliarsi, e questo può facilmente peggiorare le cose.

Infatti, se non si dispone di conoscenze particolari, si può sostenere che la propria ipotesi non è migliore di quella degli altri e che si dovrebbe semplicemente fare ciò che pensa la media delle persone.

La posizione in questo spettro di contrarietà è uno dei fattori più importanti che spiegano perché le persone adottano approcci diversi al fare del bene.

Ecco alcuni consigli più dettagliati su come trovare l'equilibrio:

- Fate una distinzione tra le vostre "impressioni" (ciò che vi sembra meglio) e la vostra "visione complessiva" (ciò che credete dopo aver preso in considerazione le opinioni degli altri). È importante sviluppare le proprie idee, in modo da contribuire alla saggezza collettiva sul da farsi, ma le azioni ad alto rischio dovrebbero generalmente essere intraprese sulla base della visione

complessiva.

- Date maggior peso alle vostre impressioni quando ci sono più ragioni per fidarsi delle vostre opinioni (come la competenza o i precedenti) e quando potete chiaramente indicare informazioni o valori che avete e che gli altri non hanno. Se una persona a caso non informata non è d'accordo con il vostro progetto, di solito non vale la pena di preoccuparsi.
- Raccogliete molti pareri. Un gran numero di persone non esperte può facilmente essere più accurato di un piccolo numero di esperti.
- Valutate le opinioni di qualcuno più in base ai precedenti che agli indicatori di competenza. In molti campi (ad esempio la maggior parte delle scienze sociali), gli esperti non sono molto più bravi del caso nel fare previsioni. Le persone con la migliore capacità di giudizio sono spesso persone con informazioni generali e con la giusta mentalità. In ogni caso, cercate di valutare l'affidabilità delle persone in base alla loro esperienza del dare giudizi simili.
- Valutate le opinioni degli altri in base alla loro forza, sia in termini di posta in gioco che di grado di convinzione. Se anche una persona affidabile ritiene che il vostro progetto sia "mediocre", non è un grosso problema. Se invece è convinto che sia molto dannoso, dovrete essere molto più cauti.
- Considerate la posta in gioco. In alcuni contesti, come le discussioni intellettuali private o la realizzazione di piccoli progetti, è lecito portare avanti le proprie idee stravaganti. Quanto maggiori sono i potenziali danni, tanto più è importante considerare un'ampia gamma di opinioni.
- Non affidatevi solo al vostro giudizio per le decisioni ad alto rischio. Ad esempio, se siete l'amministratore delegato di un progetto di grandi dimensioni, è importante avere un consiglio di amministrazione e cofondatori forti che controllino le vostre decisioni più importanti.

Un'ulteriore motivazione per l'umiltà: una linea d'azione che ritenete migliore quasi certamente non è così buona come credete.

Questo perché avete classificato le azioni potenziali in base alle vostre conoscenze attuali. Ma la vostra comprensione attuale è incompleta, quindi la vostra analisi contiene degli errori. Se pensate che qualcosa sia insolitamente valido, ciò potrebbe essere dovuto a un ragionamento corretto, oppure a un errore insolitamente grande nella vostra analisi. Questo è un esempio di regressione verso la media e significa che le azioni migliori sono in genere più vicine alla media di quanto sembrino.

Ecco un esempio speciale di quando affidarsi all'opinione altrui.

Prima abbiamo parlato della maledizione dell'unilateralismo, una situazione in cui se tutti i membri di un gruppo agiscono secondo il loro miglior giudizio, finiranno per intraprendere azioni troppo rischiose.

Per evitare questo problema:

- Esaminate il vostro progetto con altre persone esperte nello stesso campo.
- Se possibile, cercate di essere in sintonia sul valore del progetto.
- Se, dopo aver fatto questo, una minoranza significativa pensa che causerà un impatto negativo significativo, allora non fatelo.

Nel fare questo, è molto importante distinguere le persone che pensano che il vostro progetto sia semplicemente non importante da quelle che lo pensano attivamente dannoso.

Se gli altri pensano che il vostro progetto non sia d'impatto (e non ci hanno pensato molto), non è un buon motivo per non farlo. Ci saranno sempre persone che non sono entusiaste del vostro progetto.

Ma se molte persone pensano che possa essere dannoso, e se si va avanti lo stesso, allora si è unilateralisti.

4. Sviluppare competenze, ricevere formazione, creare una rete di contatti e beneficiare della saggezza accumulata nel settore

Quando si entra in un nuovo settore, è bene iniziare a lavorare per un'organizzazione consolidata che sia in grado di supervisionare il vostro lavoro. Imparate a conoscere il settore e assicuratevi di comprendere il punto di vista degli operatori affermati. Il tempo necessario varia molto a seconda dell'area in cui si lavora. In genere riteniamo che sia opportuno lavorare in un settore per almeno 1-3 anni prima di dedicarsi a progetti più impegnativi o condotti autonomamente, ma in settori particolarmente complicati - come quelli che richiedono molte conoscenze tecniche o studi universitari - questo può richiedere più tempo.

Questo non solo per farvi conoscere le questioni cruciali e le prospettive principali del settore, ma anche per assicurarvi di acquisire i consulenti necessari per compiere i passi che abbiamo descritto sopra.

Ci possono essere delle eccezioni. Per esempio, in un'emergenza come la COVID-19, non è stato possibile per alcuni di coloro che potevano contribuire alla risposta trascorrere un anno di formazione. Anche se la mancanza di competenze rilevanti dovrebbe essere un motivo di cautela, non pensiamo che tutti coloro che non hanno competenze preesistenti non debbano fare nulla.

In secondo luogo, se pensate di aver individuato un approccio trascurato all'interno di un settore fragile, cercate di capire perché le persone con maggiore esperienza non lo abbiano adottato. È possibile che abbiate trovato una lacuna promettente, ma forse il vostro approccio è già stato provato e ha fallito in precedenza, oppure altri lo stanno evitando a causa di un problema che non avete il contesto per vedere.

Se avete difficoltà a trovare opportunità di formazione o a sviluppare la rete necessaria, forse è meglio attenersi a problemi e metodi più sicuri invece di agire unilateralmente in un settore delicato.

Un modo particolarmente importante per accrescere la propria competenza è quello di migliorare la propria capacità di giudizio. Nel mondo del fare del bene, di solito non abbiamo risultati misurabili e questo significa che dobbiamo affidarci al giudizio per stimare i costi e i benefici di diversi percorsi. In un articolo a parte troverete delle note su come migliorare la vostra capacità di giudizio.

5. Seguire le regole della cooperazione

Con l'avanzare della carriera e l'aumento della vostra influenza o popolarità, potreste essere tentati di fare qualcosa di gravemente dannoso, disonesto o ampiamente considerato non etico "per il bene comune". Questa non è quasi mai una buona idea, e abbiamo scritto un articolo a parte sul perché.

Questo principio è solo un esempio della più ampia importanza della cooperazione. Se state lavorando per affrontare un importante problema sociale insieme ad altre persone, probabilmente otterrete molti più risultati se sarete:

- Onesti e con elevata integrità - in modo da potersi fidare l'uno dell'altro ed essere creduti, sia per dire la verità, sia nel senso che rispetterete gli accordi e le regole concordate.
- Utili - disposti ad aiutare gli altri che lavorano sul problema, anche se non si traggono immediatamente vantaggi personali.
- Dimostrerete capacità di compromesso e di scambio: siete disposti a fare cose che gli altri considerano preziose (o a evitare cose che gli altri considerano pessime) anche se non vi sembrano ottimali.
- Educati e rispettosi - in modo da non creare inutili drammi e cattivi sentimenti.
- Giudiziosi - disposto a ritirare la cooperazione dalle persone che non seguono le norme

Chi va in giro a violare queste norme non solo ostacola i propri progetti, ma contribuisce anche a una più ampia rottura della cooperazione nel settore, peggiorando le cose per tutti.

Inoltre, può ostacolare la capacità delle persone che si occupano del problema di lavorare con il resto del mondo. Per esempio, se qualcuno mente in nome della prevenzione degli allevamenti intensivi di animali, potrebbe creare una reputazione di disonestà per il settore in generale, rendendo più difficile la realizzazione della maggior parte dei progetti in futuro - un esempio di danno alla reputazione di cui abbiamo parlato in precedenza.

Naturalmente, la violazione di alcune norme - anche il farsi dei nemici - non sempre rende un'azione nel complesso una cattiva idea: si pensi ad alcuni casi di disobbedienza civile. Tuttavia, in questi casi vale la pena essere estremamente cauti. Le azioni non collaborative aumentano drasticamente la possibilità di fare molto più male che bene, e servono prove più solide per ritenere che sia una buona idea - tra cui, probabilmente, il fatto di provare prima altre linee d'azione.

6. Adattare le proprie capacità al proprio progetto e influenza

Cercate di adeguare le vostre capacità al vostro grado di influenza e alla fragilità dei problemi su cui state lavorando. Più la posta in gioco si fa alta, più è necessario sottoporre i progetti a verifica, competenza e cautela.

Se state organizzando qualche evento all'università, non avete bisogno di sottoporre i vostri piani a un'enorme quantità di controlli. Se invece state sostenendo un cambiamento importante nella politica del governo, allora dovete farlo.

Fatevi consigliare onestamente da esperti per capire se siete adatti a un progetto e se siete pronti ad affrontarlo. Continuate a ricevere feedback nel corso del progetto.

Man mano che il vostro progetto diventa più importante e di successo, diventa sempre più importante cercare consulenti e colleghi che vi tengano testa. Purtroppo spesso accade il contrario: con l'aumentare del successo, le persone saranno meno inclini a dubitare di voi e più preoccupate di criticarvi. Per evitare ciò, potrebbe essere necessario creare di proposito strutture e processi per limitare il ruolo del vostro giudizio.

Se avete già sviluppato la competenza, la formazione e la rete per mitigare questi rischi in un particolare settore, il vostro vantaggio comparativo consiste probabilmente nell'accettare i progetti rischiosi da cui gli altri dovrebbero stare alla larga. Potreste anche essere in grado di aggiungere valore fornendo tutoraggio, formazione e/o supervisione a promettenti nuovi arrivati nel vostro campo.

7. Evitare azioni difficili da invertire

Prima abbiamo parlato di come sia possibile bloccare accidentalmente le scelte non ottimali, per cui è importante fare attenzione e (a parità di condizioni) evitare azioni difficili da revocare.

Ad esempio, come abbiamo detto, una campagna mediatica che raggiungerà molte persone per la prima volta creerà una prima impressione che può essere difficile da cancellare.

Un altro esempio di azione difficilmente reversibile è la crescita troppo rapida, soprattutto se si lavora in un campo piccolo e fragile.

L'assunzione, l'aumento dei finanziamenti o l'incremento della popolarità sembrano normalmente cose concretamente positive dal punto di vista del progetto nel breve termine - essere più grandi significa avere un impatto maggiore - ma possono essere più ambigue se si considera il campo nel suo complesso e gli effetti a lungo termine.

Una crescita rapida è difficile da revocare, perché vorrebbe dire licenziare parecchie persone, per cui si rimane bloccati in un nuovo stato più grande. Più persone sono più difficili da coordinare e richiedono formazione e gestione, che spesso scarseggiano. Questo porta a un maggior numero di problemi, come l'unilateralismo, la rottura delle norme e gli errori di valutazione. La rapidità crea

maggiori rischi per la reputazione, che si sarebbero potuti evitare se si fosse proceduto più lentamente (ignorare questi danni più gravi è un altro esempio di ottimizzazione ingenua).

Naturalmente, una crescita più lenta comporta anche dei costi elevati; si tratta quindi di una questione di equilibrio. Tuttavia, la nostra sensazione è che i benefici della crescita siano di solito più tangibili dei costi, quindi è più probabile che le persone sopravvalutino il valore della crescita piuttosto che sottovalutarlo.

Conclusioni

Speriamo che questa discussione sui modi per fare del male non sia stata troppo demotivante.

Per concludere, è importante ricordare che poche persone cercano di affrontare i problemi più urgenti del mondo. E se più persone non ci provano, questi problemi non verranno risolti.

È anche importante ricordare che non esistono progetti perfetti. Anche i progetti che hanno un grande impatto di solito comportano una certa dose di controversie o disfunzioni.

Quando si è giovani, è facile pensare che, una volta entrati nel mondo degli adulti, si incontreranno persone competenti che gestiscono le cose, che "sanno quello che fanno" e che risolveranno i problemi del mondo al posto nostro. Ma per la maggior parte, troverete solo persone relativamente normali che fanno del loro meglio a fronte di molte limitazioni.

Se state leggendo questo articolo, probabilmente potete aiutarci.

Lo scopo di questo articolo non è quello di scoraggiarvi dal cercare di fare del bene, ma piuttosto di aiutarvi a essere più efficaci nel farlo.

Per affrontare i problemi del mondo, dobbiamo essere ambiziosi, ma non avventati; disposti a opporci allo status quo, ma non arroganti; guidati e determinati, ma non stronzi.

Trovare il giusto equilibrio non è sempre facile, ma riflettendo su entrambi gli aspetti possiamo cercare di fare del nostro meglio.